Tutorials on
Life Cycle Costing
and
Reliability
Engineering

Course Material

Course Instructor: Professor U Dinesh Kumar
Indian Institute of Management Bangalore

Chapter 1

Reliability Maintenance and Logistic Support - Introduction

All the business of war, and indeed all the business of life, is to endeavour to find out what you don't know from what you do.

Duke of Wellington

1.1.   INTRODUCTION

Ever since the Industrial Revolution began some 2½ centuries ago, customers have demanded *better, cheaper, faster, more for less,* through greater reliability, maintainability and supportability (RMS). As soon as people set themselves up in business to provide products for others and not just for themselves, their customers have always wanted to make sure they were not being exploited and that they were getting *value for money* and products that would be *fit for purpose.*

Today's customers are no different. All that has changed is that the companies have grown bigger, the products have become more sophisticated, complex and expensive and, the customers have become more demanding and even less trusting. As in all forms of evolution, the *Red Queen Syndrome* (Lewis, C. 1971, Matt, R., 1993) is forever present – in business, as in all things, you simply have to keep running faster to stand still. No matter how good you make something, it will never remain good enough for long.

Operators want infinite performance, at zero life-cycle cost, with 100% availability from the day they take to delivery to the day they dispose of it. It is the task of the designer/manufacturer/supplier/producer to get as near as possible to these extremes, or, at the very least, nearer than

their competitors. In many cases, however, it is not simply sufficient to tell the (potential) customer how well they have met these requirements, rather, they will be required to produce demonstrable evidence to substantiate these claims. In the following pages, we hope to provide you with the techniques and methodologies that will enable you to do this and, through practical examples, explain how they can be used.

The success of any business depends on the effectiveness of the process and the product that business produces. Every product in this world is made to perform a function and every customer/user would like her product to maintain its functionality until has fulfilled its purpose or, failing that, for as long as possible. If this can be done with the minimum of maintenance but, when there is a need for maintenance, that this can be done in the minimum time, with the minimum of disruption to the operation requiring the minimum of support and expenditure then so much the better. As the consumer's awareness of, and demand for, quality, reliability and, availability increases, so too does the pressure on industry to produce products, which meet these demands. Industries, over the years, have placed great importance on engineering excellence, although some might prefer to use the word "hubris". Many of those which have survived, however, have done so by manufacturing highly reliable products, driven by the market and the expectations of their customers.

The operational phase of complex equipment like aircraft, rockets, nuclear submarines, trains, buses, cars and computers is like an orchestra, many individuals, in many departments doing a set of interconnected activities to achieve maximum effectiveness. Behind all of these operations are certain inherent characteristics (design parameters) of the product that plays a crucial role in the overall success of the product. Three such characteristics are reliability, maintainability and supportability, together we call them RMS. All these three characteristics are crucial for any operation. Billions of dollars are spent by commercial and military operators every year as a direct consequence of the unreliability, lack of maintainability and poor supportability of the systems they are expected to operate.

Modern industrial systems consist of complex and highly sophisticated elements, but at the same time, users' expectations regarding trouble free operation is ever present and even increasing. A Boeing 777 has over 300,000 unique parts within a total of around 6 million parts (half of them are nuts, bolts and rivets). Successfully operating, maintaining and supporting such a complex system demands integrated tools, procedures and techniques. Failure to meet high reliability, maintainability and supportability can have costly and far-reaching effects. Losing the services

of airliners, such as the Boeing 747, can cost as high as $ 300,000 per day in forfeited revenue alone. Failure to dispatch a commercial flight on time or its cancellation is not only connected to the cost of correcting the failure, but also to the extra crew costs, additional passenger handling and loss of passenger revenue. Consequently, this will have an impact on the competitiveness, profitability and market share of the airline concerned. '*Aircraft on Ground*' is probably the most dreaded phrase in the commercial airlines' vocabulary. And, although the costs and implications may be different, it is no more popular with military operators. Costs per minute delay for different aircraft type are shown in Figure 1.1. Here the delay costs are attributable to labour charges, airport fees, air traffic control costs, rescheduling costs, passenger costs (food, accommodation, transport and payoffs).

Cost of Aircraft Delays

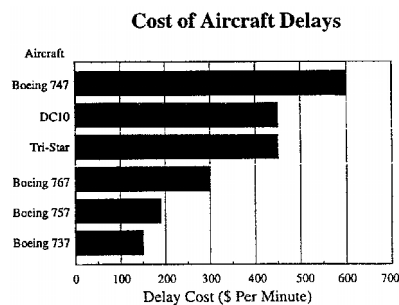| Aircraft	Delay Cost ($ Per Minute)
Boeing 747	600
DC10	430
Tri-Star	430
Boeing 767	290
Boeing 757	180
Boeing 737	140

Figure 1.1 Aircraft delay cost per minute

Industries have learned from past experience and through cutting edge research how to make their products safe and reliable. NASA, Boeing, Airbus, Lockheed Martin, Rolls-Royce, General Electric, Pratt and Whitney, and many, many more, are producing extremely reliable products. For example, over 25% of the jetliners in US have been in service for over 20 years and more than 500 over 25 years, nearing or exceeding their original design life (Lam, M., 1995). The important message is that these aircraft are still capable of maintaining their airworthiness; they are still safe and reliable. But, we cannot be complacent, even the best of organisations can have their bad days. The losses of the Challenger Space Shuttle in 1986, and Apollo 13 are still very fresh in many of our memories.
Customers' requirements generally exceed the capabilities of the producers. Occasionally, these go beyond what is practically, and sometimes even theoretically, possible. An example of this could be the

new reliability requirement, *maintenance and failure free operating period*, (Hockley *et al* 1996, Dinesh Kumar et al, 1999, 2000). High reliability is certainly a desirable function, but so to is maintainability and excellent logistic support. It is only through all three that the life-cycle cost can be driven down whilst the level of availability is driven up.

Combat aircraft are expensive and so are their crews, so no operator wants to lose either. At the same time, deploying large ground forces to maintain and support them is also expensive and, potentially hazardous. It is therefore not surprising that the operators are looking to the manufacturers to produce aircraft so reliable that they can go for weeks without any maintenance. The question is, however, can we achieve the necessary level of reliability, with sufficient confidence, at an affordable price, to meet this requirement?.

Recent projects such as the Ultra Reliable Aircraft (URA) and Future Offensive Air Systems (FOAS) place a new dimension to the reliability requirement. The operators/users would like to have Maintenance Free Operating Periods (MFOP), during which the probability that the system will need restorative maintenance is very low. Between each of these periods, sufficient maintenance will done to ensure the system will survive the next MFOP with the same probability. Only time will tell whether this policy becomes adopted but, there is no doubt that the days of the MTBF (mean time between failures) and its inverse, the [constant] failure rate are surely numbered. Science, mathematics and probability theory are slowly finding their way into the after-market business and with them will come the need for better educated people who understand these new concepts, techniques and methodologies. And, it will not just affect military aircraft, buyers of all manufactured products will demand greater value for money, at the time of purchase, of course, but more than that they will expect it throughout its life. Manufacturers who have relied on unreliability will need to re-think their policies, processes and finances.

1.2.   THE LIFE CYCLE OF A SYSTEM

Fundamental to any engineering design practice is an understanding of the cycle, which the product goes through during its life. The life cycle begins at the moment when an idea of a new system is born and finishes when the system is safely disposed. In other words, the life cycle begins with the initial identification of the needs and requirements and extends through planning, research, design, production, evaluation, operation, maintenance, support and its ultimate phase out (Figure 1.2).
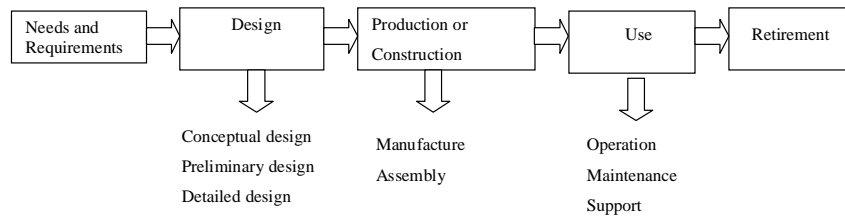
```
┌──────────────┐    ┌──────────┐    ┌──────────────┐    ┌──────────┐    ┌──────────────┐
│ Needs and    │═══▶│ Design   │═══▶│ Production or│═══▶│ Use      │═══▶│ Retirement   │
│ Requirements │    │          │    │ Construction │    │          │    │              │
└──────────────┘    └────┬─────┘    └──────┬───────┘    └────┬─────┘    └──────────────┘
                         │                 │                 │
                         ▼                 ▼                 ▼
              Conceptual design     Manufacture        Operation
              Preliminary design    Assembly           Maintenance
              Detailed design                          Support
```

Figure 1.2 Life cycle of the system.

Manufacturers who specialise in military hardware will often be approached, either directly or through an advertised "invitation to tender" to discuss the latest defence requirement. For most other manufacturers, it is generally up to them to identify a (potential) market need and decide whether they can meet that need in a profitable way. The UK MoD approached BAE Systems to bring together a consortium (including representatives of the MoD and RAF) for an air system that would out-perform all existing offensive systems, both friend and foe, and that would include all of the concepts identified as practical in the URA research project. Airbus Industries, on the other hand, decided, based on their extensive market research, that there was a sufficient market need for a very large aircraft that could carry well in excess of 500 passengers, at least across the Pacific from Tokyo to Los Angeles and possibly even non-stop between London and Sydney. It will be many years before we will know whether either of these aircraft will get off the ground and very much longer to see if they prove a business success for their manufacturers.

The first process then is a set of tasks performed to identify the needs and requirements for a new system and transform them into its technically meaningful definition. The main reason for the need of a new system could be a new function to be performed (that is there is a new market demand for a product with the specified function) or a deficiency of the present system. The deficiencies could be in the form of: 1. Functional deficiencies, 2. Inadequate performance, 3. Inadequate attributes. 4. Poor reliability, 5. High maintenance and support costs, 5. Low sales figures and hence low profits.

The first step in the conceptual design phase is to analyse the functional need or deficiency and translate it into a more specific set of qualitative and quantitative requirements. This analysis would then lead to conceptual system design alternatives. The flow of the conceptual system design process is illustrated in Figure 1.3 (D Verma and J Knezevic, 1995). The output from this stage is fed to the preliminary design stage. The

conceptual design stage is the best time for incorporating reliability, maintainability and supportability considerations. In the case of FOAS, for example, various integrated project teams with representatives of the users, suppliers and even academia will drawn together to come up with new ideas and set targets, however, impractical. It was largely a result of this activity that the concepts of the MFOP and the uninhabited combat air vehicle (UCAV) were born.
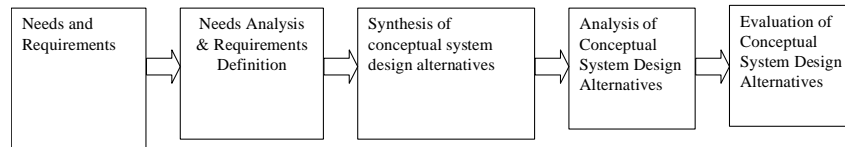

Figure 1.3 Conceptual system design process

The main tasks during the preliminary design stage are system functional analysis such as operational functions, maintenance functions, allocations of performance and effectiveness factors and the allocation of system support requirement (Blanchard, 1991). It is at this time that the concepts are brought down to earth out of the "blue sky". Groups will be required to put these ideals into reality possibly via technical development programs or abandon them until the next time.

The main tasks performed during the detailed design stage 1. Development of system/product design, 2. Development of system prototype, and 3. System prototype test and evaluation. Design is the most important and crucial stage in the product life cycle.   Reliability, maintainability and supportability depend on the design and are the main drivers of the operational availability and costs. It is during this stage that safety, reliability and maintainability demonstrations can be performed and, from these, maintenance and support plans can be decided.

The production/construction process is a set of tasks performed in order to transform the full technical definition of the new system into its physical existence.   The main tasks performed during this process are 1. Manufacture/Production/Test of prime system elements, 2. System assessment, 3. Quality Assurance, and 4. System Modification. During the production/construction process the system is physically created in accordance with the design definition. The input characteristics of the production process are the raw material, energy, equipment, facilities and other ingredients needed for the production/construction of the new

system. The output characteristics are the full physical existence of the functional system.

1.3.   CONCEPT OF FAILURE

As with so many words in the English language, *failure* has come to mean many things to many people. Essentially, a *failure of a system* is any event or collection of events that causes the system to lose its *functionability* where *functionability is the inherent characteristic of a product related to its ability to perform a specified function according to the specified requirements under the specified operating conditions.* (Knezevic 1993) Thus a system, or indeed, any component within it, can only be in one of two states: *state of functioning* or; *state of failure*.

In many cases, the transition between these states is effectively instantaneous; a windscreen shatters, a tyre punctures, a blade breaks, a transistor blows. There is insufficient time to detect the onset or prevent the consequences. However, in many other cases, the transition is gradual; a tyre or bearing wears, a crack propagates across a disc, a blade "creeps" or the performance starts to drop off. In these circumstances, some form of *health monitoring* may allow the user to take preventative measures. Inspecting the amount of tread on the tyres at regular intervals, scanning the lubricating oil for excessive debris, boroscope inspection to look for cracks or using some form trending (e.g. Kalman Filtering) on the specific fuel consumption can alert the user to imminent onset of failure. Similarly, any one of the many forms of non-destructive testing may be used (as appropriate) on components that have been exposed during the recovery of their parent component to check for damage, deterioration, erosion, corrosion or any of the other visible or physically detectable signs that might cause the component to become non-functionable.

With many highly complex systems, whose failure may have serious or catastrophic consequences, measures are taken, wherever possible, to mitigate against such events. Cars are fitted with dual braking systems, aircraft with (at least) triple hydraulic systems and numerous other instances of *redundancy*. In these cases, it is possible to have a failure of a component without a failure of the system. The recovery of the failed item, via a maintenance action, may be deferred to a time which is more convenient to the operator, safe in the knowledge that there is an acceptably high probability that the system will continue operating safely for a certain length of time. If one of the flight control computers on an aircraft fails, its functions will instantly and automatically be taken over by

one of the other computers.   The flight will generally be allowed to continue, uninterrupted to its next scheduled destination.   Depending on the level of redundancy and regulations/certification, further flights may be permitted, either until another computer fails or, the aircraft is put in for scheduled maintenance.

Most commercial airliners are fitted with two, or more, engines. Part of the certification process requires a practical demonstration that a fully loaded aircraft can take-off safely even if one of those engines fails at the most critical time; "rotation" or "weight-off-wheels". However, even though the aircraft can fly with one engine out of service, once it has landed, it would not then be permitted to take-off again until that engine has been returned to a state of functioning (except under very exceptional circumstances). With the latest large twins (e.g. Airbus 330 and Boeing 777), a change in the airworthiness rules has allowed them to fly for extended periods following the in-flight shutdown of one of the engines, generally referred to ETOPS (which officially stands for *extended twin operations over sea* or, unofficially, *engines turn or passengers swim*). This defines the maximum distance (usually expressed in minutes of flying time) the aircraft can be from a suitable landing site at any time during the flight. It also requires an aircraft that has "lost" an engine to fly to immediately divert to a landing site that is within this flying time. Again, having landed, that aircraft would not be permitted to take off until it was fitted with two functionable engines. In this case, neither engine is truly redundant but, the system (aircraft) has a limited level of fault/failure tolerance.

Most personal computers (PC) come complete with a "hard disc". During the life of the PC, it is not uncommon for small sectors of these discs to become unusable. Provided the sector did not hold the file access table (FAT) or key system's files, the computer is not only able to detect these sectors but it will mark them as unusable and avoid writing any data to them. Unfortunately, if there was already data on these sectors before they become unusable, this will no longer be accessible, although with special software, it may be possible to recover some of it. Thus, the built-in test software of the computer is able to provide a level of fault tolerance which is often totally invisible to the user, at least until the whole disc crashes or the fault affects a critical part of a program or data. Even under these circumstances, if that program or data has been backed up to another medium, it should be possible to restore the full capacity of the system usually with a level of manual intervention. So there is both fault tolerance and redundancy although the latter is usually at the discretion of the user.

Chapter 2

Probability Theory

We do not know how to predict what would happen in any given circumstances, and we believe now that it is possible, that the only thing that can be predicted is the probability of different events

Richard Feynman

Probability theory plays a leading role in modern science in spite of the fact that it was initially developed as a tool that could be used for guessing the outcome of some games of chance. Probability theory is applicable to everyday life situations where the outcome of a repeated process, experiment, test, or trial is uncertain and a prediction has to be made.

In order to apply probability to everyday engineering practice it is necessary to learn the terminology, definitions and rules of probability theory. This chapter is not intended to a rigorous treatment of all-relevant theorems and proofs. The intention is to provide an understanding of the main concepts in probability theory that can be applied to problems in reliability, maintenance and logistic support, which are discussed in the following chapters.

2.4.   PROBABILITY TERMS AND DEFINITIONS

In this section those elements essential for understanding the rudiments of elementary probability theory will be discussed and defined in a general

manner, together with illustrative examples related to engineering practice. To facilitate the discussion some relevant terms and their definitions are introduced.

Experiment

An experiment is a well-defined act or process that leads to a single well-defined outcome. Figure 2.1 illustrates the concept of random experiments. Every experiment must:

1. Be capable of being described, so that the observer knows when it occurs.
2. Have one and only one outcome, so that the set of all possible outcomes can be specified.
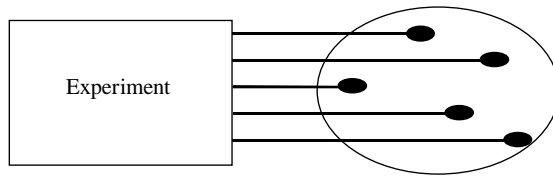
Figure 2.1 Graphical Representation of an Experiment and its outcomes.

Elementary event

An elementary event is every separate outcome of an experiment.

From the definition of an experiment, it is possible to conclude that the total number of elementary events is equal to the total number of possible outcomes, since every experiment must have only one outcome.

Sample space

The set of all possible distinct outcomes for an experiment is called the sample space for that experiment.
Most frequently in the literature the symbol *S* is used to represent the *sample space*, and small letters, *a,b,c,..*, for elementary events that are possible outcomes of the experiment under consideration. The set *S* may

contain either a finite or an infinite number of elementary events. Figure 2.2 is a graphical presentation of the sample space.
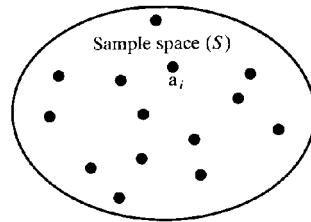

Figure 2.2 Graphical Presentation of the Sample Space

Event

Event is a subset of the sample space, that is, a collection of elementary events.

Capital letters A, B, C, ..., are usually used for denoting events. For example, if the experiment performed is measuring the speed of passing cars at a specific road junction, then the elementary event is the speed measured, whereas the sample space consists of all the different speeds one might possibly record. All speed events could be classified in, say, four different speed groups: *A* (less than 30 km/h), *B* (between 30 and 50 km/h), *C* (between 50 and 70 km/h) and *D* (above 70 km/h). If the measured speed of the passing car is, say 35 km/h, then the event *B* is said to have occurred.

2.5. ELEMENTARY THEORY OF PROBABILITY

The theory of probability is developed from axioms proposed by the Russian mathematician *Kolmogrov*. In practice this means that its elements have been defined together with several axioms which govern their relations. All other rules and relations are derived from them.

2.5.1 Axioms of Probability

In cases where the outcome of an experiment is uncertain, it is necessary to assign some measure that will indicate the chances of

occurrence of a particular event. Such a measure of events is called the *probability of the event* and symbolised by *P(.)*, (*P(A)* denotes the probability of event *A*). The function which associates each event *A* in the sample space *S*, with the probability measure *P(A)*, is called the *probability function* - the probability of that event. A graphical representation of the probability function is given in Figure 2.3.
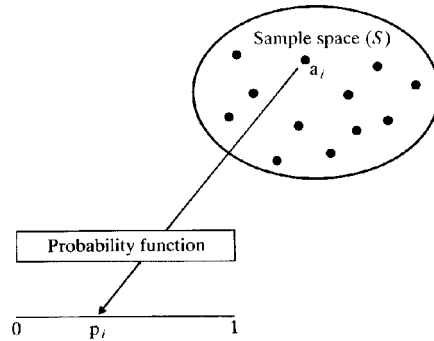

Figure 2.3 Graphical representation of probability function.

Formally, the probability function is defined as:

A function which associates with each event A, a real number, P(A), the probability of event A, such that the following axioms are true:

1. $P(A) > 0$ for every event A,
2. $P(S) = 1$, (probability of the sample space)
3. The probability of the union of mutually exclusive events is the sum of their probabilities, that is

$$P(A_1 \cup A_2 ... \cup A_n) = P(A_1) + P(A_2) + ... + P(A_n)$$

In essence, this definition states that each event A is paired with a non-negative number, probability *P(A)*, and that the probability of the sure event *S*, or *P(S)*, is always 1.

Furthermore, if A_1 and A_2 are any two mutually exclusive events (that is, the occurrence of one event implies the non-occurrence of the other) in the

sample space, the probability of their union $P(A_1 \cup A_2)$, *is simply the sum of their two probabilities,* $P(A_1) + P(A_2)$.

2.5.2   Rules of Probability

The following elementary rules of probability are directly deduced from the original three axioms, using the set theory:

a) For any event A, the probability of the complementary event, written A', is given by

$$P(A') = 1 - P(A) \tag{2.1}$$

b) The probability of any event must lie between zero and one inclusive:

$$0 \le P(A) \le 1 \tag{2.2}$$

c) The probability of an empty or impossible event, ϕ, is zero.

$$P(\phi) = 0 \tag{2.3}$$

d) If occurrence of an event A implies that an event B occurs, so that the event class A is a subset of event class B, then the probability of A is less than or equal to the probability of B:

$$P(A) \le P(B) \tag{2.4}$$

e) In order to find the probability that A or B or both occur, the probability of A, the probability of B, and also the probability that both occur must be known, thus:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{2.5}$$

f) If A and B are mutually exclusive events, so that $P(A \cap B) = 0$, then

$$P(A \cup B) = P(A) + P(B) \tag{2.6}$$

g) If n events form a partition of S, then their probabilities must add up to one:

$$P(A_1) + P(A_2) + ... + P(A_n) = \sum_{i=1}^{n} P(A_i) = 1 \tag{2.7}$$

2.5.3 Joint Events

Any event that is an intersection of two or more events is a joint event.

There is nothing to restrict any given elementary event from the sample space from qualifying for two or more events, provided that those events are not mutually exclusive. Thus, given the event A and the event B, the joint event is $A \cap B$. Since a member of $A \cap B$ must be a member of set A, and also of set B, both A and B events occur when $A \cap B$ occurs. Provided that the elements of set S are all equally likely to occur, the probability of the joint event could be found in the following way:

$$P(A \cap B) = \frac{\text{number of elementary events in } A \cap B}{\text{total number of elementary events}}$$

2.5.4 Conditional Probability

If A and B are events in a sample space which consists of a finite number of elementary events, the conditional probability of the event B given that the event A has already occurred, denoted by $P(B|A)$, is defined as:

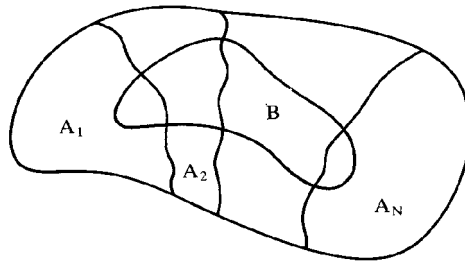$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \qquad\qquad P(A) > 0 \tag{2.8}$$

Figure 2.4 Graphical Presentation of the Bayes Theorem

The conditional probability symbol, $P(B|A)$, is read as the probability of B given A. It is necessary to satisfy the condition that $P(A)>0$, because it does not make sense to consider the probability of B given A if event A is impossible. For any two events A and B, there are two conditional probabilities that may be calculated:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \qquad and \qquad P(A|B) = \frac{P(A \cap B)}{P(B)}$$

(The probability of B, given A)   (The probability of A, given B)

One of the important application of conditional probability is due to Bayes theorem, which can be stated as follows:

If $(A_1, A_2, \backslash , A_N)$ represents the partition of the sample space (N mutually exclusive events), and if B is subset of $(A_1 \cup A_2 \cup \backslash \cup A_N)$, as illustrated in Figure 2.4, then

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1) + \backslash + P(B|A_i)P(A_i) + \backslash + P(B|A_N)P(A_N)}$$

(2.9)

2.6. PROBABILITY AND EXPERIMENTAL DATA

The classical approach to probability estimation is based on the relative frequency of the occurrence of that event. A statement of probability tells us what to expect about the relative frequency of occurrence, given that enough observations are made. In the long run, the relative frequency of occurrence of an event, say A, should approach the probability of this event, if independent trials are made at random over an indefinitely long sequence. This principle was first formulated and proved by James Bernoulli in the early eighteenth century, and is now well-known as *Bernoulli's theorem*:

If the probability of occurrence of an event A is p, and if n trials are made independently and under the same conditions, then the probability that the relative frequency of occurrence of A, (defined as $f(A) = N(A)/n$) differs

from *p* by any amount, however small, approaches zero as the number of trials grows indefinitely large. That is,

$$P(|N(A)/n) - p| > s) \to 0, \qquad\qquad as \ n \to \infty \qquad (2.10)$$

where s *is some arbitrarily small positive number. This does not mean that the proportion of* $\dfrac{N(A)}{n}$ *occurrences among any* n *trial must be* p; *the proportion actually observed might be any number between 0 and 1. Nevertheless, given more and more trials, the relative frequency of* $f(A)$ *occurrences may be expected to become closer and closer to* p.

Although it is true that the relative frequency of occurrence of any event is exactly equal to the probability of occurrence of any event only for an infinite number of independent trials, this point must not be over stressed. Even with relatively small number of trials, there is very good reason to expect the observed relative frequency to be quite close to the probability because the rate of convergence of the two is very rapid. *However, the main drawback of the relative frequency approach is that it assumes that all events are equally likely (equally probable).*

2.7.   PROBABILITY DISTRIBUTION

Consider the set of events $A_1, A_2, \setminus, A_n$, and suppose that they form a partition of the sample space *S*. That is, they are mutually exclusive and exhaustive. The corresponding set of probabilities, $P(A_1), P(A_2), \setminus, P(A_n)$, is a probability distribution. An illustrative presentation of the concept of probability distribution is shown in Figure 2.5.

As a simple example of a probability distribution, imagine a sample space of all Ford cars produced. A car selected at random is classified as a saloon or coupe or estate. The probability distribution might be:

Event	Saloon	Coupe	Estate	Total
P	0.60	0.31	0.09	1.00

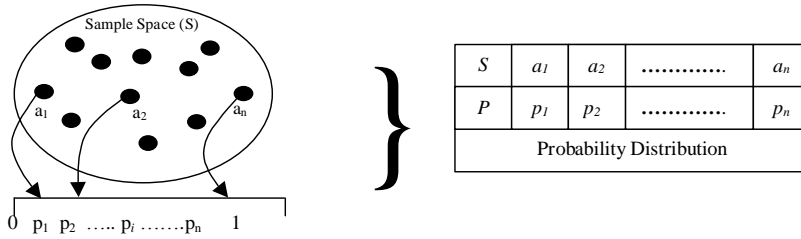All events other than those listed have probabilities of zero

Figure 2.5 Graphical representation of Probability Distribution

2.8.   RANDOM VARIABLE

A function that assigns a number (usually a real number) to each sample point in the sample space S is a random variable.

Outcomes of experiments may be expressed in numerical and non-numerical terms. In order to compare and analyse them it is much more convenient to deal with numerical terms. So, for practical applications, it is necessary to assign a numerical value to each possible elementary event in a sample space *S*. Even if the elementary events themselves are already expressed in terms of numbers, it is possible to reassign a unique real number to each elementary event. The function that achieves this is known as *the random variable*. In other words, a random variable is a real-valued function defined in a sample space. Usually it is denoted with capital letters, such as *X*, *Y* and *Z*, whereas small letters, such as *x, y, z, a, b, c*, and so on, are used to denote particular values of random variables, see Figure 2.6

If X is a random variable and r is a fixed real number, it is possible to define the event A to be the subset of S consisting of all sample points 'a' to which the random variable X assigns the number r, $A = (a : X(a) = r)$. On the other hand, the event A has a probability $p = P(A)$. The symbol p can be interpreted, generally, as the probability that the random variable X takes on the value r, $p = P(X = r)$. Thus, the symbol $P(X = r)$ represents the probability function of a random variable.
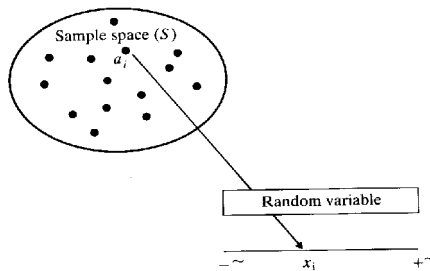
Figure 2.6 Graphical Representation of Random Variable

Therefore, by using the random variable it is possible to assign probabilities to real numbers, although the original probabilities were only defined for events of the set *S,* as shown in Figure 2.7.
The probability that the random variable X, takes value less than or equal to certain value '*x*', is called the *cumulative distribution function, F(t).* That is,
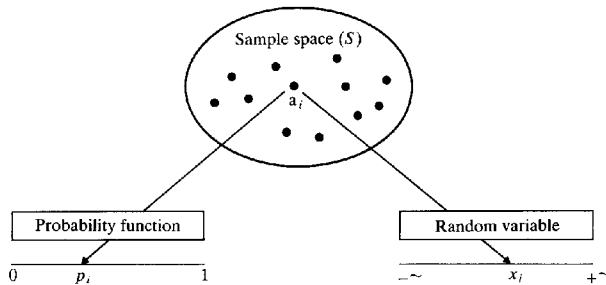
P[X ≤ x] = F(x)

Figure 2.7 Relationship between probability function and a random variable

2.8.1   Types of random variables

Depending on the values, which the random variables can assume, random variables, can be classified as discrete or continuous. The main characteristics, similarities and differences for both types will be briefly described below.

Discrete random variables

If the random variable X can assume only a particular finite or countably infinite set of values, it is said to be a discrete random variable.

There are very many situations where the random variable *X* can assume only a particular *finite* or *countably infinite* set of values; that is, the possible values of *X* are finite in number or they are infinite in number but can be put in a one-to-one correspondence with a set of real number.

Continuous random variables

 If the random variable X can assume any value from a finite or an infinite set of values, it is said to be a continuous random variable.

Let us consider an experiment, which consists of recording the temperature of a cooling liquid of an engine in the area of the thermostat at a given time. Suppose that we can measure the temperature exactly, which means that our measuring device allows us to record the temperature to any number of decimal points. If *X* is the temperature reading, it is not possible for us to specify a finite or countably infinite set of values. For example, if one of the finite set of values is 75.965, we can determine values 75.9651, 75.9652, and so on, which are also possible values of *X*. What is being demonstrated here is that the possible values of *X* consist of the set of real numbers, a set which contains an infinite (and uncountable) number of values.

Continuous random variables have enormous utility in reliability, maintenance and logistic support as the random variables time to failure, time to repair and the logistic delay time are continuous random variables.

2.9. THE PROBABILITY DISTRIBUTION OF RANDOM VARIABLE

Taking into account the concept of the probability distribution and the concept of the random variable, it could be said that the probability distribution of the random variable is a set of pairs, $\{r_i, P(X = r_i),\ i = 1, n\}$ as shown in Figure 2.8.
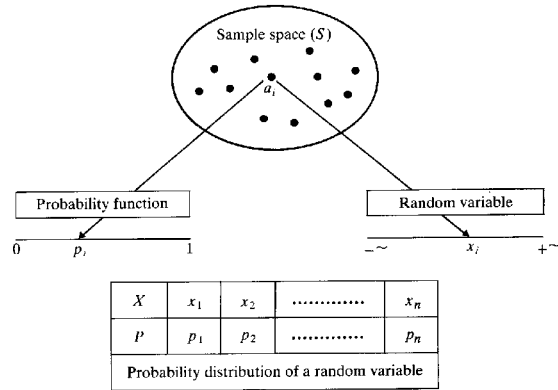
Figure 2.8 Probability Distribution of a Random Variable

The easiest way to present this set is to make a list of all its members. If the number of possible values is small, it is easy to specify a probability distribution. On the other hand, if there are a large number of possible values, a listing may become very difficult. In the extreme case where we have an infinite number of possible values (for example, all real numbers between zero and one), it is clearly impossible to make a listing. Fortunately, there are other methods that could be used for specifying a probability distribution of a random variable:

a) Functional method, where a specific mathematical functions exist from which the probability of any value or interval of values can be calculated.
b) Parametric method, where the entire distribution is represented through one or more parameters known as summary measures.

2.9.1   Functional Method

By definition, a function is a relation where each member of the domain is paired with one member of the range. In this particular case, the relation between numerical values which random variables can have and their probabilities will be considered. The most frequently used functions for the description of probability distribution of a random variable are the probability mass function, the probability density function, and the cumulative distribution function. Each of these will be analysed and defined in the remainder of this chapter.

Probability mass function

This function is related to a discrete random variable and it represents the probability that the discrete random variable, X, will take one specific value x_i, $p_i = P(X = x_i)$. Thus, a probability mass function, which is usually denoted as $PMF(.)$, places a mass of probability p_i at the point of x_i on the X-axis.

Given that a discrete random variable takes on only n different values, say $a_1, a_2, \backslash, a_n$, the corresponding $PMF(.)$ must satisfy the following two conditions:

1. $P(X = a_i) \geq 0$ *for $i = 1, 2, \backslash, n$*

2. $\sum_{i=1}^{n} P(X = a_i) = 1$

(2.11)

In practice this means that the probability of each value that X can take must be non-negative and the sum of the probabilities must be 1. Thus, a probability distribution can be represented by the set of pairs of values (a_i, p_i), where $i = 1, 2, \backslash, n$, as shown in Figure 2.9. The advantage of such a graph over a listing is the ease of comprehension and a better provision of a notion for the nature of the probability distribution.
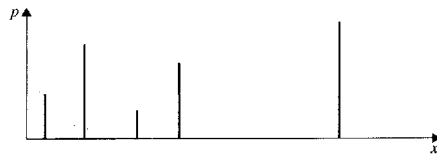

Figure 2.9 Probability Mass Function

Probability density function

In the previous section, discrete random variables were discussed in terms of probabilities $P(X = x)$, the probability that the random variables take on an *exact* value. However, consider the example of an infinite set for a specific type of car, where the volume of the fuel in the fuel tank is measured with only some degree of accuracy. What is the probability that a car selected at random will have *exactly* 16 litres of fuel? This could be considered as an event that is defined by the interval of values between, say 15.5 and 16.5, or 15.75 and 16.25, or any other interval $\pm 16 \times 0.1i$, where i is not exactly

zero.   Since the smaller the interval, the smaller the probability, the probability of exactly 16 litres is, in effect, zero.

In general, for continuous random variables, the occurrence of any exact value of X may be regarded as having zero probability.

The Probability Density Function, $f(x)$, which represents the probability that the random variable will take values within the interval $x \le X \le x + \Delta(x)$, when $\Delta(x)$ approaches zero, is defined as:

$$f(x) = \lim_{\Delta(x) \to 0} \frac{P(x \le X \le x + \Delta(x))}{\Delta x} \tag{2.12}$$

As a consequence, the probabilities of a continuous random variable can be discussed only for *intervals* of *X* values. Thus, instead of the probability that *X* takes on a specific value, say *'a'*, we deal with the so-called *probability density* of *X* at *'a'*, symbolised by $f(a)$.   In general, the probability distribution of a continuous random variable can be represented by its *Probability Density Function, PDF*, which is defined in the following way:

$$P(a \le X \le b) = \int_{a}^{b} f(x)dx \tag{2.13}$$

A fully defined probability density function must satisfy the following two requirements:

$$f(x) \ge 0 \qquad\qquad for\ all\ x$$

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

The *PDF* is always represented as a smooth curve drawn above the horizontal axis, which represents the possible values of the random variable *X.* A curve for a hypothetical distribution is shown in Figure 2.10 where the two points *a* and *b* on the horizontal axis represent limits which define an interval.
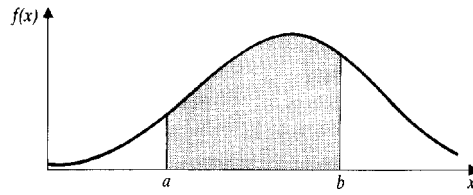
Figure 2.10 Probability Density Function for a Hypothetical Distribution

The shaded portion between *'a'* and *'b'* represents the probability that *X* takes on a value between the limits *'a'* and *'b'*.

Cumulative distribution function

The probability that a random variable *X* takes on a value at or below a given number 'a' is often written as:

$$F(a) = P(X \leq a) \tag{2.14}$$

The symbol $F(a)$ denotes the particular probability for the interval $X \leq a$. The general symbol $F(x)$ is sometimes used to represent the function relating the various values of *X* to the corresponding cumulative probabilities. This function is called the *Cumulative Distribution Function, CDF*, and it must satisfy certain mathematical properties, the most important of which are:

1. $0 \leq F(x) \leq 1$
2. *if* $a < b,$ $\quad F(a) \leq F(b)$
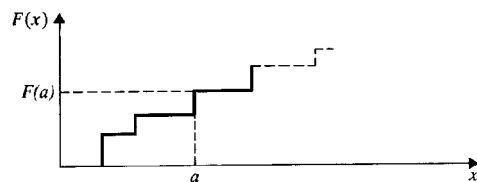3. $F(\infty) = 1$ $\quad$ *and* $\quad F(-\infty) = 0$

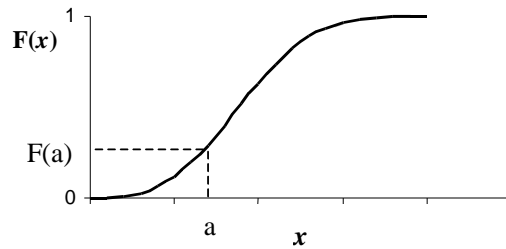Figure 2.11 Cumulative Distribution Function for Discrete Variable

Figure 2.12 Cumulative Distribution Function for Continuous Variable

The symbol $F(x)$ can be used to represent the cumulative probability that X is less than or equal to x. It is defined as:

$$F(a) = \sum_{i=1}^{n} P(X = x_i) \qquad (2.15)$$

For the discrete random variables, whereas in the case of continuous random variables it will take the following form:

$$F(a) = \int_{-\infty}^{a} f(x)dx \qquad (2.16)$$

Hypothetical cumulative distribution functions for both types of random variable are given in Figures 2.11 and 2.12.

2.9.2   Parametric Method

In some situations it is easier and even more efficient to look only at certain characteristics of distributions rather than to attempt to specify the distribution as a whole. Such characteristics summarise and numerically describe certain features for the entire distribution. Two general groups of such characteristics applicable to any type of distribution are:

a) *Measures of central tendency* (or location) which indicate the typical or the average value of the random variable.

b) *Measures of dispersion* (or variability) which show the spread of the
   difference among the possible values of the random variable.

In many cases, it is possible to adequately describe a probability distribution
with a few measures of this kind. It should be remembered, however, that
these measures serve only to summarise some important features of the
probability distribution. In general, they do not completely describe the
entire distribution.

One of the most common and useful summary measures of a probability
distribution is *the expectation* of a random variable, *E(X)*. It is a unique
value that indicates a location for the distribution as a whole (In physical
science, expected value actually represents the Centre of gravity).   The
concept of expectation plays an important role not only as a useful
measure, but also as a central concept within the theory of probability and
statistics.

If a random variable, say *X*, is discrete, then its expectation is defined as:

$$E(X) = \sum_x x \times P(X = x) \tag{2.17}$$

Where the sum is taken for all the values that the variable *X* can assume. If
the random variable is continuous, the expectation is defined as:

$$E(X) = \int_{-\infty}^{+\infty} x \times f(x)dx \tag{2.18}$$

Where the sum is taken over all values that *X* can assume. For a continuous
random variable the expectation is defined as:

$$E(X) = \int_{-\infty}^{+\infty} [1 - F(x)]dx \tag{2.19}$$

If *c* is a constant, then

$$E(cX) = c \times E(X) \tag{2.20}$$

Also, for any two random variables *X* and *Y*,

$$E(X + Y) = E(X) + E(Y)$$

Measures of central tendency

The most frequently used measures are:

The mean of a random variable is simply the expectation of the random variable under consideration. Thus, for the random variable, *X*, the mean value is defined as:

$$Mean = E(X) \tag{2.21}$$

The median, is defined as the value of X which is midway (in terms of probability) between the smallest possible value and the largest possible value. The median is the point, which divides the total area under the *PDF* into two equal parts. In other words, the probability that X is less than the median is $1/2$, and the probability that X is greater than the median is also $1/2$. Thus, if $P(X \leq a) \geq 0.50$ and $P(X \geq a) \geq 0.50$ then $'a'$ is the *median* of the distribution of *X*. In the continuous case, this can be expressed as:

$$\int_{-\infty}^{a} f(x)dx = \int_{a}^{+\infty} f(x)dx = 0.50 \tag{2.22}$$

The mode, is defined as the value of *X* at which the *PDF* of *X* reaches its highest point. If a graph of the *PMF (PDF)*, or a listing of possible values of *X* along with their probabilities is available, determination of the mode is quite simple.

A central tendency parameter, whether it is mode, median, mean, or any other measure, summarises only a certain aspect of a distribution. It is easy to find two distributions which have the same mean but which are not at all similar in any other respect.

Measures of dispersion

The mean is a good indication of the location of a random variable, but *no single value need be exactly like the mean*. A deviation from the mean, D, expresses the measure of error made by using the mean as a particular value:

$$D = x - M$$

Where, x, is a possible value of the random variable, X. The deviation can be taken from other measures of central tendency such as the median or mode. It is quite obvious that the larger such deviations are from a measure of central tendency, the more the individual values differ from each other, and the more apparent the spread within the distribution becomes. Consequently, it is necessary to find a measure that will reflect the spread, or variability, of individual values.

The expectation of the deviation about the mean as a measure of variability, $E(X - M)$, will not work because the expected deviation from the mean must be zero for obvious reasons. The solution is to find the *square* of each deviation from the mean, and then to find the expectation of the squared deviation. This characteristic is known as a *variance of the distribution, V*, thus:

$$V(X) = E(X - Mean)^2 = \sum (X - Mean)^2 \times P(x) \qquad \text{if X is discrete} \quad (2.23)$$

$$V(X) = E(X - Mean)^2 = \int_{-\infty}^{+\infty} (X - Mean)^2 \times f(x)dx \quad \text{if X is continuous} \quad (2.24)$$

The positive square root of the variance for a distribution is called the *Standard Deviation, SD*.

$$SD = \sqrt{V(X)} \qquad\qquad\qquad\qquad\qquad\qquad (2.25)$$

Probability distributions can be analysed in greater depth by introducing other summary measures, known as *moments*. Very simply these are expectations of different powers of the random variable. More information about them can be found in texts on probability.

X	x_1 x_2 .... x_n
P	p_1 p_2 .... p_n

Probability distribution of a random variable
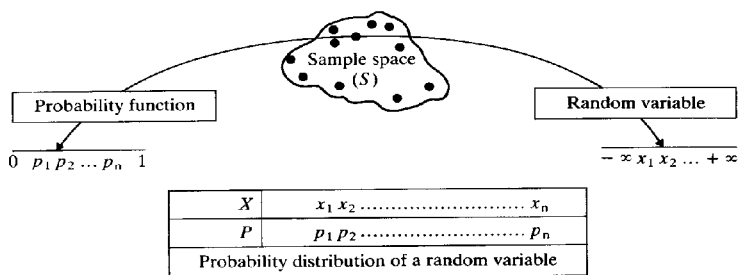
	Discrete random variable	Continuous random variable
Functional presentation	$p_i = P(X = x_i), \quad i = 1,2,3 \dots n$	$P(a < X < b) = \int_a^b f(x)dx$
	Probability mass function (*PMF*)	Probability density function (*PDF*)
	$F(a) = P(X \le a) = \sum_{i=1} P(X = x_i)$	$F(a) = P(X \le a) = \int_{-\infty}^a f(x)dx$
	Cumulative distribution function (*CDF*)	
Summary measure	$M = E(X) = \sum_x x \times P(X = x)$	$M = E(X) = \int_{-\infty}^{+\infty} x \times f(x)dx$
	Measure of central tendency (location) – mean value	
	$V = E(X - M)^2 = \sum (x - M)^2 \times P(x)$	$V = E(X - M)^2 = \int_{-\infty}^{+\infty} (x - M)^2 \times f(x)dx$
	Measure of dispersion (spread) – variance	

Figure 2.13 Probability System for Continuous Random Variable

Variability

The standard deviation is a measure that shows how closely the values of random variables are concentrated around the mean. Sometimes it is difficult to use only knowledge of the standard deviation, to decide whether the dispersion is considerably large or small, because this will depend on the mean value. In this case the parameter known as coefficient of variation, CV_X, defined as

$$CV_X = \frac{SD}{M}$$
(2.26)

Coefficient of variation is very useful because it gives better information regarding the dispersion. The concept thus discussed so far is summarised in Figure 2.13.

In conclusion it could be said that the probability system is wholly abstract and axiomatic. Consequently, every fully defined probability problem has a unique solution.

2.10.   DISCRETE THEORETICAL PROBABILITY DISTRIBUTIONS

In probability theory, there are several rules that define the functional relationship between the possible values of random variable *X* and their probabilities, *P(X)*. As they are purely theoretical, i.e. they do not exist in reality, they are called *theoretical probability distributions*. Instead of analysing the ways in which these rules have been derived, the analysis in this chapter concentrates on their properties. It is necessary to emphasise that all theoretical distributions represent the family of distributions defined by a common rule through unspecified constants known as *parameters of distribution*. The particular member of the family is defined by fixing numerical values for the parameters, which define the distribution.

The probability distributions most frequently used in reliability, maintenance and the logistic support are examined in this chapter.

Among the family of theoretical probability distributions that are related to discrete random variables, the Binomial distribution and the Poisson distribution are relevant to the objectives set by this book. A brief description of each now follows.

2.10.1   Bernuolli Trials

The simple probability distribution is one with only two event classes. For example, a car is tested and one of two events, pass or fail, must occur, each with some probability. The type of experiment consisting of series of independent trials, each of which can eventuate in only one of two outcomes are known as *Bernuolli Trials*, and the two event classes and their associated probabilities a *Bernuolli Process.* In general, one of the two events is called a "success" and the other a "failure" or "nonsuccess". These names serve only to tell the events apart, and are not meant to bear any connotation of "goodness" of the event. The symbol *p,* stands for the probability of a success, *q* for the probability of failure *(p + q =1).* If 5 independent trials are made *(n = 5),* then 2^5 = 32 different sequences of possible outcomes would be observed.

The probability of given sequences depends upon *p* and *q,* the probability of the two events. Fortunately, since trials are independent, it is possible to compute the probability of any sequence.

If all possible sequences and their probabilities, are written down the following fact emerges: *The probability of any given sequences of n independent Bernuolli Trials depends only on the number of successes and p.* This is regardless of the order in which successes and failure occur in sequence, the probability is

$$p^r q^{n-r}$$

where r is the number of successes, and $n - r$ is the number of failures. Suppose that in a sequence of 10 trials, exactly 4 success occurs. Then the probability of that particular sequence is $p^4 q^6$. If $p = \frac{2}{3}$, then the probability can worked out from:

$$\left(\frac{2}{3}\right)^4 \left(\frac{1}{3}\right)^6$$

The same procedure would be followed for any *r* successes out of *n* trials for any *p*. Generalising this idea for any *r*, *n*, and *p*, we have the following principle:

In sampling from the Bernuolli Process with the probability of a success equal to p, the probability of observing exactly r successes in n independent trials is:

$$P(r \ successes | n, p) = \binom{n}{r} p^r q^{n-r} = \frac{n!}{r!(n-r)!} p^r q^{n-r} \qquad (2.27)$$

2.10.2   The Binomial Distribution

The theoretical probability distribution, which pairs the number of successes in n trials with its probability, is called the binominal distribution.

This probability distribution is related to experiments, which consist of a series of independent trials, each of which can result in only one of two outcomes: success and or failure. These names are used only to tell the events apart. By convention the symbol *p* stands for the probability of a success, *q* for the probability of failure $(p + q = 1)$.

The number of successes, *x* in *n* trials is a discrete random variable which can take on only the whole values from 0 through *n*. The PMF of the Binomial distribution is given by:

$$PMF(x) = P(X = x) = \binom{n}{x} p^x q^{n-x}, \qquad\qquad 0 < x < n \qquad (2.28)$$

where:

$$\binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x} \qquad\qquad (2.29)$$

The binomial distribution expressed in cumulative form, representing the probability that *X* falls at or below a certain value '*a*' is defined by the following equation:

$$P(X \leq a) = \sum_{i=o}^{a} P(X = x_i) = \sum_{i=0}^{a} \binom{n}{i} p^i q^{n-i} \qquad (2.30)$$

As an illustration of the binomial distribution, the PMF and CDF are shown in Figure 2.14 with parameters $n = 10$ and $p = 0.3$.
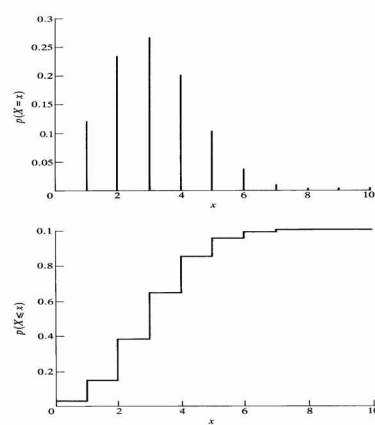

Figure 2.14 PMF and CDF For Binomial Distribution, $n = 10, p = 0.3$

$$E(X) = np \qquad (2.31)$$

Similarly, because of the independence of trials, the variance of the binomial distribution is the sum of the variances of the individual trials, or $p(1 - p)$ summed n times:

$$V(X) = np(1 - p) = npq \qquad (2.32)$$

Consequently, the standard deviation is equal to:

$$Sd(X) = \sqrt{npq} \qquad\qquad\qquad (2.33)$$

Although the mathematical rule for the binomial distribution is the same regardless of the particular values which parameters n and p take, the shape of the probability mass function and the cumulative distribution function will depend upon them. The *PMF* of the binomial distribution is symmetric if $p = 0.5$, positively skewed if $p < 0.5$, and negatively skewed if $p > 0.5$.

2.10.3   The Poisson Distribution

The theoretical probability distribution which pairs the number of occurrences of an event in a given time period with its probability is called the Poisson distribution. There are experiments where it is not possible to observe a finite sequence of trials. Instead, observations take place over a continuum, such as time. For example, if the number of cars arriving at a specific junction in a given period of time is observed, say for one minute, it is difficult to think of this situation in terms of finite trials. If the number of binomial trials n, is made larger and larger and p smaller and smaller in such a way that np remains constant, then the probability distribution of the number of occurrences of the random variable approaches the Poisson distribution.

The probability mass function in the case of the Poisson distribution for random variable X can be expressed as follows:

$$P(X = x|\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!} \qquad \text{where } x = 0, 1, 2, \dots \qquad (2.34)$$

λ is the *intensity of the process* and represents the expected number of occurrences in a time period of length t. Figure 2.15 shows the *PMF* of the Poisson distribution with $\lambda = 5$
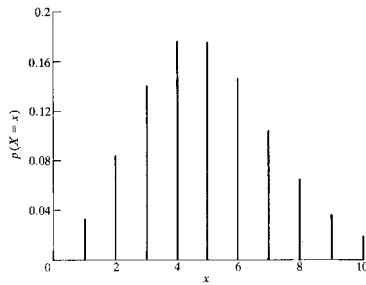
Figure 2.15 PMF of the Poisson Distribution with $\lambda = 5$

The Cumulative Distribution Function for the Poisson distribution

$$F(x) = P(X \leq x) = \sum_{i=o}^{x} \frac{e^{i} \lambda^{x}}{i!} \qquad (2.35)$$

The CDF of the Poisson distribution with $\lambda = 5$ is presented in Figure 2.16.

Expected value of the distribution is given by

$$E(X) = \sum_{x=0} xP(X = x) = \sum_{x=0} x \frac{e^{-\lambda} \lambda^{x}}{x!}$$

Applying some simple mathematical transformations it can be proved that:

$$E(X) = \lambda \qquad (2.36)$$

which means that the expected number of occurrences in a period of time *t* is equal to *np*, which is equal to λ .

The variance of the Poisson distribution is equal to the mean:

$$V(X) = \lambda \qquad (2.37)$$

Thus, the Poisson distribution is a single parameter distribution because it is completely defined by the parameter λ. In general, the Poisson

distribution is positively skewed, although it is nearly symmetrical as λ becomes larger.
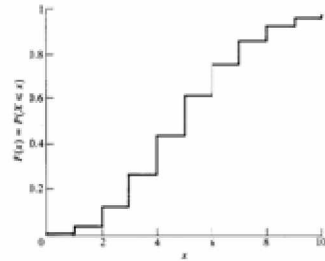
Figure 2.16 CDF of the Poisson Distribution $\lambda = 5$

The Poisson distribution can be derived as a limiting form of the binomial if the following three assumptions were simultaneously satisfied:

1. n becomes large (that is, $n \rightarrow \infty$).
2. p becomes small (that is, $p \rightarrow 0$).
3. np remains constant.

Under these conditions, the binomial distribution with the parameters n and p, can be approximated to the Poisson distribution with parameter $\lambda = np$. This means that the Poisson distribution provides a good approximation to the binomial distribution if p is very small and n is large. Since p and q can be interchanged by simply interchanging the definitions of success and failure, the Poisson distribution is also a good approximation when p is close to one and n is large.

As an example of the use of the Poisson distribution as an approximation to the binomial distribution, the case in which n = 10 and p = 0.10 will be considered. The Poisson parameter for the approximation is then $\lambda = np = 10 \times 0.10 = 1$. The binomial distribution and the Poisson approximation are shown in Table 2.2.

The two distributions agree reasonably well. If more precision is desired, a possible rule of thumb is that the Poisson is a good approximation to the binomial if $n / p > 500$ (this should give accuracy to at least two decimal places).

Table 2.2 Poisson Distribution as an Approximation to the Binomial Distribution

	Binomial $P(X = x\|n = 10, p = 0.1)$	Poisson $P(X = x\|\lambda = 1)$
0	0.598737	0.606531
1	0.315125	0.303265
2	0.074635	0.075816
3	0.010475	0.012636
4	0.000965	0.001580
5	0.000061	0.000158

2.11. CONTINUOUS THEORETICAL PROBABILITY DISTRIBUTIONS

It is necessary to emphasise that all theoretical distributions represent the family of distributions defined by a common rule through unspecified constants known as *parameters of distribution*. The particular member of the family is defined by fixing numerical values for the parameters, which define the distribution. The probability distributions most frequently used in reliability, maintainability and supportability engineering are examined in this chapter. Each of the above mentioned rules define a family of distribution functions. Each member of the family is defined with a few parameters, which in their own way control the distribution. Parameters of a distribution can be classified in the following three categories (note that not all distributions will have all the three parameters, many distributions may have either one or two parameters):

1. *Scale parameter*, which controls the range of the distribution on the horizontal scale.
2. *Shape parameter*, which controls the shape of the distribution curves.
3. *Source parameter or Location parameter*, which defines the origin or the minimum value which random variable, can have. Location parameter also refers to the point on horizontal axis where the distribution is located.

Thus, individual members of a specific family of the probability distribution are defined by fixing numerical values for the above parameters.

2.11.1   Exponential Distribution

Exponential distribution is fully defined by a single one parameter that governs the scale of the distribution. The probability density function of the exponential distribution is given by:

$$f(x) = \lambda \exp(-\lambda x), x > 0 \tag{2.38}$$

In Figure 2.17 several graphs are shown of exponential density functions with different values of λ. Notice that the exponential distribution is positively skewed, with the mode occurring at the smallest possible value,
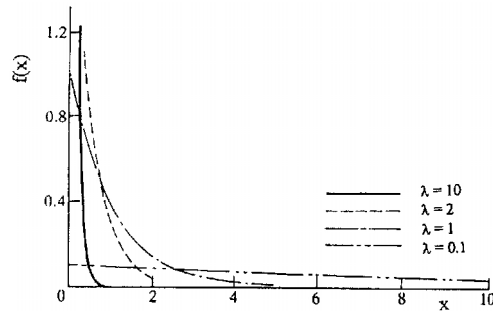
zero.

Figure 2.17. Probability density function of exponential distribution for different values of λ

The cumulative distribution of exponential distribution is given by:

$$F(x) = P(X < x) = 1 - \exp(-(\lambda x)) \tag{2.39}$$

It can be shown that the mean and variance of the exponential distribution are:

$$E(X) = 1/\lambda \tag{2.40}$$

$$V(X) = (1/\lambda)^2 \tag{2.41}$$

The standard deviation in the case of the exponential distribution rule has a numerical value identical to the mean and the scale parameter, $SD(X) = E(X) = 1/\lambda$.

11.1.1   Memory-less Property of Exponential Distribution

One of the unique property of exponential distribution is that it is the only continuous distribution that has *memory less* property. Suppose that the random variable X measures the duration of time until the occurrence of failure of an item and that it is known that X has an exponential distribution with parameter λ. Suppose the present age of the item is t, that is X > t. Assume that we are interested in finding the probability that this item will not fail for another *s* units of time. This can be expressed using the conditional probability as:

$$P\{X > s + t \,|\, x > t\}$$

Using conditional probability of events, the above probability can be written as:

$$P\{X > s + t \,|\, X > s\} = \frac{P\{X > s + t \cap X > t\}}{P\{X > t\}} = \frac{P\{X > s + t\}}{P\{X > t\}} \quad (2.42)$$

However we know that for exponential distribution

$$P[X > s + t] = \exp(-\lambda(s + t)) \text{ and } P[X > t] = \exp(-\lambda t)$$

Substituting these expressions in equation (2.42), we get

$$P[X > s + t \,|\, X > t] = P[X > s] = \exp(-\lambda s)$$

That is, the conditional probability depends only on the remaining duration and is independent of the current age of the item. *This property is exploited to a great extend in reliability theory.*

2.11.2   Normal Distribution (Gaussian Distribution)

This is the most frequently used and most extensively covered theoretical distribution in the literature. The Normal Distribution is continuous for all values of *X* between $-\infty$ and $+\infty$. It has a characteristic symmetrical

shape, which means that the mean, the median and the mode have the same numerical value. The mathematical expression for its probability density function is as follows:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \qquad (2.43)$$

Where μ is a location parameter (as it locates the distribution on the horizontal axis) and σ is a scale parameter (as it controls the range of the distribution). μ and σ also represents the mean and the standard deviation of this distribution.

The influence of the parameter μ on the location of the distribution on the horizontal axis is shown in Figure 2.18, where the values for parameter σ are constant.

As the deviation of x from the location parameter μ is entered as a squared quantity, *two* different x values, showing the same absolute deviation from μ, will have the same probability density according to this rule. This dictates the symmetry of the normal distribution. Parameter μ can be any finite number, while σ can be any *positive* finite number.

The cumulative distribution function for the normal distribution is:

$$F(a) = P(X \le a) = \int_{-\infty}^{a} f(x)dx$$

where $f(x)$ is the normal density function. Taking into account Eq. (2.43) this becomes:

$$F(a) = \int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{a-\mu}{\sigma}\right)^2\right)dx \qquad (2.44)$$
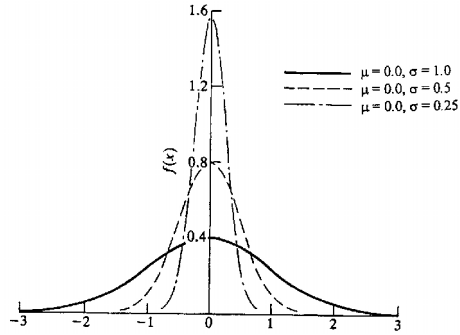
Figure 2.18 Probability density of normal distribution for different σ values

In Figure 2.19 several cumulative distribution functions are given of the Normal Distribution, corresponding to different values of μ *and* σ .

As the integral in Eq. (2.44) cannot be evaluated in a closed form, statisticians have constructed the table of probabilities, which complies with the normal rule for the standardised random variable, Z. This is a theoretical random variable with parameters μ = 0 and σ = 1. The relationship between standardised random variable Z and random variable X is established by the following expression:
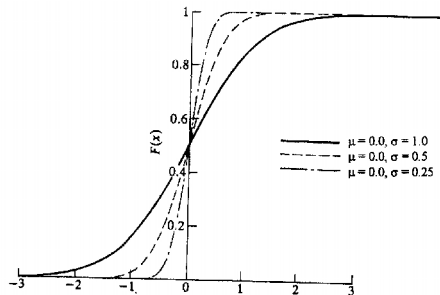

Figure 2.19 Cumulative distribution of normal distribution for different values of μ and σ.

$$z = \frac{x - \mu}{\sigma} \qquad\qquad (2.45)$$

Making use of the above expression the equation (2.43) becomes simpler:

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad (2.46)$$

The standardised form of the distribution makes it possible to use only one table for the determination of *PDF* for any normal distribution, regardless of its particular parameters (see Table in appendix).

The relationship between *f(x)* and *f(z)* is :

$$f(x) = \frac{f(z)}{\sigma} \qquad (2.47)$$

By substituting $\frac{x-\mu}{\sigma}$ with z Eq. (2.44) becomes:

$$F(a) = \int_{-\infty}^{z} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right) dz = \Phi\left(\frac{x-\mu}{\sigma}\right) \qquad (2.48)$$

where Φ is the standard normal distribution Function defined by

$$\Phi(z) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right) dx \qquad (2.49)$$

The corresponding standard normal probability density function is:

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \qquad (2.50)$$

Most tables of the normal distribution give the cumulative probabilities for various *standardised* values. That is, for a given *z* value the table provides the cumulative probability up to, and including, that standardised value in a normal distribution. In *Microsoft EXCEL®*, the cumulative distribution

function and density function of normal distribution with mean μ and standard deviation σ can be found using the following function.

 $F(x) = NORMDIST\ (x,\ \mu,\ \sigma,\ TRUE),\ and\ f(x) = NORMDIST\ (x,\ \mu,\ \sigma,\ FALSE)$

The expectation of a random variable, is equal to the location parameter μ thus:

$$E(X) = \mu \tag{2.51}$$

Whereas the variance is

$$V(X) = \sigma^2 \tag{2.52}$$

Since normal distribution is a symmetrical about its mean, the area between μ - $k\sigma$, μ + $k\sigma$ (k is any real number) takes a unique value, which is shown in Figure 2.20.
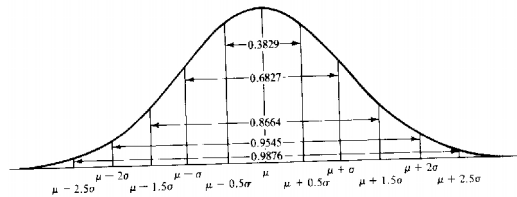

Figure 2.20 The areas under a normal distribution between
μ - $k\sigma$ and μ + $k\sigma$

11.2.1   Central Limit Theorem

Suppose X_1, X_2, ... X_n are mutually independent observations on a random variable X having a well-defined mean μ_x and standard deviation σ_x. Let

$$Z_n = \frac{\overline{X} - \mu_x}{\sigma_x / n} \tag{2.53}$$

Where,

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{2.54}$$

and $F_{z_n}(z)$ be the cumulative distribution function of the random variable Z_n. Then for all z, $-\infty < z < \infty$,

$$\lim_{n \to \infty} F_{Z_n}(z) = F_Z(z) \tag{2.55}$$

where $F_Z(z)$ is the cumulative distribution of standard normal distribution *N(0,1)*. The X values have to be from the same distribution but the remarkable feature is that this distribution does not have to be normal, it can be uniform, exponential, beta, gamma, Weibull or even an unknown one.

2.11.3 Lognormal Distribution

The lognormal probability distribution, can in some respects, be considered as a special case of the normal distribution because of the derivation of its probability function. If a random variable $Y = \ln X$ is normally distributed then, the random variable *X* follows the lognormal distribution. Thus, the probability density function for a random variable *X* is defined as:

$$f_X(x) = \frac{1}{x\sigma_l \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\ln x - \mu_l}{\sigma_l} \right)^2 \right) \quad \geq 0 \tag{2.56}$$

The parameter μ_l is called the *scale parameter* (see Figure 2.21) and parameter σ_l is called the *shape parameter*. The relationship between parameters μ (location parameter of the normal distribution) and μ_l is defined:

$$\mu = \exp\left(\mu_l + \frac{1}{2}\sigma_l^2 \right) \tag{2.57}$$

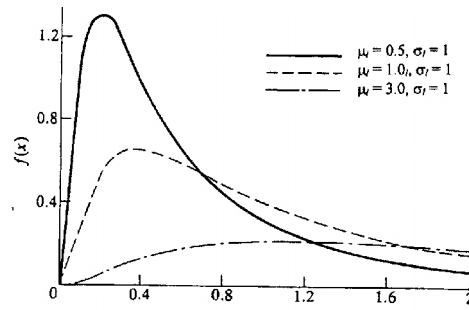The cumulative distribution function for the lognormal distribution is defined with the following expression:

Figure 2.21 Probability density of log-normal distribution

$$F_X(x) = P(X \leq x) = \int\limits_0^x \frac{1}{x\sigma_l \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln x - \mu_l}{\sigma_l}\right)^2\right) dx \quad (2.58)$$

As the integral cannot be evaluated in close form, the same procedure is applied as in the case of normal distribution. Then, making use of the standardised random variable Equation (2.61) transforms into:

$$F_X(x) = P(X \leq x) = \Phi\left(\frac{\ln x - \mu_l}{\sigma_l}\right) \quad (2.59)$$

The measures of central tendency in the case of lognormal distributions are defined by the:

(a) Location parameter (Mean)

$$M = E(X) = \exp\left(\mu_l + \frac{1}{2}\sigma_l^2\right) \quad (2.60)$$

(b) Deviation parameter (the variance)

$$V(X) = \exp\left(2\mu_l + \sigma_l\right)^2 \left[\exp(\sigma_l^2 - 1)\right] \tag{2.61}$$

2.11.4   Weibull Distribution

This distribution originated from the experimentally observed variations in the yield strength of Bofors steel, the size distribution of fly ash, fibre strength of Indian cotton, and the fatigue life of a *St*-37 steel by the Swedish engineer W.Weibull. As the Weibull distribution has no characteristic shape, such as the normal distribution, it has a very important role in the statistical analysis of experimental data. The shape of this distribution is governed by its parameter.

The rule for the probability density function of the Weibull distribution is:

$$f(x) = \frac{\beta}{\eta}\left(\frac{x-\gamma}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{x-\gamma}{\eta}\right)^{\beta}\right] \tag{2.65}$$

where $\eta, \beta, \gamma > 0$. As the location parameter ν is often set equal to zero, in such cases:

$$f(x) = \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\eta}\right)^{\beta}\right] \tag{2.66}$$

By altering the shape parameter β, the Weibull distribution takes different shapes. For example, when $\beta = 3.4$ the Weibull approximates to the normal distribution; when $\beta = 1$, it is identical to the exponential distribution. Figure 2.22 shows the Weibull probability density function for selected parameter values.

The cumulative distribution functions for the Weibull distribution is:
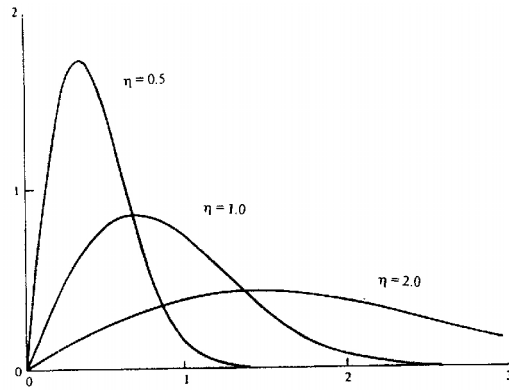
$$F(x) = 1 - \exp\left[-\left(\frac{x-\gamma}{\eta}\right)^{B}\right]$$ (2.67)

Figure 2.22. Probability density of Weibull distribution with β = 2.0,
γ = 0, η = 0.5, 1, 2

For $\gamma = 0$, the cumulative distribution is given by

$$F(x) = 1 - \exp\left[-\left(\frac{x}{\eta}\right)^{\beta}\right]$$ (2.68)

The expected value of the Weibull distribution is given by:

$$E(X) = \gamma + \eta \times \Gamma\left(\frac{1}{\beta} + 1\right)$$ (2.69)

where Γ is the gamma function, defined as

$$\Gamma(n) = \int\limits_{0}^{\infty} e^{-x} \times x^{n-1} dx$$

When n is integer then $\Gamma(n) = (n-1)!$. For other values, one has to solve the above integral to the value. Values for this can be found in Gamma function table given in the appendix. In *Microsoft EXCEL,* Gamma function, $\Gamma(x)$ can be found using the function, *EXP[GAMMALN(x)]*.

The variance of the Weibull distribution is given by:

$$V(X) = (\eta)^2 \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right) \right] \qquad (2.70)$$

Chapter 3

Reliability Measures

I have seen the future; and it works

Lincoln Steffens

In this chapter we discuss various measures by which hardware and software reliability characteristics can be numerically defined and described. Manufacturers and customers use reliability measure to quantify the effectiveness of the system. Use of any particular reliability measure depends on what is expected of the system and what we are trying measure. Several life cycle decision are made using reliability measure as one of the important design parameter. The reliability characteristics or measures used to specify reliability must reflect the operational requirements of the item. Requirements must be tailored to individual item considering operational environment and mission criticality. In broader sense, the reliability metrics can be classified (Figure 3.1) as: 1. Basic Reliability Measures, 2. Mission Reliability Measures, 3. Operational Reliability Measures, and 4. Contractual Reliability Measures.

Basic Reliability Measures are used to predict the system's ability to operate without maintenance and logistic support. Reliability measures like reliability function and failure function fall under this category.

Mission Reliability Measures are used to predict the system's ability to complete mission. These measures consider only those failures that cause mission failure. Reliability measures such as mission reliability, maintenance free operating period (MFOP), failure free operating period (FFOP), and hazard function fall under this category.

Operational Reliability Measures are used to predict the performance of the system when operated in a planned environment including the combined effect of design, quality, environment, maintenance, support policy, etc.

Measures such as Mean Time Between Maintenance (MTBM), Mean Time Between Overhaul (MTBO), Maintenance Free Operating Period (MFOP), Mean Time Between Critical Failure (MTBCF) and Mean Time Between Unscheduled Removal (MTBUR) fall under this category.

Contractual Reliability Measure is used to define, measure and evaluate the manufacturer's program. Contractual reliability is calculated by considering design and manufacturing characteristics. Basically it is the inherent reliability characteristic. Measures such as Mean Time To Failure (MTTF), Mean Time Between Failure (MTBF) and Failure rate fall under this category.
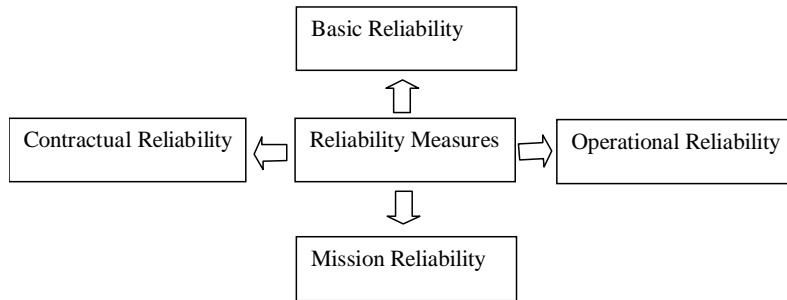

Figure 3.1 Classifications of Reliability Measures

Though we classify the reliability measures into four categories as mentioned above, one may require more than one reliability metric in most of the cases for specifying reliability requirements. Selection of specific measure to quantify the reliability requirements should include mission and logistic reliability along with maintenance and support measures. Currently, many manufacturers specify reliability by using mean time between failure (MTBF) and failure rate. However, MTBF and failure rates have several drawbacks. Recent projects such as Future Offensive Air Systems (FOAS) drive maintenance free operating periods (MFOP) as the preferred reliability requirement.

In the next Section, we define various reliability measures and how to evaluate them in practical problems. All the measures are defined based on the assumption that the time-to-failure (TTF) distribution of the system is known. Procedures for finding the time-to-failure distribution by analysing the failure data that are discussed in Chapter 12.

3.12.   FAILURE FUNCTION

Failure function is a basic (logistic) reliability measure and is defined as the probability that an item will fail before or at the moment of operating time t. *Here time t is used in a generic sense and it can have units such as miles, number of landings, flying hours, number of cycles, etc., depending on the operational profile and the utilisation of the system*. That is, Failure function is equal to the probability that the time-to-failure random variable will be less than or equal a particular value t (in this case operating time, see Figure 3.2a). The failure function is usually represented as *F(t)*.

F(t) = P (failure will occur before or at time t) = P (*TTF* ≤ *t*)

$$= \int_0^t f(u)\,du \qquad\qquad\qquad (3.1)$$

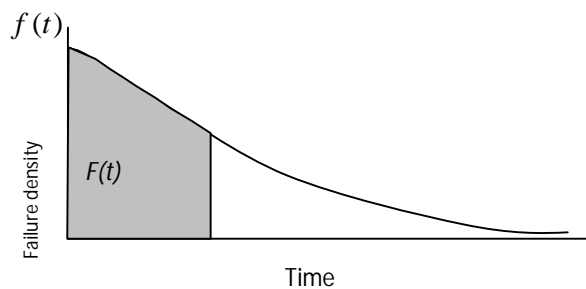Figure 3.2a. Failure function of a hypothetical distribution

Where $f(t)$ is the probability density function of the time-to-failure random variable *TTF*. Exponential, Weibull, normal, lognormal, Gamma and Gumbel are few popular theoretical distributions that are used to represent failure function. Equation (3.1) is derived by assuming that no maintenance is performed to the system, and gives the probability of failure free operation without maintenance up to time t.   However, most of the complex systems will require maintenance at frequent intervals. In such cases, equation (3.1) has to be modified, to incorporate the behaviour of the system under maintenance.   Failure functions of few popular theoretical distributions are listed in Table 3.1.

It should be noted that in case of normal distribution the failure function exists between -∞ and +∞, so may have significant value at t ≤ 0. Since negative time is meaningless in reliability, great care should be taken in using normal distribution for the failure function. For μ >> 3σ, probability values for t ≤ 0 can be considered negligible.

Table 3.1 Failure function, *F(t)*, of few theoretical distributions

Distribution	Failure Function, F(t)
Exponential	$1 - \exp(-\lambda t) \qquad t > 0, \lambda > 0$
Normal	$\int_0^t \frac{1}{\sigma\sqrt{2\pi}} e^{-[\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2]} dx \quad or \quad \Phi\left(\frac{t-\mu}{\sigma}\right)$ *or NORMDIST(t, μ, σ, TRUE) in EXCEL®*
Lognormal	$\int_0^t \frac{1}{\sigma_l x\sqrt{2\pi}} e^{-\left(\frac{1}{2}\left(\frac{\ln(x)-\mu_l}{\sigma_l}\right)^2\right)} dx \quad or \quad \Phi\left(\frac{\ln(t)-\mu_l}{\sigma_l}\right)$ *or NORMDIST(ln(t), μ, σ, TRUE) in EXCEL®*
Weibull	$1 - \exp(-(\frac{t-\gamma}{\eta})^\beta) \qquad \eta, \beta, \gamma > 0, t \geq \gamma$
Gamma	$\frac{1}{\Gamma(\alpha)} \int_0^t \beta^\alpha x^{\alpha-1} e^{-\beta x} dx$

Note that the failure function of normal distribution is defined between 0 and t, since t is greater than 0 for reliability purposes (against the usual limit -∞) Applications of failure function are listed below (Figure 3.2b). Failure functions of various theoretical distributions for different parameter values are shown in Figures 3.3a-3.3c.

Characteristics of failure function

1. Failure function is an increasing function. That is, for $t_1 < t_2$, F $(t_1) \leq$ F (t_2).
2. For modelling purposes it is assumed that the failure function value at time t = 0, *F(0) = 0*. However, this assumption may not be valid always.

For example, systems can be *dead on arrival.* The value of failure function increases as the time increases and for t = ∞, *F(∞) = 1*.

Applications of failure function

1. *F(t)* is the probability that an individual item will fail by time t.
2. *F(t)* is the fraction of items that fail by time t.
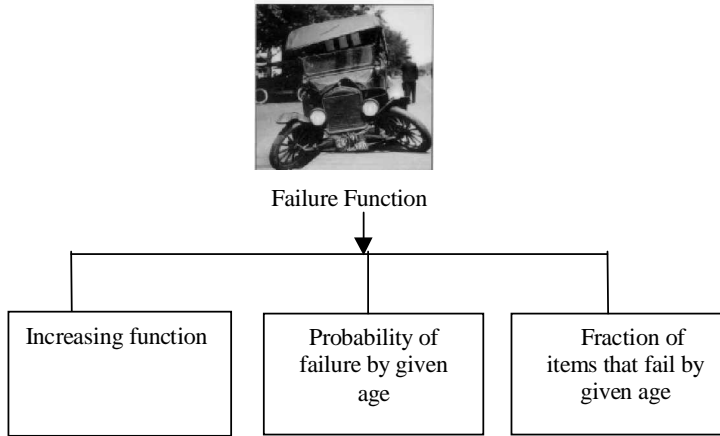3. 1 - *F(t)* is the probability that an individual item will survive up to time t.

Figure 3.2b. Properties of failure function
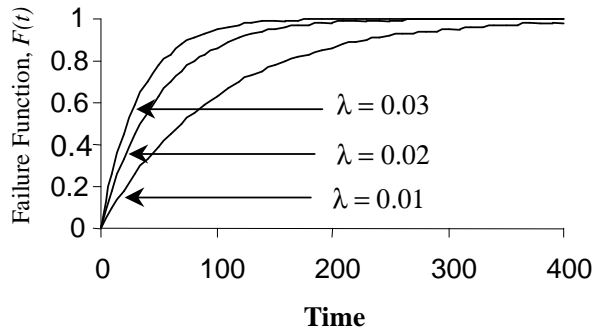
Figure 3.3a: Failure function of exponential distribution for different values of λ
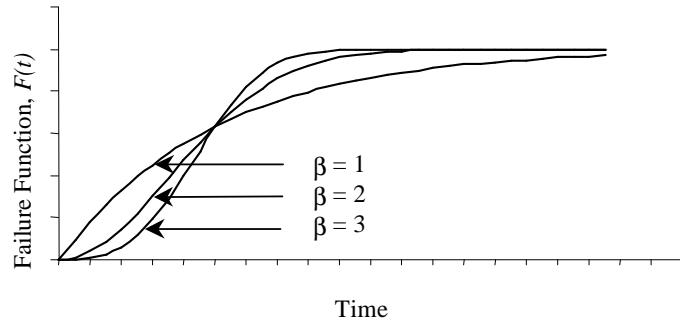
Figure 3.3b Failure function of Weibull distribution for different β values
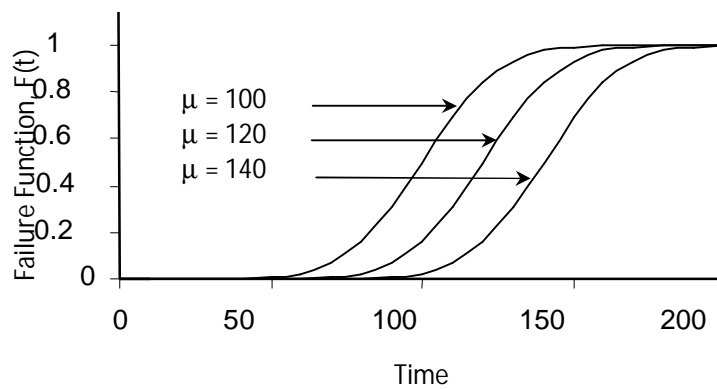
Figure 3.3c Failure function of normal distribution for different μ values

Example 3.1

The time to failure distribution of a sub-system in an aircraft engine follows Weibull distribution with scale parameter η = 1100 flight hours and the shape parameter β = 3. Find:

a) Probability of failure during first 100 flight hours.
b) Find the maximum length of flight such that the failure probability is less than 0.05.

SOLUTION:

a) The failure function for Weibull distribution is given by:

$$F(t) = 1 - \exp(-(\frac{t-\gamma}{\eta})^\beta)$$

It is given that: t = 100 flight hours, η = 1100 flight hours, β = 3 and γ = 0.

Probability of failure within first 100 hours is given by:

$$F(100) = 1 - \exp(-(\frac{100-0}{1100})^3) = 0.00075$$

b) If t is the maximum length of flight such that the failure probability is less than 0.05, we have

$$F(t) = 1 - \exp(-(\frac{t-0}{1100})^3) < 0.05$$

$$= \exp(-(\frac{t}{1100})^3) > 0.95$$

$$= (\frac{t}{1100})^3 > -\ln 0.95 \Rightarrow t = 1100 \times [-\ln(0.95)]^{1/3}$$

Now solving for t, we get t = 408.70 flight hours. The maximum length of flight such that the failure probability is less than 0.05 is 408.70 flight hours.

Example 3.2

The time to failure distribution of a Radar Warning Receiver (RWR) system in a fighter aircraft follows Weibull distribution with scale parameter 1200 flight hours and shape parameter 3. The time to failure distribution of the same RWR in a helicopter follows exponential distribution with scale parameter 0.001. Compare the failure function of the RWR in the fighter aircraft and the helicopter. If the supplier gives a warranty for 750 flight hours, calculate the risk involved with respect to fighter aircraft and the helicopter. (Although we have a same system, the operating conditions have significant impact on the failure function. In this case, RWR in helicopter is subject to more vibrations compared to aircraft).

SOLUTION:

The failure function of RWR on the fighter aircraft is given by:

$$F(t) = 1 - \exp(-(\frac{t}{1200})^3)$$

The failure function of RWR on the helicopter is given by:

$$F(t) = 1 - \exp(-(0.001 \times t))$$

Figure 3.4 depicts the failure function of RWR in fighter aircraft and the helicopter.
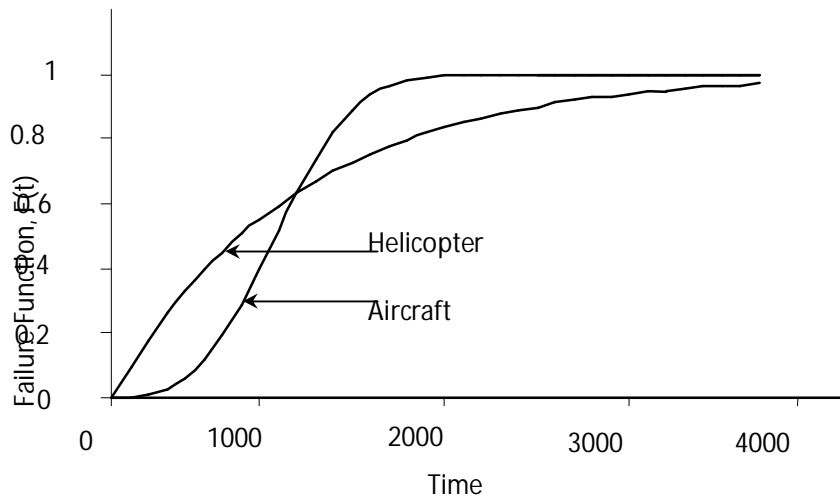
Figure 3.4 Failure function of RWR in fighter aircraft and helicopter

If the supplier provides warranty for 750 flight hours the risk associated with aircraft is given by:

$$F(750) = 1 - \exp(-(\frac{750}{1200})^3) = 0.2166$$

That is, just above 21% percent of RWR are likely to fail if the RWR is installed in the aircraft.

If the RWR is installed in helicopter then the associated risk is given by:

$$F(750) = 1 - \exp(-0.001 \times 750) = 0.5276$$

In the case of helicopter, more than 52% of the RWR's are likely to fail before the warranty period.

3.12.1   Failure function of system under multiple failure mechanisms

It is seldom true that an item's failure is caused by a single failure mechanism. In most of the cases there will be more than one (some times

hundreds) mechanism that causes the failure of an item. The expression (3.1) is more appropriate when the failure is caused by a single failure mechanism. However, most of the practical systems fail due to different causes or different failure mechanisms. Assume that the system failure is due to two different failure mechanisms. Let $f_1(t)$ and $f_2(t)$ be the probability density function of the system due to failure mechanism 1 and 2 respectively. Now the probability density function of the time-to-failure of the system caused by either of the failure mechanisms:

$$f(t) = f_1(t)[1 - F_2(t)] + f_2(t)[1 - F_1(t)]$$

where, $F_1(t)$ and $F_2(t)$ the are failure function for failure mechanism 1 and 2 respectively. The failure function of the item under two different failure mechanism is given by:

$$F(t) = \int_0^t \{f_1(x)[1 - F_2(x)] + f_2(x)[1 - F_1(x)]\}dx \qquad (3.2)$$

Example 3.3

Failure of an item is caused by two different failure mechanisms (say failure mechanism A and B). The time-to-failure distribution of the item due to failure mechanism A can be represented by exponential distribution with parameter $\lambda_A = 0.002$ hours. The time-to-failure distribution of the item due to failure mechanism B can be represented by exponential distribution with parameter $\lambda_B = 0.005$ hours. Find the probability that the item will fail before 500 hours of operation.

SOLUTION:
Assume that $f_A(t)$ and $f_B(t)$ represent probability density function of the time-to-failure random variable due to failure mechanism A and B respectively. Thus,

$$f_A(t) = \lambda_A \exp(-\lambda_A t), \quad 1 - F_A(t) = \exp(-\lambda_A t)$$
$$f_B(t) = \lambda_B \exp(-\lambda_B t), \quad 1 - F_B(t) = \exp(-\lambda_B t)$$

Now the failure function of the item is given by:

$$F(t) = \int_0^t \{\lambda_A \exp(-(\lambda_A + \lambda_B)x) + \lambda_B \exp(-(\lambda_A + \lambda_B)x)dx$$

$$= (\lambda_A/\lambda_A + \lambda_B)[1 - \exp(-(\lambda_A + \lambda_B)t]$$

$$+ (\lambda_B/\lambda_A + \lambda_B)[1 - \exp(-(\lambda_A + \lambda_B)t]$$

$$= [1 - \exp(-(\lambda_A + \lambda_B)t]$$

Figure 3.5 represents the failure function due to failure mechanism 1, 2 and the system failure function. The probability that the item will fail by 500 hours is given by:

$$F(500) = 1 - \exp(-((0.005 + 0.002) \times 500)) = 0.9698$$
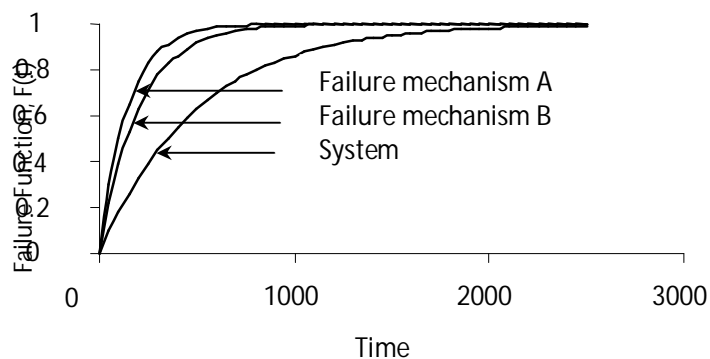
Figure 3.5 Failure function due to different failure mechanisms

3.13.   RELIABILITY FUNCTION

Reliability is the ability of the item to maintain the required function for a specified period of time (or mission time) under given operating conditions. *Reliability function, R(t), is defined as the probability that the system will not fail during the stated period of time, t, under stated operating conditions.* If *TTF* represents the time-to-failure random variable with failure function (cumulative distribution function) *F(t)*, then the reliability function *R(t)* is given by:

$$R(t) = P\{\text{the system doesn't fail during } [0, t]\} = 1 - F(t) \qquad (3.3)$$

In equation (3.3) we assume that the age of the system before the start of the mission is zero. Thus the equation (3.3) is valid only for new systems or those systems whose failures are not age related (that is, the time-to-failure follows exponential distribution due to *memory less* property of exponential distribution). However, in most of the cases this assumption may not be valid. If the system age is greater than zero at the beginning of the mission, then we have to calculate mission reliability function, which will be discussed later. Figure 3.6 depicts the relation between reliability function and the *TTF* density function. *R(t)* is the area under *TTF* density between t and ∞.
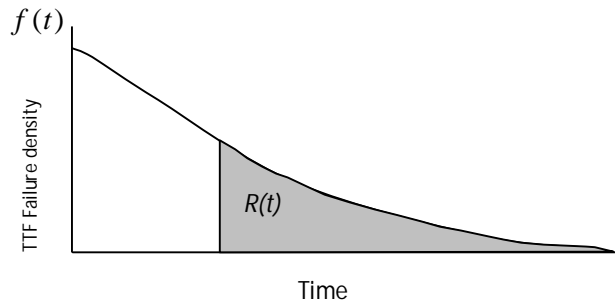
Figure 3.6 Reliability function of a hypothetical probability distribution

Properties of reliability function:

1. Reliability is a decreasing function with time t. That is, for $t_1 < t_2$; $R(t_1) \geq R(t_2)$.
2. It is usually assumed that R (0) = 1. As t becomes larger and larger R(t) approaches zero, that is, R(∞).

Applications of reliability function

1. R(t) is the probability that an individual item survives up to time t.
2. R(t) is the fraction of items in a population that survive up to time t.
3. R(t) is the basic function used for many reliability measures and system reliability prediction.

Reliability function for some important life distributions are given in Table 3.2. Figure 3.7a-c represents reliability function of various theoretical distributions for different parameter values.

Table 3.2. Reliability function, *R(t)*, for popular theoretical distributions

Distribution	Reliability function, R(t)
Exponential	$\exp(-\lambda t) \qquad t > 0, \lambda > 0$
Normal	$\Phi(\dfrac{\mu - t}{\sigma}) = 1 - \int\limits_0^t \dfrac{1}{\sigma\sqrt{2\pi}}\, e^{-\left(\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)} dx$ *or NORMDIST (μ, t, σ, TRUE) in EXCEL*
Lognormal	$\Phi(\dfrac{\mu_l - \ln t}{\sigma_l}) = 1 - \int\limits_0^t \dfrac{1}{\sigma_l x\sqrt{2\pi}}\, e^{-\left(\frac{1}{2}\left(\frac{\ln(x)-\mu_l}{\sigma_l}\right)^2\right)} dx$ or NORMDIST (μ, ln(t), σ, TRUE) in EXCEL
Weibull	$\exp(-(\dfrac{t-\gamma}{\eta})^{\beta}) \qquad \eta, \beta, \gamma > 0, t \geq \gamma$
Gamma	$1 - \dfrac{1}{\Gamma(\alpha)}\int\limits_0^t \beta^{\alpha} x^{\alpha-1} e^{-\beta x}\, dx$
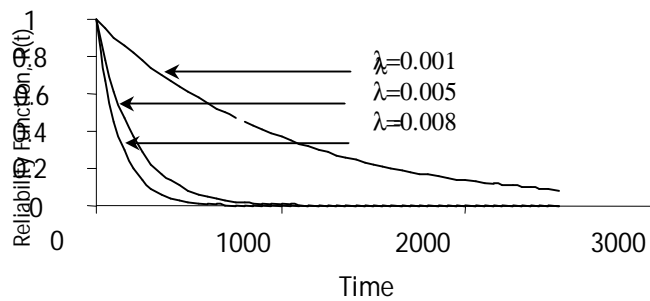
Figure 3.7 a. Reliability function of exponential distribution for different values of λ
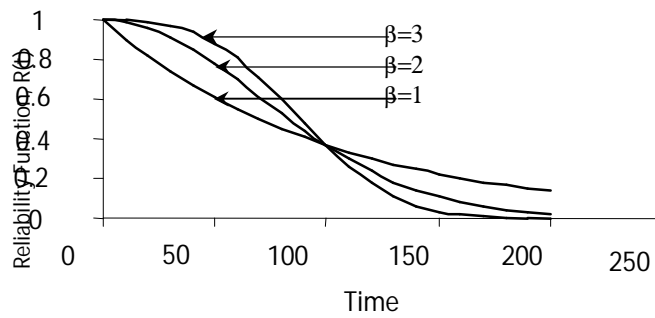
Figure 3.7 b. Reliability function of Weibull distribution for different values of β
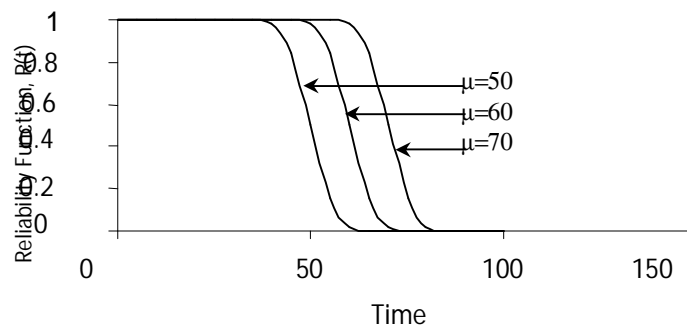
Figure 3.7c. Reliability function of Normal distribution for different values of μ

Example 3.4

Time to failure distribution of a computer memory chip follows normal distribution with mean 9000 hours and standard deviation 2000 hours. Find the reliability of this chip for a mission of 8000 hours.

SOLUTION

Using Table 3.2, the reliability for a mission of 8000 hours is given by:

$$R(t) = \Phi(\frac{\mu - t}{\sigma}) = \Phi(\frac{9000 - 8000}{2000}) = \Phi(0.5) = 0.6915$$

Example 3.5

The time to failure distribution of a steam turbo generator can be represented using Weibull distribution with η = 500 hours and β = 2.1. Find the reliability of the generator for 600 hours of operation.

SOLUTION:

Again using Table 3.2, reliability of the generator for 600 hours of operations is given by:

$$R(t) = \exp(-(600/500)^{2.1}) = 0.2307$$

3.13.1   Reliability function for items under multiple failure mechanisms

Assume that the failure of the item is caused due to two different failure mechanisms. Let $f_1(t)$ and $f_2(t)$ be the probability density function of the time-to-failure random variable due to failure mechanism 1 and 2 respectively. The probability density function of the time-to-failure of the item is given by caused by either of the failure mechanisms:

$$f(t) = f_1(t)[1 - F_2(t)] + f_2(t) \cdot [1 - F_1(t)]$$

Where $F_1(t)$ and $F_2(t)$ are failure function for failure mechanism 1 and 2 respectively. The Reliability function of the item under two different failure mechanism is given by:

$$R(t) = 1 - F(t) = 1 - \int_0^t \{ f_1(x)[1 - F_2(x)] + f_2(x)[1 - F_1(x)] \} dx \ (3.4)$$

The above result can be extended to obtain expression for reliability function due to more than two failure mechanisms.

Example 3.6

For the example 3.3, find the reliability of the item for 200 hours.

SOLUTION:

Using the expression for failure function obtained in example 3.3, the reliability function can be written as:

$$R(t) = \exp(-(\lambda_A + \lambda_B) \times t)$$
$$R(200) = \exp(-(0.002 + 0.005) \times 200) = 0.2465$$

3.13.2   Mission Reliability Function

In many practical situations, one might be interested in finding the probability of completing a mission successfully. Success probability of hitting an enemy target and returning to the base is an example where mission reliability function can be used.   The main difference between reliability function and the mission reliability function is that, in mission reliability we recognise the age of the system before the mission. *Mission reliability is defined, as the probability that the system aged t_b is able to complete mission duration of t_m successfully.*   We assume that no maintenance is performed during the mission. The expression for mission reliability MR (t_b , t_m) is given by

$$MR(t_b, t_m) = \frac{R(t_b + t_m)}{R(t_b)} \tag{3.5}$$

where, t_b is the age of the item at the beginning of the mission and t_m is the mission period. If the time to failure distribution is exponential, then the following relation is valid.

$$MR(t_b, t_m) = R(t_m)$$

Application of mission reliability function

1. Mission reliability, MR(t_a, t_m) gives the probability that an individual item aged t_a will complete a mission duration of t_m hours without any need for maintenance.
2. Mission reliability is the appropriate basic reliability measure for ageing items or items whose time-to-failure distribution is other then exponential.

Example 3.7

Time-to-failure distribution of the gearbox within an armoured vehicle can be modelled using Weibull distribution with scale parameter η = 2400 miles and shape parameter β = 1.25. Find the probability that that gearbox will not fail during a mission time of 200 miles. Assuming that the age of the gearbox is 1500 miles.

SOLUTION:

Given, t_b = 1500 miles and t_m = 200 miles

$$MR(t_b, t_m) = \frac{R(t_m + t_b)}{R(t_b)} = \frac{R(1700)}{R(1500)}$$

$$R(1700) = \exp(-(\frac{1700}{2400})^{1.25}) = 0.5221$$

$$R(1500) = \exp(-(\frac{1500}{2400})^{1.25}) = 0.5736$$

$$MR(1500,200) = \frac{R(1700)}{R(1500)} = \frac{0.5221}{0.5736} = 0.9102$$

That is, the gearbox aged 1500 miles has approximately 91% chance of surviving a mission of 200 miles.

3.14.   DESPATCH RELIABILITY

Despatch reliability (DR) is one of popular reliability metrics used by commercial airlines around the world. Despatch reliability is defined as the percentage of revenue departures that do not occur in a delay or cancellation due to technical problems. For most airlines, the delay means that the aircraft is delayed more than 15 minutes. Technical delays occur can be caused due to some unscheduled maintenance. Airlines frequently seek DR guarantees where the aircraft manufactures face penalties if DR levels are not achieved. For commercial airlines despatch reliability is an important economic factor, it is estimated that delay cost per minute for

large jets can be as high as 1000 US dollars. The expression for despatch reliability is given by:

$$DR(\%) = \frac{100 - ND_{15} - NC}{100} \times 100\% \qquad (3.6)$$

Where,

ND_{15} = Number of delays with more than 15 minutes delay

NC = the number of cancellations

Equation (3.6) is applied only to technical delays. DR is a function of equipment reliability, system and component maintainability, and overall logistic support.

3.15. HAZARD FUNCTION (HAZARD RATE OR INSTANTANEOUS FAILURE RATE)

Hazard function (or hazard rate) is used as a parameter for comparison of two different designs in reliability theory. Hazard function is the indicator of the effect of ageing on the reliability of the system. It quantifies the risk of failure as the age of the system increases. Mathematically, it represents the conditional probability of failure in an interval t to t + δt given that the system survives up to t, divided by δt, as δt tends to zero, that is,

$$h(t) = \lim_{\delta t \to 0} \frac{1}{\delta t} \cdot \frac{F(t + \delta t) - F(t)}{R(t)} = \lim_{\delta t \to 0} \frac{R(t) - R(t + \delta t)}{\delta t R(t)} \qquad (3.7)$$

Note that hazard function, h(t), is not a probability, it is the limiting value of the probability. However, h(t)δt, represents the probability that the item will fail between ages t and t+δt as δt →0. The above expression can be simplified so that

$$h(t) = \frac{f(t)}{R(t)} \qquad (3.8)$$

Thus, the hazard function is the ratio of the probability density function to the reliability function. Integrating both sides of the above equation, we get:

$$\int_0^t h(x)dx = \int_0^t \frac{f(x)}{R(x)}dx$$

$$= \int_0^t -\frac{R'(x)}{R(x)}dx = -\ln R(t)$$

Thus reliability can be written as:

$$R(t) = \exp\left[-\int_0^t h(x)dx\right] \qquad (3.9)$$

From equation (3.9), it immediately follows that:

$$f(t) = h(t)\exp(-\int_0^t h(x)dx \qquad (3.10)$$

The expression (3.10), which relates reliability and hazard function, is valid for all types of time to failure distribution. Hazard function shows how the risk of the item in use changes over time (hence also called *risk rate)*. The hazard functions of some important theoretical distributions are given in Table 3.3.

Characteristics of hazard function

1. Hazard function can be increasing, decreasing or constant.
2. Hazard function is not a probability and hence can be greater than 1.

Table 3.3. Hazard function, $h(t)$, of few theoretical distributions

Distribution	Hazard function, h(t)
Exponential	λ

Distribution	Hazard function, h(t)
Normal	$f(t)/\Phi(\dfrac{\mu-t}{\sigma})$, *f(t)* is the pdf of normal distribution.
Lognormal	$f_l(t)/\Phi(\dfrac{\mu_l-t}{\sigma_l})$, *fl(t)* is the pdf of lognormal distribution.
Weibull	$\dfrac{\beta}{\eta}(\dfrac{t}{\eta})^{\beta-1}$
Gamma	$[\dfrac{\beta^{\alpha}}{\Gamma(\alpha)}t^{\alpha-1}e^{-\beta t}]/1-\dfrac{1}{\Gamma(\alpha)}\int\limits_{0}^{t}\beta^{\alpha}x^{\alpha-1}e^{-\beta x}dx$

Applications of hazard function

1. h(t) is loosely considered as failure rate at time t (time-dependent)
2. h(t) quantifies the amount of risk a system is under at time t.
3. For h(t) $\leq$ 1, it is not recommended to carry out preventive maintenance.

Figures 3.8a-c show hazard function of various theoretical distributions for different parameter values.
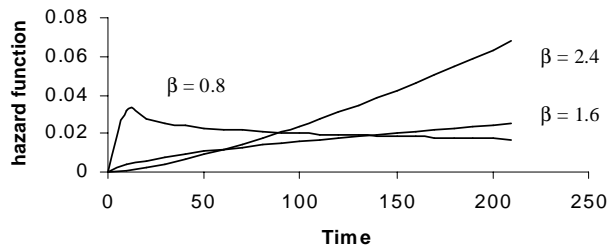
Figure 3.8a Hazard function of Weibull distribution for different values of β
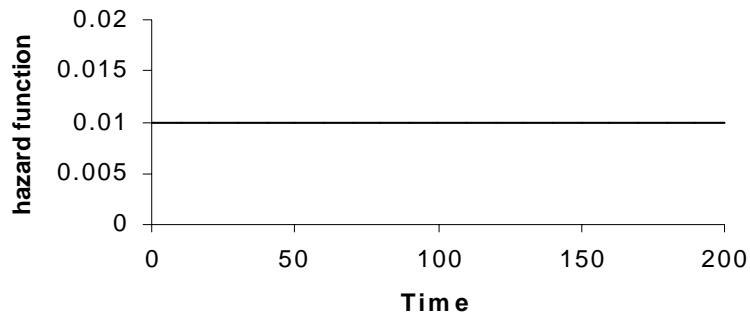
Figure 3.8b Hazard function of exponential distribution
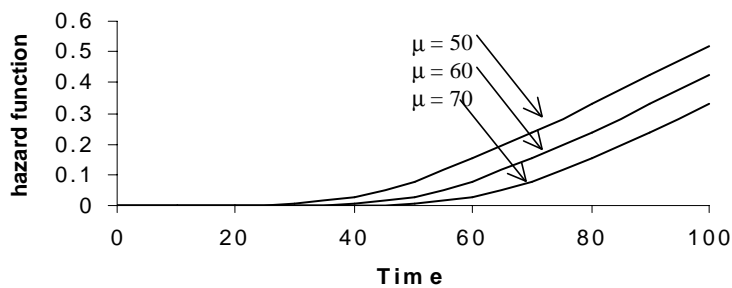
Figure 3.8c Hazard function of normal distribution for different values of μ

Example 3.8

Time to failure distribution of a gas turbine system can be represented using Weibull distribution with scale parameter η = 1000 hours and shape parameter β = 1.7. Find the hazard rate of the gas turbine at time t = 800 hours and t = 1200 hours.

SOLUTION:

The hazard rate for Weibull distribution is given by:

$$h(t) = \frac{\beta}{\eta}(\frac{t}{\eta})^{\beta-1}$$

$$h(800) = \frac{1.7}{1000}(\frac{800}{1000})^{0.7} = 0.00145$$

$$h(1200) = \frac{1.7}{1000}(\frac{1200}{1000})^{0.7} = 0.0019$$

3.15.1 Cumulative hazard function

Cumulative hazard function represents the cumulative hazard or risk of the item during the interval [0,t]. Cumulative hazard function, H(t), is given by:

$$H(t) = \int_0^t h(x)dx \qquad (3.11)$$

Reliability of an item can be conveniently written using cumulative hazard as:

$$R(t) = e^{-H(t)} \qquad (3.12)$$

3.15.2 Cumulative hazard function and the expected number of failures

Consider an item, which upon failure is subject to minimal repair. That is, the hazard rate after repair is same as the hazard rate just before failure. If *N(t)* is the total number of failures by time t, then *M(t) = E [N(t)]* is the expected number of failures by time t. It can be shown that under the assumption that the item receives minimal repair[*] ('*as-bad -as-old*') after each failure, then

$$E[N(t)] = M(t) = \int_0^t h(x)dx \qquad (3.13)$$

---

[*] Mathematically minimal repair or '*as bad as old*' means that the hazard rate of the item after repair will be same as the hazard rate just prior to failure.

The above expression can be used to model different maintenance/replacement policies. In case of exponential and Weibull time to failure distributions we get the following simple expressions for the expected number of failures of an item subject to minimal repair.

Exponential time to failure distribution

For exponential distribution, the expected number of failures is given by

$$E[N(t)] = \int_0^t h(x)dx = \int_0^t \lambda dx = \lambda t \qquad (3.14)$$

Weibull time to failure distribution

$$E[N(t)] = \int_0^t h(x)dx = \int_0^t \frac{\beta}{\eta}(\frac{x}{\eta})^{\beta-1} dx = (\frac{t}{\eta})^{\beta} \qquad (3.15)$$

Example 3.9
An item is subject to minimal repair whenever it failed. If the time to failure of the item follows Weibull distribution with η = 500 and β = 2. Find: 1. The number of times the item is expected to fail by 1500 hours, and 2. The cost of the item is \$ 200. If the cost of minimal repair is \$ 100 per each repair, is it advisable to repair or replace the item upon failure.

SOLUTION:

1. The expected number of failures is given by:

$$E[N(t)] = [\frac{t}{\eta}]^{\beta} = [\frac{1500}{500}]^2 = 3^2 = 9$$

2. Using the above result the cost associated with repair, C_{repair} (t) = 9 × 100 = \$ 900.
If the item is replaced, then the expected number of failures is given by the renewal function, *M(t)* [refer chapter 4], where

$$M(t) = \sum_{i=1}^{\infty} F^i(t)$$

For the above case, the value of M(t) < 4 (The actual calculation of the above function will be discussed in Chapter 4). Thus the cost due to replacement will be less than $4 \times 200 = \$ 800$. Thus, it is better to replace the item upon failure rather using minimal repair.

3.15.3 Typical Forms of Hazard Function

In practice, hazard function can have different shapes. Figure 3.9 shows most general forms of hazard function. Recent research in the field of reliability centred maintenance (RCM) shows that the hazard rate mostly follows six different patterns. Depending on the equipment and its failure mechanism, one can say that the hazard function may follow any one of these six patterns. *However, one should not blindly assume that hazard rate of any item will follow any one of these six patterns.* These are only possible cases based on some data.

Pattern A is called the *bathtub curve* and consist of three distinct phases. It starts with early failure region (known as burn-in or infant mortality) characterised by decreasing hazard function. Early failure region is followed by constant or gradually increasing region (called useful life). The constant or gradually increasing region is followed by wear out region characterised by increasing hazard function. The reason for such as shape is that the early decreasing hazard rate results from manufacturing defects. Early operation will remove these items from a population of like items. The remaining items have a constant hazard for some extended period of time during which the failure cause is not readily apparent. Finally those items remaining reach a wear-out stage with an increasing hazard rate. One would expect bathtub curve at the system level and not at the part or component level (unless the component has many failure modes which have different *TTF* distribution). It was believed that bathtub curve represents the most general form of the hazard function. However, the recent research shows that in most of the cases hazard function do not follow this pattern.

Pattern B starts with high infant mortality and then follows a constant or very slowly increasing hazard function. Pattern C starts with a constant or slowly increasing failure probability followed by wear out (sharply increasing) hazard function. Pattern D shows constant hazard throughout the file. Pattern E represents a slowly increasing hazard without any sign of

wear out. Pattern F starts with a low hazard initially followed by a constant hazard.
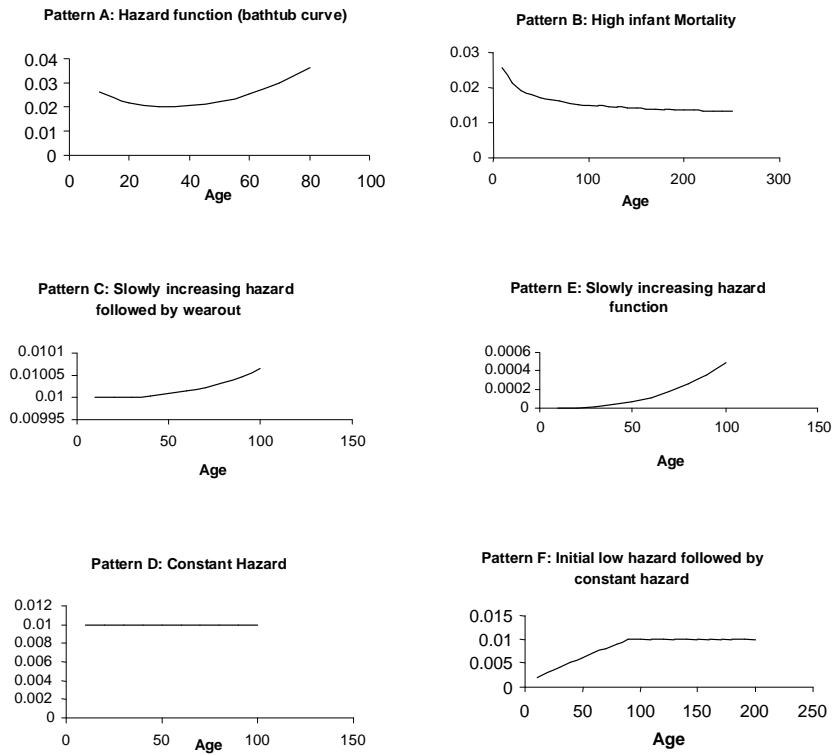
Figure 3.9. Different forms of hazard function

Table 3.4 shows the relationship between failure function, reliability function and hazard function.

Table 3.4. Relationship between F(t), R(t) and h(t)

	F(t)	*R(t)*	*h(t)*
F(t)	-----	1 - *R(t)*	$1 - \exp(-\int_{0}^{t} h(x)dx)$

	F(t)	R(t)	h(t)
R(t)	1 - F(t)	------	$\exp(-\int\limits_{0}^{t} h(x)dx$
h(t)	$F^{'}(t)/[1-R(t)]$	$-R^{'}(t)/R(t)$	-------

3.15.4 Failure rate

Whenever the hazard function is constant, we call it as failure rate. That is, failure rate is a special case of hazard function (which is time dependent failure rate). Failure rate is one of the most widely used contractual reliability measures in the defence and aerospace industry. By definition, it is appropriate to use failure rate only when the time-to-failure distribution is exponential. Also, failure rate can be used only for a non-repairable system. Many defence standards such as MIL-HDBK-217 and British DEF-STAN 00-40 recommend the following equation for estimating the failure rate.

$$Failure\ rate = \frac{Total\ number\ of\ failures\ in\ a\ sample}{Cumulative\ operating\ time\ of\ the\ sample} \quad (3.16)$$

Care should be taken in using the above equation, for good estimation one has to observe the system failure for a sufficiently large operating period.

Applications of failure rate

1. Failure rate represents the number of failures per unit time.
2. If the failure rate is λ, then the expected number of items that fail in [0,t] is λt.
3. Failure rate is one of the popular contractual reliability measures among many industries including aerospace and defence.

3.16.   MEAN TIME TO FAILURE (MTTF)

MTTF represents the expected value of a system's time to first failure. It is used as a measure of reliability for *non-repairable* items such as bulb, microchips and many electronic circuits. Mathematically, MTTF can be defined as:

$$MTTF = \int_0^\infty tf(t)dt = \int_0^\infty R(t)dt \qquad (3.17)$$

Thus, MTTF can be considered as the area under the curve represented by the reliability function, R(t), between zero and infinity. If the item under consideration is repairable, then the expression (3.17) represents mean time to first failure of the item. Figure 3.10 depicts the MTTF value of an item.

For many reliability functions, it is difficult to evaluate the integral (3.17). One may have to use numerical approximation such as trapezium approach to find MTTF value.
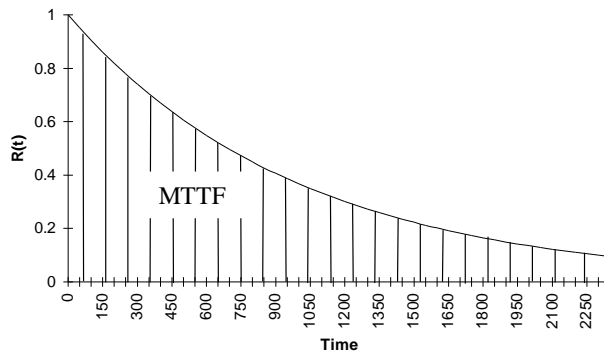
Figure 3.10 MTTF of an item as a function of Reliability

MTTF is one of the most popular measures for specifying reliability of non-repairable items among military and Government organisations throughout the world. Unfortunately there are many misconception about MTTF among reliability analysts. During the *Gulf War*, one of Generals from a defence department said, '*We know exactly how many tanks to send, we measured the distance from the map and divided that by MTTF'*. What many people do not realise is that MTTF is only a measure of central tendency.

For example, if the time-to-failure distribution is exponential, then 63% of the items will fail before their age reaches MTTF value.

MTTF is one of the important contractual reliability measures for non-repairable (consumable) items. However, it is important to understand what MTTF value really means. For example let us assume that we have two items A and B with same MTTF (say 500 days). One might think that both the components have equal reliability. However, if the time to failure of the item A is exponential is that of item B is normal then there will be a significant variation in the behaviour of these items. Figure 3.11 shows the cumulative distribution of these two items up to 500 days. The figure clearly shows that items with exponential failure time show higher chance of failure during the initial stages of operation.
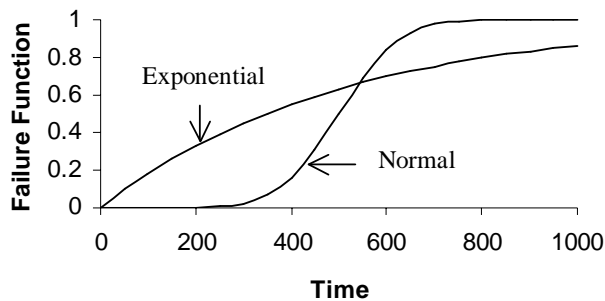
Figure 3.11 Comparison of item with same MTTF

Using the equation (3.17), the MTTF of various failure distributions are listed in Table 3.5.

It is easy to check that if the time to failure of the item is exponential then more than 63% of the items will fail by the time the age of the item reaches MTTF. In the case of normal distribution, it will be 50%.

Applications of MTTF

1. MTTF is the average life of a non-repairable system.
2. For a repairable system, MTTF represents the average time before the first failure.

3. MTTF is one of the popular contractual reliability measures for non-repairable systems.

Table 3.5. MTTF of different time-to-failure distributions

Distribution	MTTF
Exponential	$1/\lambda$
Normal	μ
Lognormal	$\exp(\mu_l + \dfrac{\sigma_l}{2})$
Weibull	$\eta \times \Gamma(1 + \dfrac{1}{\beta})$
Gamma	α / β

3.16.1   Mean Residual Life

In some cases, it may be of interest to know the expected value of the remaining life of the item before it fails from an arbitrary time t_0 (known as, *mean residual life*). We denote this value as MTTF(t_0), which represents the expected time to failure of an item aged t_0. Mathematically, MTTF(t_0) can be expressed as:

$$MTTF(t_0) = \int\limits_{t_0}^{\infty} (t - t_0) f(t|t_0) dt \tag{3.18}$$

$f(t|t_0)$ is the density of the conditional probability of failure at time t, provided that the item has survived over time t_0. Thus,

$$f(t \mid t_0) = h(t) \times R(t \mid t_0)$$

where, $R(t|t_0)$, is the conditional probability that the item survives up to time t, given that it has survived up to time t_0. Now, the above expression can be written as:

$$f(t \mid t_0) = h(t) \times \frac{R(t)}{R(t_0)}$$

The expression for MTTF(t_0) can be written as:

$$MTTF(t_0) = \int_{t_0}^{\infty} (t - t_0) h(t) \frac{R(t)}{R(t_0)} dt \qquad (3.19)$$

substituting for h(t) in the above equation, we have

$$MTTF(t_0) = \int_{t_0}^{\infty} \frac{(t - t_0) f(t)}{R(t_0)} dt = \frac{1}{R(t_0)} \int_{t_0}^{\infty} (t - t_0) f(t) dt$$

The above equation can be written as (using integration by parts):

$$MTTF(t_0) = \frac{\int_{t_0}^{\infty} R(t) dt}{R(t_0)} \qquad (3.20)$$

The concept of mean residual life can be successfully applied for planning maintenance and inspection activities.

Example 3.10

Companies A and B manufacture car tyres. Both the companies claim that the MTTF of their car tyre is 2000 miles. After analysing the field failure data of these two tyres it was found that the time to failure distribution of A is exponential with $\lambda = 0.0005$ and the time to failure distribution of B is normal with $\mu = 2000$ miles and $\sigma = 200$ miles. If the maintenance policy of the Exeter city car rentals is to replace the tyres as soon as it reaches 2000 miles which tyre they should buy:

SOLUTION:

Reliability of the car tyre produced by company A for 2000 miles, $R_A(2000)$, is given by:

$$R_A(2000) = \exp(-0.0005 \times 2000) = 0.3678$$

Reliability of the car tyre produced by company B for 2000 miles, $R_B(2000)$, is given by:

$$R_B(2000) = \Phi(\frac{\mu - 2000}{\sigma}) = \Phi(\frac{2000 - 2000}{200}) = \Phi(0) = 0.5$$

Thus, it is advisable to buy the tyres produced by company B.

Example 3.11

The time to failure of an airborne navigation radar can be represented using Weibull distribution with scale parameter η = 2000 hours and β = 2.1. It was told that the age of the existing radar is 800 hours. Find the expected value of the remaining life for this radar.

SOLUTION:
Using Equation (3.20), The MTTF(800) can be written as:

$$MTTF(800) = \frac{\int_{800}^{\infty} R(t)dt}{R(800)} = \frac{\int_{0}^{\infty} R(t)dt - \int_{0}^{800} R(t)dt}{R(800)}$$

$$\text{MTTF(800)} = MTTF(800) = \frac{MTTF - \int_{0}^{800} \exp(-(\frac{t}{2000})^{2.1} dt}{0.8641}$$

$$MTTF = \eta \times \Gamma(1 + \frac{1}{\beta}) = 2000 \cdot \Gamma(1 + \frac{1}{2.1}) = 1771.2$$

The value of $\Gamma(1 + \frac{1}{\beta})$ can be found from Gamma function table (see appendix).

Using numerical approximation, $\int\limits_{0}^{800} \exp(-(\frac{t}{2000})^{2.1} dt \approx 763.90$

Thus MTTF(800) $\approx$ (1771.2 - 763.90) / 0.8641 = 1165.72 hours

Thus, expected remaining life of the radar aged 800 hours is 1165.72 hours.

3.16.2   MTTF of a maintained system

Assume that an item is subject to preventive maintenance after every T_{pm} units, that is, at T_{pm}, $2T_{pm}$, $3T_{pm}$, etc. The expected time to failure, $MTTF_{pm}$, (MTTF of an subject to preventive) of the item is given by:

$$MTTF_{pm} = \int\limits_{0}^{\infty} R_{pm}(t)dt \tag{3.21}$$

Using additive property of integration, the above integral can be written as:

$$MTTF_{pm} = \int\limits_{0}^{T_{pm}} R_{pm}(t)dt + \int\limits_{T_{pm}}^{2T_{pm}} R_{pm}(t)dt + \int\limits_{2T_{pm}}^{3T_{pm}} R_{pm}(t)dt + ...$$

where R_{pm} (t) is the reliability of the item subject to preventive maintenance. If the item is restored to '*as-good-as-new*' state after each maintenance activity, then the reliability function between any two maintenance tasks can be written as:

$$R_{pm}(t) = R[T_{pm}]^k R(t), \qquad kT_{pm} \le t \le (k+1)T_{pm}$$

Using the above expression for $R_{pm}(t)$ in the integral (3.21) we have:

$$MTTF_{pm} = \int\limits_{0}^{T_{pm}} R(t)dt + \int\limits_{0}^{T_{pm}} R(T_{pm})R(t)dt + \int\limits_{0}^{T_{pm}} [R(T_{pm})]^2 R(t)dt + ...$$

$$= \{1 + R(T_{pm}) + [R(T_{pm})]^2 + ...\} \int\limits_{0}^{T_{pm}} R(t)dt$$

As R(t) ≤ 1, the above expression can be written as:

$$MTTF_{pm} = \frac{\int\limits_0^{T_{pm}} R(t)dt}{1 - R(T_{pm})} = \frac{\int\limits_0^{T_{pm}} R(t)dt}{F(T_{pm})} \qquad (3.22)$$

Similar logic can be used to derive the expression for $MTTF_{pm}$ when the repair is not perfect (that is, when the item is not as good as new after maintenance). $MTTF_{pm}$ can be used to quantify the effectiveness of the maintenance action. If $MTTF_{pm} > MTTF$, then one can say that the reliability can be improved by carrying out maintenance. If $MTTF_{pm} \leq MTTF$ then, the maintenance will not improve the reliability of the item. Figure 3.12 shows $MTTF_{pm}$ values of an item for different T_{pm} whose time-to-failure can be represented using Weibull distribution with $\eta = 200$ and $\beta = 2.5$. It can be noticed that as the value of T_{pm} increases, the $MTTF_{pm}$ converges to that of corrective maintenance.

Example 3.12

A solid state radar is subject to preventive maintenance after every 400 flight hours. The time to failure of the radar follows exponential distribution with mean life 800 flight hours. Find the $MTTF_{pm}$ of the radar.

SOLUTION:

We have: $T_0 = 500$ flight hours and $(1/\lambda) = 800$

$\lambda = (1/800) = 0.00125$

$$MTTF_{pm} = \frac{\int\limits_0^{400} \exp(-0.00125 \times t)dt}{1 - \exp(-0.00125 \times 400)} = 800$$

There is no improvement in the $MTTF_{pm}$ because the time to failure is exponential. *Thus, preventive maintenance will not improve the reliability of the system, if the time to failure is exponential. This example is used to demonstrate this well known fact mathematically.*

MTTFpm for different Tpm values
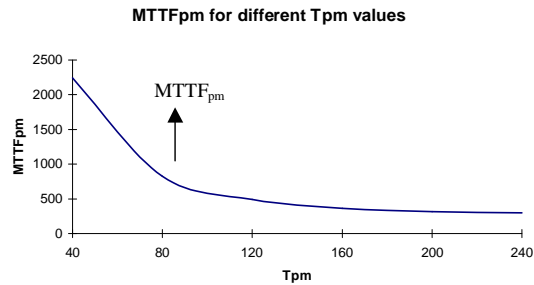
Figure 3.12. *MTTF$_{pm}$* of an item for different *T$_{pm}$* values

Example 3.13

A manufacturing company buys two machines A and B. The time to failure of machine A can be represented by Weibull distribution with η = 1000 hours and β = 2. The time to failure of machine B can be represented by Weibull distribution with η = 1000 hours and β = 0.5. The maintenance manager in charge of operation plan to apply preventive maintenance for both the machines for every 200 hours, so that he can improve the expected time to failure of the machines. Check whether the manager's decision is correct.

SOLUTION:

The MTTF$_{pm}$ for machine A is given by:

$$\text{MTTF}_{pm} = \; MTTF_{pm} \; \frac{\int_0^{200} \exp(-(t/1000)^2)\,dt}{\exp(-(200/1000)^2)} \approx 5033 \, hours$$

$$\text{MTTF for machine A is } \eta \times \Gamma(1+\frac{1}{B}) = 1000 \times \Gamma(1+\frac{1}{2}) = 886.2 \; hours$$

Thus for machine A, preventive maintenance will improve the mean time to failure of the system.

The MTTF$_{pm}$ for machine B is given by:

$$MTTF_{pm} = \frac{\int\limits_{0}^{200} \exp(-(t/1000)^{0.5})dt}{\exp(-(200/1000)^{0.5})} \approx 414 \text{ hours}$$

MTTF for machine B is $\eta \times \Gamma(1 + \frac{1}{B}) = 1000 \times \Gamma(1 + \frac{1}{0.5}) = 2000 \text{ hours}$

Thus for machine B, preventive maintenance will decrease the mean time to failure of the system. Thus, it is better not to apply preventive maintenance for machine B.

3.16.3   Variance of Mean Time To Failure

It is important to know the variance of mean time to failure for better understanding of the item. From definition variance *V(t)* is given by:

$$V(t) = E(t^2) - [E(t)]^2$$
$$= \int\limits_{0}^{\infty} t^2 f(t) - MTTF^2$$

Applying integration by parts:

$$V(t) = 2\int\limits_{0}^{\infty} tR(t)dt - MTTF^2 \tag{3.23}$$

3.17.   MEAN OPERATING TIME BETWEEN FAILURES (MTBF)

MTBF stands for *mean operating time between failures* (wrongly mentioned as *mean time between failures* throughout the literature) and is used as a reliability measure for repairable systems. In British Standard (BS 3527) MTBF is defined as follows:

For a stated period in the life of a functional unit, the mean value of the lengths of time between consecutive failures under stated condition.

MTBF is extremely difficult to predict for fairly reliable items. However, it can be estimated if the appropriate failure data is available. In fact, it is very rarely predicted with an acceptable accuracy. In 1987 the US Army conducted a survey of the purchase of their SINCGARS radios that had been subjected to competitive procurement and delivery from 9 different suppliers. They wanted to establish how the observed Reliability In-service compared to that which had been predicted by each supplier (using MIL-HDBK-217). The output of this exercise is shown in Table 3.6 (Knowles, 1995). It is interesting to note that they are all same radio, same design, same choice of components (but different manufacturers) and the requirement set by the Army was MTBF of 1250 hours with a 80% confidence. Majority of the suppliers' observed MTBF was no where near their prediction.

Table 3.6 SINCGARS radios 217 prediction and the observed MTBF

Vendor	MIL-HDBK-217 (hours)	Observed MTBF (hours)
A	7247	1160
B	5765	74
C	3500	624
D	2500	2174
E	2500	51
F	2000	1056
G	1600	3612
H	1400	98
I	1000	472

Let us assume that the sequence of random variables X_1, X_2, X_3, ...X_n represent the operating time of the item before i-th failure (Figure 3.13). MTBF can be predicted by taking the average of expected values of the random variables X_1, X_2, X_3,..., X_n etc. To determine these expected values it is necessary to determine the distribution type and parameters. As soon as an item fails, appropriate maintenance activities will be carried out. This involves replacing the rejected components with either new ones or ones that have been previously recovered (repaired). Each of these components will have a different wear out characteristic governed by a different distribution. To find the expected value of the random variable X_2 one should take into account the fact that not all components of the item are new and, indeed, those, which are not new, may have quite different ages. This makes it almost impossible to determine the distribution of the random variable X_2 and hence the expected value.
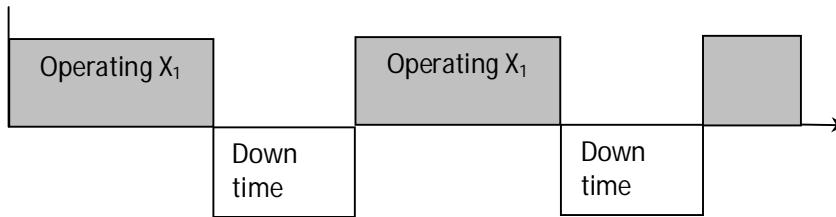
Figure 3.13 operating profile of a generic item

The science of failures has not advanced sufficiently, as yet, to be able to predict failure time distribution in all cases. This is currently done empirically by running a sample of items on test until they fail, or for an extended period, usually under 'ideal' conditions that attempt to simulate the operational environment. Military aircraft-engines, for example, are expected to operate while subjected to forces between -5 and + 9 'g', altitudes from zero to 50000 feet (15000 meters) and speeds from zero to Mach 2+. One has to test the equipment with some new and some old components to find the expected values of the random variables X_2, X_3, etc. In practice most of the testing is done on new items with all new components in pristine condition. The value derived by these type of testing will give the expected value of the random variable X_1 . In practice, the expected value of X_1 is quoted as MTBF. In fact, the expected value of X_1 will give only the Mean Time To First Failure (as the testing is done on new items and the times reflect the time to first failure) and not the MTBF. To calculate MTBF one should consider the expected values of the random variables X_2, X_3, etc.

If the time to failure distribution of the system is exponential then the MTBF can be estimated using the following equation (recommended by MIL-HDBK-217 and DEF-STAN-00-40):

$$MTBF = \frac{T}{n} \tag{3.24}$$

where, T is the total operating period and `n' is the number of failures during this period. Note that the above relation is valid only for large value of T. If n = 0, then MTBF becomes infinity, thus one should be careful in using the above relation. The above expression can be used only when sufficient amount of data is available.

Characteristics of MTBF

1. The value of MTBF is equal to MTTF if after each repair the system is as good as new.
2. MTBF = $1 / \lambda$ for exponential distribution, where λ is the scale parameter (also the hazard function).

Applications of MTBF

1. For a repairable system, MTBF is the average time in service between failures. Note that, this does not include the time spent at repair facility by the system.
2. MTBF is used to predict steady-state availability measures like inherent and operational availability.

3.18.   PERCENTILE LIFE (TTF$_P$ OR B$_{P\%}$)

Percentile life or B$_{p\%}$ is a measure of reliability which is popular among industries. This is the life by which certain proportion of the population (p %) can be expected to have failed. B$_{10\%}$ means the life (time) by which 10% of the items will be expected to have failed. Percentile life is now frequently used among aerospace industries as a design requirement. Mathematically percentile life can be obtained by solving the following equation for t:

$$F(t) = \int_{0}^{t} f(x)dx = p\% \tag{3.25}$$

Assume that *F(t)* is a exponential distribution with parameter λ = 0.05, and we are interested in finding B$_{10}$. Then from above equation we have:

$$1 - \exp(-0.05t) = 0.10 \Rightarrow t = 2.107$$

Thus 2.107 is the B$_{10}$ life for exponential distribution with parameter 0.05. The main application of percentile life lies in prediction of initial spares requirement (*initial spares provisioning, IP).*

System Reliability and Maintainability

Course material

Course Instructor: Professor U Dinesh Kumar
Indian Institute of Management Bangalore

Chapter 4

Systems Reliability

'A Bird is an instrument working according to a mathematical law. It lies within the power of man to make this instrument with all its motion'

Leonardo da Vinci

In this chapter, we present methodologies that can be used to evaluate systems reliability using simple mathematical tools. The chapter discusses two approaches that can be used to predict the reliability metrics of the system. First, we study the models that are based on simple probability theory, assuming that the time-to-failure distributions of different components within the system are known. These models can be used only for non-repairable items. The second approach is based on Markov models, for predicting different reliability measures. The models for repairable items will be discussed using the Markov models. Throughout the Chapter, the word 'system' is used to represent the complete equipment and the word 'item' is used as a generic term that stands for subsystem, module, component, part or unit. Any reliability prediction methodology using time-to-failure approach will involve the following steps:

1. Construct the reliability block diagram (RBD) of the system. This may involve performing failure modes and effect analysis (FMEA).
2. Determine the operational profile of each block in the reliability block diagram.
3. Derive the time-to-failure distribution of each block.
4. Derive the life exchange rate matrix (LERM) for the different components within the system.
5. Compute reliability function of each block.

6. Compute the reliability function of the system.

4.19.   RELIABILITY BLOCK DIAGRAM

Reliability block diagram, RBD, of an item is a logical diagrammatic illustration of the system in which each item (hardware/software) within the system is represented by a block. RBD forms a basis for calculation of system reliability measures. Each block within a RBD can represent a component, subsystem, module or system. The structure of a RBD is determined by the effect of failure of each block on the functionality of the system as a whole. A block does not have to represent physically connected hardware in the actual system to be connected in the block diagram. In an RBD the items whose failure can cause system failure irrespective of the remaining items of the system are connected in series. Items whose failure alone cannot cause system failure are connected in parallel. Depending on the item, a RBD can be represented by a series, parallel, series-parallel, r-out-of-n or complex configuration. Construction of RBD requires functional analysis of various parts within the system. Each block within a RBD should be described using time-to-failure distribution for the purpose of calculating system reliability measures. The RBD can also have network structures (e.g. communication systems, water network and Internet). In the following sections we address how to evaluate various reliability measures for different reliability block diagrams.

4.20.   RELIABILITY MEASURES FOR SERIES
CONFIGURATION

In a series configuration, all the consisting items of the system should be available or functional to maintain the required function of the system. Thus, failure of any one item of the system will cause failure of the system as whole. Series configuration is probably the most commonly encountered RBD in engineering practice. The RBD of a hypothetical system whose items are connected in series is given in Figure 4.1.

```
  ┌─────┐   ┌─────┐             ┌─────┐
──┤  1  ├───┤  2  ├── .. ───────┤ $n^n$├──
  └─────┘   └─────┘             └─────┘
```
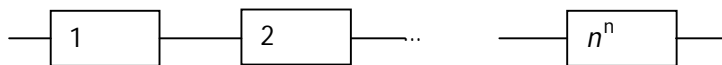
Figure 4.1. Reliability block diagram of a system with series configuration
Reliability function of series configuration

Reliability function of a system with series configuration can be derived from the reliability function of its consisting items. Let $R_s(t)$ represent the reliability function of a series system with n items. Let $R_i(t)$ denote the reliability function of the item *i*. If TTF_i is the time-to-failure random variable for the item *i*, then the reliability function of system for 't' hours of operation is given by:

$$R_S(t) = P[\ TTF_1 \geq t,\ TTF_2 \geq t,\ ....,\ TTF_n \geq t\ ] \tag{4.1}$$

The equation (4.1) clearly states that the system under consideration will maintain the required function if and only if all the n items of the system are able to maintain the required function for at least t hours of operation. Assuming that the random variables TTF_i are independent of each other, the expression (4.1) can be written as:

$$R_s(t) = P[TTF_1 \geq t\,] \times P[TTF_2 \geq t\,] \times\ ...\ \times P[\ TTF_n \geq\ t\,]$$
$$= R_1(t) \times R_2(t) \times ... \times R_n(t)$$

Thus, the reliability of a series configuration with n items is given by:

$$R_s(t) = \prod_{i=1}^{n} R_i(t) \tag{4.2}$$

Note that in the above equation (4.2), it is assumed that the connecting media (such as solder joints) between different items is 100% reliable (unless this is specifically included in the RBD). However, this need not be true. In the equation (4.2) time *t* is used as a generic term. In most case time actually represents age or utilisation of the item under consideration. It can have different units such as hours, miles, landings, cycles etc for different items. One has to normalise the 'time' before calculating the reliability function in such cases. One method of normalising the different life units of the items is using Life Exchange Rate Matrix (LERM), which will be discussed later in this chapter. When the life units of items are different (or different items have different utilisation), we use the following equation to find the reliability of the series system.

$$R_s = P[TTF_1 \geq t_1, TTF_2 \geq t_2,]\ ,TTF_n \geq t_n] = R_1(t_1) \times R_2(t_2) \times]\ \times R_n(t_n)$$

That is,

$$R_S(t) = \prod_{i=1}^{n} R_i(t_i) \qquad (4.3)$$

In equation (4.3), t_i is the age of the item i, which is equivalent to age t of the system. That is, for the system to survive up to age t, the item i should survive up to t_i. Throughout this book we use equation (4.3) unless otherwise specified.

Characteristics of reliability function of a series configuration

1. The value of the reliability function of the system, $R_S(t)$, for a series configuration is less than or equal to the minimum value of the individual reliability function of the constituting items. That is:

$$R_S(t) \le \underset{i=1,2,..n}{Min} \{R_i(t)\}$$

2. If $h_i(t)$ represent the hazard function of item i, then the system reliability of a series system can be written as:

$$R_s(t) = \prod_{i=1}^{n} \exp(-\int_0^t h_i(x)dx$$
$$= \exp(-\int_0^t [\sum_{i=1}^{n} h_i(x)]dx$$

Example 4.1

A system consists of four items, each of them are necessary to maintain the required function of the system. The time to failure distribution and their corresponding parameter values are given in Table 4.1. Find the reliability of the system for 500 and 750 hours of operation.

Table 4.1 Time to failure distribution and their parameter of the items

Item	Time to failure distribution	Parameter values
Item 1	Exponential	$\lambda = 0.001$
Item 2	Weibull	$\eta = 1200$ hours $\beta = 3.2$
Item 3	Normal	$\mu = 800$ hours $\sigma = 350$
Item 4	Weibull	$\eta = 2000$ hours $\beta = 1.75$

SOLUTION:

From the information given in Table 4.1, the reliability function of various items can be written as:

$$R_1(t) = \exp(-0.001 \times t)$$

$$R_2(t) = \exp[-(\frac{t}{1200})^{3.2}]$$

$$R_3(t) = \Phi(\frac{800 - t}{350})$$

$$R_4(t) = \exp[-(\frac{t}{2000})^{1.75}]$$

Since the items are connected in series, the reliability function of the system is given by:

$$R_s(t) = \exp(-0.001 \times t) \times \exp[-(\frac{t}{1200})^{3.2}] \times \Phi(\frac{800 - t}{350}) \times \exp[-(\frac{t}{2000})^{1.75}]$$
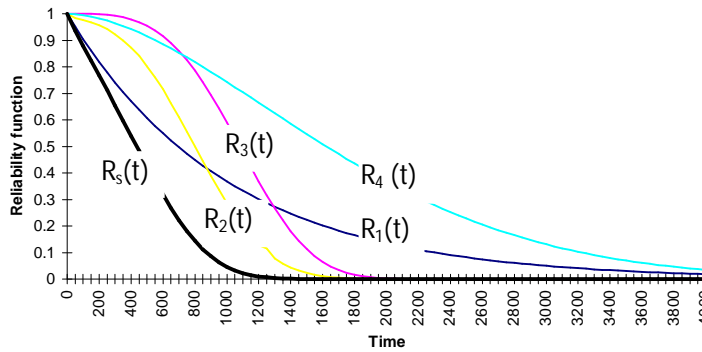
Figure 4.2 Reliability function of the system and its constituent items.

Substituting t = 500 and 750 in the above equation, we get:

R(500) = 0.6065 × 0.9410 × 8043 × 0.9154 = 0.4202

R(750) = 0.4723 × 0.8003 × 0.5568 × 0.8355 = 0.1759

Figure 4.2 shows the reliability function of the system and various items of the system. Note that the system reliability value is always less than or equal to any of the constituting items.

Example 4.2

Avionics system of an aircraft consists of digital auto-pilot, integrated global positioning system, weather and ground mapping radar, digital map display and warning system. Apart from the above items, the avionics system has control software. The time-to-failure distributions of various items are given in Table 4.2. Find the reliability of the avionics system for 100 hours of operation if all the items are necessary to maintain the required function of the avionics system.

Table 4.2 Time-to-failure distribution of various items of the avionics system

Item	Time-to-failure distribution	Parameter values
Digital autopilot	Exponential	$\lambda = 0.003$
Integrated global positioning system	Weibull	$\eta = 1200, \beta = 3.2$
Weather and ground mapping radar	Weibull	$\eta = 1000, \beta = 2.1$
Digital map display	Normal	$\mu = 800, \sigma = 120$
Warning System	Normal	$\mu = 1500, \sigma = 200$
Software	Exponential	$\lambda = 0.001$

SOLUTION:

From the data given in Table 4.2, we can derive the reliability function of various items as follows:

1. Reliability of digital auto-pilot

$$R_1(t) = \exp(-\lambda \times t) \Rightarrow R_1(100) = \exp(-0.003 \times 100) = 0.7408$$

2. Reliability of integrated global positioning system.

$$R_2(100) = \exp(-(t/\eta)^\beta) \Rightarrow R_2(100) = \exp(-(100/1200)^{3.2}) = 0.9996$$

3. Reliability of weather and ground mapping system radar

$$R_3(100) = \exp(-(t/\eta)^\beta) \Rightarrow R_3(100) = \exp(-(100/1000)^{2.1}) = 0.9920$$

4. Reliability of digital map display

$$R_4(100) = \Phi(\frac{\mu - t}{\sigma}) \Rightarrow R_4(100) = \Phi(\frac{800 - 100}{120}) = \Phi(5.8) = 1$$

5. Reliability of warning system

$$R_5(100) = \Phi(\frac{\mu - t}{\sigma}) \Rightarrow R_4(100) = \Phi(\frac{1500 - 100}{200}) = \Phi(7) = 1$$

6. Reliability of software

$$R_6(t) = \exp(-\lambda t) \Rightarrow \exp(-0.001 \times 100) = 0.9048$$

Thus, the reliability of the avionics system for 100 hours of operation is given by:

$$R_s(100) = \prod_{i=1}^{6} R_i(100) = 0.7408 \times 0.9996 \times 0.9920 \times 1 \times 1 \times 0.9048 = 0.6646$$

Hazard function of a series configuration

Let $R_S(t)$ denote the reliability function of the system. From definition, the hazard rate of the system, $h_S(t)$, can be written as:

$$h_S(t) = -\frac{dR_S(t)}{dt} \times \frac{1}{R_S(t)} \tag{4.4}$$

Using equation (4.2), the expression for $R_S(t)$ can be written as:

$$R_S(t) = \prod_{i=1}^{n} R_i(t) = \prod_{i=1}^{n} [1 - F_i(t)] \tag{4.5}$$

where $F_i(t)$ is the failure function of the item i. Differentiating the above expression for reliability function with respect to t, we get:

$$\frac{dR(t)}{dt} = -\sum_{i=1}^{n} f_i(t) \prod_{\substack{j=1 \\ j \neq i}}^{n} [1 - F_i(t)] \tag{4.6}$$

Substituting equation (4.6) in equation (4.4), we get

$$h_S(t) = \sum_{i=1}^{n} \frac{f_i(t)}{R_i(t)} = \sum_{i=1}^{n} h_i(t) \tag{4.7}$$

Table 4.3 Hazard rate of series configuration with n items.

Probability density function of i-th item, $f_i(t)$ (Exponential) $\lambda_i \exp(-\lambda_i t)$	Hazard function of the system, $h_S(t)$ $h_S(t) = \sum_{i=1}^{n} \lambda_i$
(Weibull) $\frac{\beta_i}{\eta_i}(\frac{t}{\eta_i})^{\beta_i-1}\exp(-(\frac{t}{\eta_i})^{\beta_i})$	$h_S(t) = \sum_{i=1}^{n}(\frac{\beta_i}{\eta_i})(\frac{t}{\eta_i})^{\beta_i-1}$
(Normal) $\frac{1}{\sigma_i\sqrt{2\pi}}\exp(-(\frac{1}{2}(\frac{t-\mu_i}{\sigma_i})^2))$	$h_S(t) = \sum_{i=1}^{n} f_i(t)\Big/\Phi(\frac{\mu_i-t}{\sigma_i})$

Thus the hazard function of a series system is given by the sum of the hazard function of individual items. Table 4.3 gives hazard function of a series configuration with n item under the assumption that the time-to-failure of the items follows same distribution but have different parameter.

Figure 4.3 shows hazard rate of a series system with two items where the time-to-failure of individual items follow Weibull distribution.
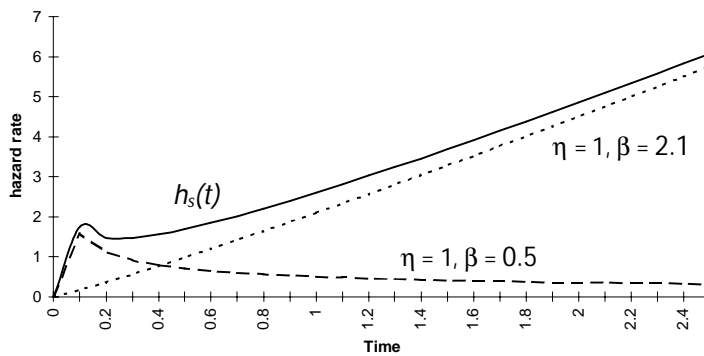
Figure 4.3 Hazard rate of series system with two items with Weibull time-to-failure distribution.

In most cases, the hazard function of a series configuration will be a increasing function. For example, consider a series system with 10 items. Let 9 out of 10 items be identical and have exponential time-to-failure distribution with parameter with rate λ = 0.01. Now we consider two different cases for the time-to-failure distribution of the remaining one item.
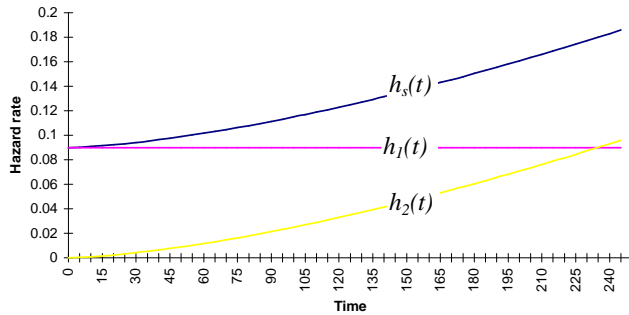
Figure 4.4 Hazard rate the system with 10 items where 9 of them have constant hazard.

Case 1:

Let the time-to-failure of the remaining one item be represented by using Weibull distribution with scale parameter η = 100 and β = 2.5. Now the hazard rate of this system is given by:

$$h_s(t) = 9 \times 0.01 + \frac{\beta}{\eta} (\frac{t}{\eta})^{\beta-1}$$

It is obvious from the above expression that the hazard rate of the system is not constant. Figure 4.4 shows the effect of non-constant hazard function on the system hazard function even when most of the items have constant hazard function. In Figure 4.4, $h_1(t)$ represents the hazard rate for the nine items with exponential time-to-failure and $h_2(t)$ represent the hazard rate of the item with Weibull time-to-failure distribution.

Let the time-to-failure of the remaining one item can be represented by using Weibull distribution with scale parameter η = 100 and β = 0.5. Now the hazard rate of this system is given by:

$$h_s(t) = 9 \times 0.01 + \frac{\beta}{\eta}(\frac{t}{\eta})^{\beta-1}$$

It is obvious from the above expression that the hazard rate of the system is not constant. Figure 4.5 shows the effect of non-constant hazard function on the system hazard function even when most of the items have constant hazard function. In Figure 4.5, $h_1(t)$ represent the hazard rate for the nine item with exponential time-to-failure and $h_2(t)$ represent the hazard rate of the items with Weibull time-to-failure distribution.

Note: The hazard function of complex repairable system may converge to a constant hazard function under certain conditions (mainly under steady-state conditions). This result proved by Drenick (1961) may not be true for today's highly reliable systems. Thus, one has to be very careful in using constant hazard function and thus exponential time to failure for complex systems. This problem will be further discussed in Chapter 8.
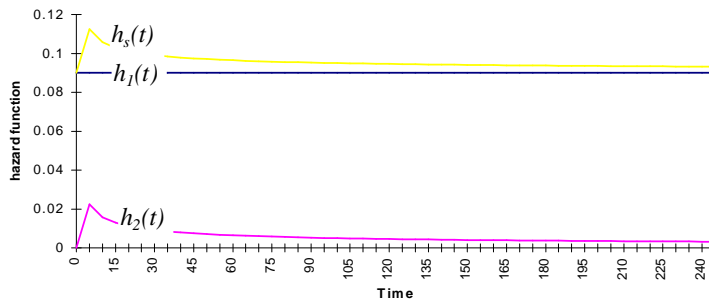
Figure 4.5 Hazard function of the system with 10 items where 9 of them
have constant hazard.

Example 4.3

A system has two items A and B connected in series. The time-to-failure of
item A follows exponential distribution with parameter $\lambda = 0.002$. The time-
to-failure of item B follows Weibull distribution with parameter $\eta = 760$ and
$\beta = 1.7$.   Find the hazard rate of this system at time t = 100 and t = 500.

SOLUTION:

Let $h_A(t)$ and $h_B(t)$ represent the hazard rate of item A and B respectively.
Since the items are connected in series, the hazard rate of the system, $h_S(t)$
is given by:

$$h_S(t) = h_A(t) + h_B(t) = \lambda + (\frac{\beta}{\eta})(\frac{t}{\eta})^{\beta-1} = 0.002 + (\frac{1.7}{760})(\frac{t}{760})^{0.7}$$

Substituting t = 100 and t = 500 in the above equation,

$h_S(100) = 0.00254$

$h_S(500) = 0.0036$

Mean time to failure of a series configuration

The mean time to failure, MTTF, of a series configuration, denoted by
$MTTF_S$, can be written as:

$$MTTF_S = \int_0^\infty R_S\,dt = \int_0^\infty \prod_{i=1}^n R_i(t)dt \qquad\qquad (4.8)$$

The above integral can be evaluated using numerical integration if the
failure distribution is Weibull, normal, lognormal or Gamma. However, in
case of exponential distribution the expression for system $MTTF_S$ can be
obtained as follows. Assume that the time-to-failure distribution of
component i is given by, $1 - \exp(-\lambda_i t)$. Substituting $R_i(t) = \exp(-\lambda_i t)$ in
equation (4.8) we have,

$$MTTF_s = \int_0^\infty \prod_{i=1}^n R_i(t)dt = \int_0^\infty \prod_{i=1}^n \exp(-\lambda_i t)dt = \int_0^\infty \exp(-\sum_{i=1}^n \lambda_i t)dt$$

$$MTTF_s = \frac{1}{\sum_{i=1}^n \lambda_i} \tag{4.9}$$

Thus, the $MTTF_S$ of a series configuration with n items where the time-to-failure of the items are represented by exponential distribution is given by the inverse of the system's hazard function. *Note that this result is true only when the time-to-failure distribution is exponential.* The following equation derived using trapezium approximation of equation (4.8) can be used whenever the time-to-failure of at least one item is non-exponential.

$$MTTF_S \approx \frac{h}{2} \times (R[0] + R[M * h]) + \sum_{i=1}^{M-1} h \times R[i \times h] \tag{4.10}$$

Where h is a small value (e.g. 0.01 or 0.1), the value of M is selected such that $R_S(M \times h)$ is almost zero.

Example 4.4

A system consists of three items connected in series. The time-to-failure distribution and their corresponding parameter values are given in Table 4.4. Find the mean time to failure of the system. Compare the value of $MTTF_S$ with mean time to failure of individual items.

Table 4.4 Time-to-failure distribution of different items

Item	Distribution	Parameter values
Item 1	Weibull	$\eta_1 = 10, \beta_1 = 2.5$
Item 2	Exponential	$\lambda = 0.2$
Item 3	Weibull	$\eta_2 = 20, \beta_2 = 3$

SOLUTION:

Mean time to failure of the system is given by:

$$MTTF_S = \int_0^\infty \prod_{i=1}^3 R_i(t)dt$$

$$= \int_0^\infty \exp(-(\frac{t}{\eta_1})^{\beta_1} \times \exp(-\lambda t) \times \exp(-(\frac{t}{\eta_2})^{\beta_2})dt$$

$$MTTF_S = \int_0^\infty \exp(-(\frac{t}{10})^{2.5}) \times \exp(-0.2t) \times \exp(-(\frac{t}{20})^3)dt$$

Using numerical integration, the $MTTF_S$ is given by:

$MTTF_S \approx 3.48$

Table 4.5 gives the mean time to failure of various items. Note that the mean time to failure of the system is always less than that of the components when the items are connected in series.

Table 4.5 Comparison of $MTTF$ of individual items and $MTTF_S$

Item 1	Item 2	Item 3	System
$MTTF = 8.87$	$MTTF = 5$	$MTTF = 17.86$	$MTTF_S \approx 3.48$

Characteristics of $MTTF_S$ of series system

1. The $MTTF_S \le MTTF_i$, where $MTTF_i$ is the mean time to failure of the item *i*. Thus, the mean time to failure of a system with series RBD will be less than the mean time to time failure of any of its constituting items.

$$MTTF_S \le \underset{i=1,2,...,n}{Min} \{MTTF_i\}$$

Where $MTTF_i$ denote the mean time to failure of the item *i*.

2. For complex repairable systems, *MTTF$_S$*, represents the mean time to first failure.

4.21.   LIFE EXCHANGE RATE MATRIX

Not all the components of the item will have the same utilisation or life unit. In some cases, if the actual mission period is t hours, some items of the system may have to operate more than t hours (in many cases it can be less than *t* hours). An aircraft jet engine will be switched on at least 20 minutes before the actual flight. Thus, for 10 hours flight, the engine may have to operate for more than 10 hours. Operational environment can also change the ageing pattern of different components within a system. For example, the average flight of a domestic flight within Japan is around 30 minutes compared to that of around 3 hours in US. Thus the aircraft used in Japan lands more often than the one in USA. This means that the usage of landing gears, tyres etc of aircraft used in domestic flights in Japan will be much higher than that of USA. It is very common that different items within a system may have different life units such as hour, miles, flying hours, landings, cycles etc. Thus, to find the reliability of a system whose items have different life units it is necessary to normalise the life units. In this section we introduce the concept of life exchange rate matrix, which can be used to describe the exchange rates between various life units.

Life exchange rate matrix (LERM) is a square matrix of size *n*, where *n* is the number of items in the system. Let us denote the life exchange rate matrix as R = [r$_{i,j}$], where r$_{i,j}$ is the (*i,j*) th element in the LERM. Thus, for a system with n items connected in series, the LERM can be represented as:

$$LERM = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,n} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,n} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ r_{n,1} & r_{n,2} & \cdots & r_{n,n} \end{bmatrix}$$

The elements of LERM are interpreted as follows:

r$_{i,j}$ denotes that:

1 life unit of $i = r_{i,j} \times 1$ life unit of j.

Any LERM will satisfy the following conditions:

$$r_{i,i} = 1 \text{ for all } i.$$

$$r_{i,j} = r_{i,k} \times r_{k,j} \text{ for all } i, j, k$$

$$r_{i,j} = \frac{1}{r_{j,i}}$$

As an example, let us consider a system with three items connected in series (Figure 4.6). Let the life unit of items 1, 2 and 3 be hours, miles and cycles respectively.
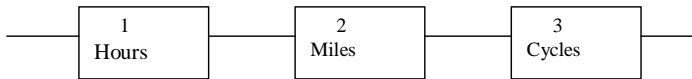
Figure 4.6. Series system with three items where each item has different life units

Assume that:

1 hour = 10 miles

1 hour = 5 cycles

Using the above data, it is easy to construct the life exchange rate matrix for the above system. The LERM for the above matrix is:

$$R = \begin{bmatrix} 1 & 10 & 5 \\ 1/10 & 1 & 0.5 \\ 1/5 & 2 & 1 \end{bmatrix}$$

One can easily verify that the above matrix satisfies all three conditions for a life exchange rate matrix. Using the above matrix, one can easily measure

reliability characteristics in normalised life unit. For the RBD shown in Figure 4.6, reliability of the system for 5 cycles is given by $R_1(1) \times R_2(1) \times R_3(5)$.

Example 4.5

Reliability block diagram of a system consists of three modules A, B and C connected in series.   The time-to-failure of module A follows Weibull distribution with scale parameter η = 100 hours and β = 3.2. The time-to-failure of module B follows Normal distribution with parameter μ = 400 cycles and σ = 32 cycles.   The time-to-failure of module C follows exponential distribution with parameter λ = 0.00015 per mile. It was also noted that, during 1 hour, the module B performs 12 cycles and module C performs 72 miles. Find the probability that the system will survive up to 240 cycles of module B.

SOLUTION:

For the system to survive 240 cycles, module A should survive up to 20 hours and module C should survive up to 1440 miles.

The reliability of individual modules are given by:

$$R_A(t_A) = \exp(-(\frac{t_A}{\eta})^\beta) = \exp(-(\frac{20}{100})^{3.2}) = 0.9942$$

$$R_B(t_B) = \Phi(\frac{\mu - t_B}{\sigma}) = \Phi(\frac{400 - 240}{32}) = 1$$

$$R_C(t_C) = \exp(-\lambda \times t_C) = \exp(-0.00015 \times 1440) = 0.8174$$

The system reliability for 240 cycles is given by:

$$R_S(240) = R_A(20) \times R_B(240) \times R_C(1440) = 0.9942 \times 1 \times 0.8174 = 0.8126$$

4.22.   PARALLEL CONFIGURATION

In a parallel configuration the system fails only when all the items of the system fail. In other words, to maintain the required function only one item

of the system is required to function. The reliability block diagram for a system consisting of items connected in parallel is shown in Figure 4.7
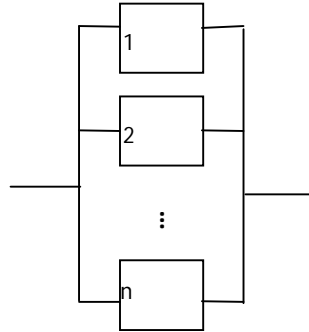
Figure 4.7 Reliability block diagram for a parallel configuration

Parallel components are introduced when the reliability requirements for the system are very high. The use of more than one engine in aircraft is one of the obvious examples of parallel configuration (In practice an aircraft would not be allowed to fly if any of the engine fails. If an engine fails during a flight, the pilot would normally be expected to divert to the nearest airport). However, parallel items will increase cost, complexity and weight of the system. Hence, the number of parallel items required should be carefully determined and if possible optimised.

Reliability function of parallel configuration

Reliability function of a parallel configuration can be obtained using the following arguments. As the system fails only when all the items fail, the failure function, $F_S(t)$, of the system is given by:

$$F_S(t) = P[TTF_1 \leq t, TTF_2 \leq t, ... TTF_n \leq t] \qquad (4.11)$$

where TTF_i represents the time-to-failure random variable of item i. Assuming independence among different items, the above expression can be written as:

$$F_S(t) = F_1(t) \times F_2(t) \times ... \times F_n(t)$$
$$(4.12)$$

where $F_i(t)$ is the time to failure distribution of item i. Substituting $F_i(t)=1 - R_i(t)$ in equation (4.12), the expression for failure function of a parallel configuration can be written as:

$$F_S(t) = [1 - R_1(t)] \times [1 - R_2(t)] \times ... \times [1 - R_n(t)]$$
(4.13)

Now, the reliability function, $R_S(t)$, of a parallel configuration can be written as:

$$R_S(t) = 1 - F_S(t) = 1 - [1 - R_1(t)] \times [1 - R_2(t)] \times ... \times [1 - R_n(t)]$$

or

$$R_S(t) = 1 - \prod_{i=1}^{n}[1 - R_i(t)]$$
(4.14)

Characteristics of a parallel configuration

1. The system reliability, $R_S(t)$, is more than reliability of the any of the consisting items. That is,

$$R_S(t) \geq \underset{i=1,...,n}{Max}\{R_i(t)\}$$

2. If $h_i(t)$ represent the hazard rate of item i, then the reliability function of a parallel configuration can be written as:

$$R_S(t) = 1 - \prod_{i=1}^{n}[1 - \exp(-\int_{0}^{t}h_i(t)dt]$$

Example 4.6
A fly-by-wire aircraft has four flight control system electronics (FCSE) connected in parallel. The time-to-failure of FCSE can be represented by Weibull distribution with scale parameter η=2800 and β = 2.8. Find the reliability of flight control system for 1000 hours of operation.

SOLUTION:

Reliability function for a parallel system with four identical items is given by:

$$R_S(t) = 1 - \prod_{i=1}^{4}[1 - R_i(t)]$$
$$= 1 - [1 - R(t)]^4$$

where $R(t)$ is the reliability function of each item. For $t = 1000$, $R(t)$ is given by:

$$R(t) = \exp(-(t/\eta)^\beta) = \exp(-(1000/2800)^{2.8}) = 0.9455$$

Thus the reliability of flight control system for 1000 hours of operation is given by:

$$R_S(1000) = 1 - [1 - 0.9455]^4 = 0.999991$$

Hazard function of a parallel configuration

Hazard function, $h_S(t)$, of the parallel configuration can be written as:

$$h_S(t) = \frac{-dR_S(t)}{dt} \times \frac{1}{R_S(t)} \qquad (4.15)$$

Substituting the expression for $R_S(t)$ from equation (4.14) in the above equation, we get

$$h_S(t) = \{-\frac{d}{dt}[1 - \prod_{i=1}^{n}(1 - R_i(t))]\} \times \frac{1}{[1 - \prod_{i=1}^{n}(1 - R_i(t))]} \qquad (4.16)$$

It is easy to verify that the above equation can be written as:

$$h_S(t) = \frac{\sum_{j=1}^{n}\{f_j(t) \times \prod_{i=1,i \neq j}^{n} F_i(t)\}}{1 - \prod_{i=1}^{n}[1 - R_i(t)]} \qquad (4.17)$$

Where, $f_i(t)$ is the probability density function of item i.

Example 4.7

For the flight control system electronics discussed in the example 3.5, find the hazard function of the system at time t = 100.

SOLUTION:

Since all the four items are identical, the hazard rate of the system can be written as (using equation (3.15)):

$$h_S(t) = \frac{4 \times f(t) \times [F(t)]^3}{1 - [F(t)]^4}$$

where,

$$f(t) = \frac{\beta}{\eta} (\frac{t}{\eta})^{\beta-1} \exp(-(\frac{t}{\eta})^{\beta})$$

$$F(t) = \exp(-(\frac{t}{\eta})^{\beta})$$

Substituting t = 100, we get

$h_S(t) = 8.0 \times 10^{-8}$

Mean time to failure of parallel configuration

The mean time to failure of a parallel configuration, denoted by $MTTF_S$, can be written as:

$$MTTF_S = \int_0^\infty R_S \, dt = \int_0^\infty \{1 - \prod_{i=1}^n [1 - R_i(t)]\} dt \tag{4.18}$$

For most of the failure distributions one may have to use numerical integration to evaluate the above integral. However, in case of exponential distribution we can get simple expression for system's *MTTF*.

Assume that the time-to-failure distribution of component i is exponential with mean $(1/\lambda_i)$. Then the mean time to failure of the system, $MTTF_S$, is given by:

$$MTTF_s = \int_0^\infty \prod_{i=1}^n R_i(t)dt = \int_0^\infty \{1 - \prod_{i=1}^n [1 - \exp(-\lambda_i t)]\}dt \tag{4.19}$$

For particular values of *n,* we can simplify the above integral to derive the expression for the $MTTF_S$.

Case 1: Assume $n = 2.$ Equation (4.19) can be written as:

$$
\begin{aligned}
MTTF_S &= \int_0^\infty \{1 - [(1 - \exp(-\lambda_1 t)) \cdot (1 - \exp(-\lambda_2 t))]\}dt \\
&= \int_0^\infty [\exp(-\lambda_1 t) + \exp(-\lambda_2 t) - \exp(-(\lambda_1 + \lambda_2)t)]dt \\
&= \frac{1}{\lambda_1} + \frac{1}{\lambda_2} - \frac{1}{\lambda_1 + \lambda_2}
\end{aligned}
$$

Case 2: Assume $n = 3$, the expression for $MTTF_S$ can be written as:

$$
\begin{aligned}
MTTF_S &= \int_0^\infty \{1 - [\prod_{i=1}^3 [(1 - \exp(-\lambda_i t))]\}dt \\
&= \frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} - \frac{1}{\lambda_1 + \lambda_2} - \frac{1}{\lambda_1 + \lambda_3} - \frac{1}{\lambda_2 + \lambda_3} + \frac{1}{\lambda_1 + \lambda_2 + \lambda_3}
\end{aligned} \tag{4.20}
$$

4.23.   R-OUT-OF-N SYSTEMS

In an r-out-of-n (or r-out-of-n:G) system, at least r items out of the total n items should maintain their required function for the system to be operational. Following are few examples of r-out-of-n systems:
1. Control software in a space shuttle has four programs. For the successful completion of the mission, at least three of them should maintain the required function and also the output from at least three programs should agree with each other. This is an example of a 3-out-of-4 system.
2. Most of the telecommunication system can be represented as a r-out-of-n systems.

The reliability function of r-out-of-n system can be derived as stated below.

Reliability function of an r-out-of-n system

Consider an r-out-of-n system with identical items. That is, $R_1(t)=R_2(t)=... = R_n(t)$. Then the system reliability, $R_S(t,r,n)$, is given by:

$$R_S(t,r,n) = \sum_{i=r}^{n} \binom{n}{i} [R(t)]^i [1-R(t)]^{n-i} \qquad (4.21)$$

For the cases when the time-to-failure distribution is exponential or Weibull we have the following expressions for reliability function.

1. Exponential time-to-failure distribution

$$R_S(t,r,n) = \sum_{i=r}^{n} \binom{n}{i} [\exp(-\lambda t)]^i [1-\exp(-\lambda t)]^{n-i}$$

2. Weibull time-to-failure distribution

$$R_S(t,r,n) = \sum_{i=r}^{n} \binom{n}{i} [\exp(-(\frac{t}{\eta})^\beta)]^i [1-\exp(-(\frac{t}{\eta})^\beta)]^{n-i}$$

However, if the items are not identical then one may have to use other mathematical models such as enumeration to evaluate the reliability. For example consider a 2-out-of-3 system with non-identical items. The reliability function of the system can be derived as follows.
Let E_i denote the event that the item i successfully completes the mission (or survives t hours of operation). Then the reliability function for the system can be written as:

$R_S(t) = P [\{E_1 \cap E_2\} \cup \{E_1 \cap E_3 \} \cup \{E_2 \cap E_3 \}]$

By putting, $A = E_1 \cap E_2$, $B = E_1 \cap E_3$ and $C = E_2 \cap E_3$, the above expression can be written as:

$R_S(t,2,3) = P [\{A \cup B \cup C \}]$

$= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C)$
$+ P(A \cap B \cap C)$

$$= P(E_1 \cap E_2) + P(E_1 \cap E_3) + P(E_2 \cap E_3) - 2\,P(E_1 \cap E_2 \cap E_3)$$

Let $R_i(t)$ represent the reliability function for the item i. Now the above expression can be written as:

$$R_S(t, 2, 3) = R_1(t)\,R_2(t) + R_1(t)\,R_3(t) + R_2(t)\,R_3(t) - 2 \times R_1(t)\,R_2(t)\,R_3(t)$$

The above approach becomes complex when the number of items n increases. However, there are several approaches available to tackle complex r-out-of-n systems with non-identical items. The reliability function of r-1-out-of-n and r-out-of-n system with identical items satisfies the following relation:

$$R_s(t, r-1, n) = \binom{n}{r-1} [R(t)]^{r-1} [1 - R(t)]^{n-r+1} + R_S(t, r, n) \qquad (4.22)$$

Mean Time to Failure of r-out-of-n Systems

The mean time to failure, MTTF, of an r-out-of-n system, $MTTF_S(r,n)$, can be obtained using the following expression:

$$MTTF_S(r, n) = \int_0^\infty R_S(t, r, n)\, dt$$

One may have to use numerical integration in most of the cases to evaluate the above integral. However, if the time-to-failure distribution is exponential, then the above integral reduces to a simple expression. For example, consider a 2-out-of-3 system with identical items where the time-to-failure distribution of the item is represented by exponential distribution with parameter λ. The reliability function of 2-out-of-3 system with exponential items are given by:

$$R_S(t) = \sum_{i=2}^{3} \binom{3}{i} [\exp(-\lambda t)]^i [1 - \exp(-\lambda t)]^{n-i}$$
$$= 3\exp(-2\lambda t)(1 - \exp(-\lambda t)) + \exp(-3\lambda t)$$

Now the $MTTF_S$ is given by,

$$MTTF_S = \int\limits_0^\infty [3\exp(-2\lambda t)(1-\exp(-\lambda t)) + \exp(-3\lambda t)]dt$$

$$= \frac{5}{6\lambda}$$

Using equation (4.22), we get the following relation between $MTTF_S(r\text{-}1,n)$ and $MTTF_S (r,n)$ (Misra, 1992):

$$MTTF_S (r-1,n) = \int\limits_0^\infty \binom{n}{r-1} R(t)]^{r-1}[1-R(t)]^{n-r+1}dt + MTTF_s (r,n) \quad (4.23)$$

4.24.   SERIES AND PARALLEL CONFIGURATION

In this Section we discuss two types of series and parallel structures, which have wide application in reliability theory.

Model 1. Series-Parallel Configuration

Here the system has a series structure with n items where each item has parallel redundant items. Assume that item i has m_i components in parallel. Figure 4.8 shows a series-parallel configuration.
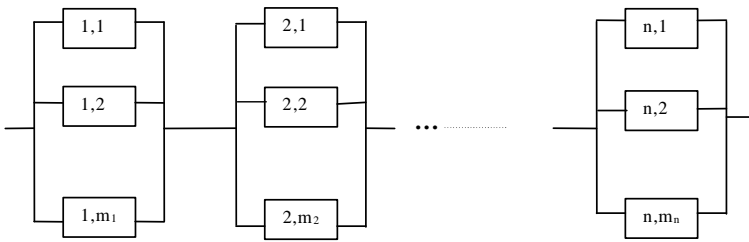
Figure 4.8 Series-parallel structure with n items subsystem where subsystem i has m_i parallel components

In Figure 4.8, (i,j) represent j-th parallel component of the item i. If $R_{i,j}(t)$ denote the corresponding reliability of the component, then the reliability of item i of the system is given by:

$$R_i(t) = 1 - \prod_{j=1}^{m_i}[1-R_{i,j}(t)] \qquad (4.24)$$

Now the system reliability can be written as:

$$R_S(t) = \prod_{i=1}^{n} R_i(t) = \prod_{i=1}^{n}[1 - \prod_{j=1}^{m_i}(1 - R_{i,j}(t))]$$ (4.25)

Model 2. Parallel-Series System
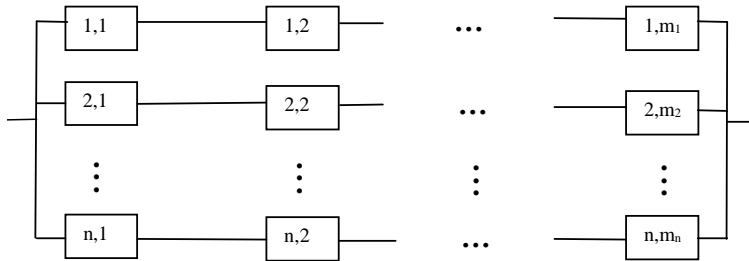
Figure 4.9. Parallel series structure with n sub-systems where subsystem *i* has m$_i$ components

Assume that the system has *n* items connected in parallel where each item has components connected in series. An aircraft with more than one engine, is a typical example for these type of configuration. Figure 4.9 shows parallel-series structure.

Since item *i* has m_i components in series, the reliability of item *i* is given by:

$$R_i(t) = \prod_{j=1}^{m_i} R_{i,j}(t)$$ (4.26)

where $R_{i,j}(t)$ is the reliability function of the component *j* in item *i*. Now the reliability of the parallel-series system is given by:

$$R_S(t) = 1 - \prod_{i=1}^{n}[1 - R_i(t)] = 1 - \prod_{i=1}^{n}[1 - \prod_{j=1}^{n} R_{i,j}(t)]$$ (4.27)

4.25. REDUNDANT SYSTEMS

In systems, redundancy is a means of maintaining system integrity if critical parts of it fail. In some cases this means replicating parts of the system, in others, alternatives are used. A commercial aircraft has to be able to complete a take-off and landing with one of its engines shutdown but, except under very special circumstances, no such aircraft would be allowed to leave the departure gate if any of its engines are not functioning. And yet, ETOPS, extended twin engine operations allows certified twin-engine aircraft (e.g. Boeing 777 and Airbus 330) to fly up to 180 minutes from a suitable landing site. This is based on the probability that even if one of the engines fails that far from land, the other is sufficiently, reliable to make the probability of not reaching a landing site an acceptable risk. It should be noted that in normal flight, i.e. at cruising speed and altitude, the engines are generally doing very little work and usually are throttled back. If an engine fails, it would normally be wind-milled to minimise 'parasitic' drag but, even then, it still offers a considerable resistance and, of course, produces an in-balance which has to be offset by the rudder and other controllable surfaces all of which means the functional engine has to work considerably harder thus increasing its probability of failure.

If the aircraft only had one engine and it failed, the probability of landing safely with no engines is not very high, at least, for fast military jets. In

most cases ultimately, if the engine cannot be re-lit, the only option is to eject after directing the aircraft away from inhabited areas, if there is time. With commercial airlines, neither the pilot, the crew nor the passengers have the option of ejecting or baling out if the aircraft suffers a total engine failure (i.e. all engines fail). These aircraft will glide, to a certain extend but, with no power, none of the instruments will function and, there will be no power assistance for the control surfaces or to deploy the landing gear. For this reason, they are fitted with wind turbines that should drop down and start functioning if there is prolonged loss of power. This gives the pilots some control, but even then, large airliners are not going to rise on a thermal, however good the pilot may be.

A Boeing 767, on one of its first flights, had a total engine failure some 1500 miles from its intended destination, Ontario. All attempts to re-light the engines failed simply because it had run out of fuel. There was a total blackout in the cockpit and, even when the co-pilot managed to find a torch (flashlights) all this showed was that none of the instruments were working (being all digital and computer controlled). The pilot, by pure chance, happened to be an extremely accomplished glider pilot and, again by pure chance, the co-pilot happened to be particularly familiar with this part of Canada, some 200 miles outside Winnipeg. For several minutes the pilot manhandled the controls and managed to stop the aircraft from loosing height too quickly. Eventually the wind turbine deployed which gave them enough power for the instruments, radio and power assisted controls to work again. Unfortunately the aircraft had lost too much height to reach Winnipeg but, it had just enough to get to an ex-military runway (used as a strip for drag racing). There was just enough power to lock the main undercarriage down, but not the nose wheel. The *Gimli Glider* as it became known, landed safely with no serious casualties. But, out of eleven other pilots, who later tried to land the aircraft in the same circumstances on a flight simulator all crashed. Had it not been for the 'redundant' wind turbine, it is almost certain even this experienced glider pilot would have crashed killing all on board.

If the Boeing 777, say, was fitted with three or four Rolls-Royce Trent 800s, Pratt & Whitney 4084s or General Electric GE 90's (instead of the two it currently has) then there would be true redundancy since it needs only two to achieve *ETOPS* (Extended Twin-engine Operations). There are, however, a number of problems with this design. Firstly, it would add very significantly to both weight and drag, to the point where it would seriously reduce the payload and range, probably making the aircraft uneconomical to operate and hence undesirable to the airlines. Secondly such an increase in weight and drag would probably mean the normal two engines would

provide insufficient thrust therefore either more powerful engines would be needed or, the extra engines would have to be used rendering them no longer truly redundant.

On the Boeing 767, for example, the IFSD (In Flight Shut Down) rate after 10 million hours was less than 0.02 per thousand flying hours (the standard measure in the aerospace industry). And, none of these had led to the loss of a single life, let alone an aircraft with its full complement of passengers and crew. It is quite likely that, in some of the instances, flights would have been diverted from their scheduled destinations to alternatives, for safety reasons. The inconvenience to passengers (and airlines) would have cost the airline but, the amount would, almost certainly, have been significantly less than the loss of revenue resulting from the reduced payload had truly redundant engines been fitted.

In many cases, the redundant items may not be functioning simultaneously as in the case of parallel or r-out-of-n configurations. The redundant items will be turned on only when the main item fails. In some cases, the items may be functioning simultaneously but one of them may be sharing much higher load compared to the other. Such types of systems are called standby redundant systems. Whenever the main item fails, a built-in switch senses the failure and switches on the first standby item. It is important that the switch has to maintain its function. Failure of the switch can cause the system failure. The standby redundant systems are normally classified as *cold standby*, *warm standby* and *hot standby.*

Cold Standby System

In a cold standby, the redundant part of the system is switched on only when the main part fails. For example, to meet the constantly changing demand for electricity from the 'National Grid' it is necessary to keep a number of steam turbines ready to come on stream whenever there is a surge in demand. The failure of a generator would result in instantaneous reduction in capacity, which would be rectified by bringing one of these 'redundant' turbines up to full power. In the event of a power cut to a hospital, batteries may switch in instantly to provide emergency lighting and keep emergency equipment, e.g. respirators and monitors running. Petrol and diesel generators would then be started up to relieve the batteries and provide additional power.

In a cold standby system, a redundant item is switched on only when the operating item fails. That is, initially one item will be operating and when this item fails, one item from the redundant items will be switched on to

maintain the function. In a cold standby, the hazard function of the item in standby mode is zero.
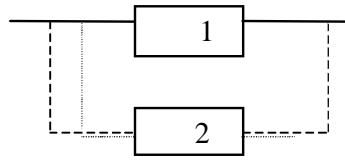
Figure 4.10 Cold standby redundant system

Consider a cold standby system with two identical items (see Figure 4.10). The reliability function of this system can be derived as follows (assuming that the switch is perfect):

$R_S(t)$ = P{The main item survives up to time t}
+ P{The main item fails at time u ($u < t$) and the standby items survives the remaining interval ($t - u$) }

Thus,

$$R_S(t) = R(t) + \int_0^t f(u)R(t-u)du \qquad (4.28)$$

where $f(t)$ is the probability density function of time-to-failure random variable.

As an example consider a cold standby system with two items where the time-to-failure distribution is exponential with parameter λ. Using the equation (4.29) the expression for reliability function is given by:

$$R_S(t) = \exp(-\lambda t) + \int_0^t \lambda \exp(-\lambda u) \times \exp(-\lambda(t-u))du$$
$$= \exp(-\lambda t) + \lambda t \exp(-\lambda t) = \exp(-\lambda t)[1 + \lambda t]$$

For a cold standby system with n identical items with exponential time-to-failure distribution, the expression for reliability function is given by:

$$R_S(t) = \exp(-\lambda t) \sum_{i=0}^{n-1} \frac{(\lambda t)^i}{i!} \tag{4.29}$$

The equation (4.30) is the cumulative distribution of Poisson distribution with mean λt. One can also derive the expression for non-identical standby units using the arguments presented in equation (4.29). For a cold-standby system with non-identical items, the system reliability function is given by:

$$R_s = R_1(t) + \int_0^t f_1(x) R_2(t-x) dx \tag{4.30}$$

Where $R_1(t)$ and $f_1(t)$ are the reliability function and failure density function of item 1 and $R_2(t)$ is the reliability function of item 2. Assume that the time-to-failure items 1 and 2 can be modelled using exponential distribution with mean $(1/\lambda_1)$ and $(1/\lambda_2)$ respectively. Using equation (4.31), the reliability function of cold-standby system with non-identical items is with exponential failure time is given by:

$$R_s(t) = \exp(-\lambda_1 t) + \int_0^t \lambda_1 \exp(-\lambda_1 x) \times \exp(-\lambda_2(t-x)) dx$$

$$R_s(t) = \exp(-\lambda_1(t)) + \frac{\lambda_1}{\lambda_1 - \lambda_2}[\exp(-\lambda_2 t) - \exp(-\lambda_1 t)]$$

The *MTTF* of a cold-standby system can be evaluated by integrating the reliability function between 0 and ∞. The *MTTF* of a cold-standby system with n identical units with exponential failure time is given by:

$$MTTF = \frac{n}{\lambda} \tag{4.31}$$

Equation (4.31) can be easily derived from equation (4.30). For the non-identical MTTF is given by:

$$MTTF_s = \sum_{i=1}^{n} \frac{1}{\lambda_i} \tag{4.31a}$$

Warm Standby System

In a warm standby system, the redundant item will be sharing partial load along with the main item. Thus, in a warm standby, the hazard function of the standby item will be less than that of the main item.

That is, a standby system can deteriorate even when it is not in use. Consider a system with two warm standby items. Assume that $R(t)$ and $R^s(t)$ represent the reliability of the item in operating mode and standby mode respectively. Now the reliability function of the system can be written as:

$$R_S(t) = R(t) + \int_0^t f(x) \times R^s(x) \times R(t-x) du \qquad (4.32)$$

For a particular case where $R(t) = \exp(-\lambda t)$ and $R^s(t) = \exp(-\lambda_s t)$ the reliability function of a warm standby system is given by:

$$R_S(t) = \exp(-\lambda t) + \int_0^t \lambda \exp(-\lambda u) \times \exp(-\lambda_s u) \times \exp(-\lambda(t-u)) du$$

$$= \exp(-\lambda t) + \frac{\lambda \exp(-\lambda t)}{\lambda_s}(1 - \exp(-\lambda_s t))$$

Hot Standby System

In a hot standby, the main item and the standby item will be sharing equal load, and hence will have the same hazard rate. Thus, a hot standby can be treated as a parallel system to derive reliability expressions. If $h_o(t)$ and $h_s(t)$ represent the hazard rate of a operating and standby item respectively. The Table 4.6 gives the various redundancies and the properties of hazard rate.

Table 4.6 Types of standby redundancy and the corresponding properties of hazard rate

Type of Redundancy	Properties of hazard rate

Cold Standby	$h_S(t) = 0$
Warm Standby	$h_O(t) > h_S(t)$
Hot Standby	$h_O(t) = h_S(t)$

4.26.   COMPLEX RELIABILITY BLOCK DIAGRAMS

In many cases, the reliability block diagram will have complex combinations of series and parallel blocks. In such cases, one has to reduce the block to either a series structure or a parallel structure before one can predict the reliability characteristics of the system. Reducing a complex reliability structure will involve the following steps:

1. Replace all purely series (parallel) with an equivalent (reliability wise) single block.
2. Repeat step 1 up till the RBD reduces to either a series or parallel structure.
3. Compute the reliability of resulting RBD.

For example, consider the RBD shown in Figure 4.11.
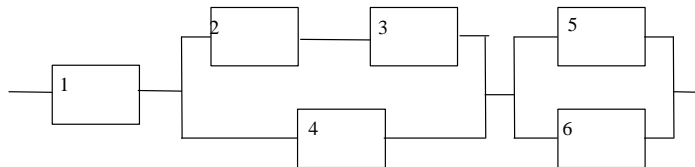
Figure 4.11 Reliability block diagram with combination of series-parallel structures

The time-to-failure of the six items within the system shown in Figure 4.11 are shown in Table 4.7.

Table 4.7. Time-to-failure of items shown in Figure 4.12

Item	Distribution with parameter values

1	Weibull, η = 450 hours, β = 2.4
2	Lognormal μ_l = 4.5, σ_l = 0.75
3	Weibull, η = 890 hours, β = 1.75
4	Exponential, λ = 0.001
5	Normal μ = 800, σ = 120
6	Exponential, λ = 0.00125

The reliability block diagram shown in Figure 4.11 can be evaluated using the three steps explained above. The RBD in Figure 4.11 can be replaced by a series structure with three blocks as shown in Figure 4.11a.
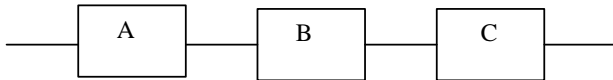
Figure 4.11a Reliability block diagram equivalent to Figure 4.11
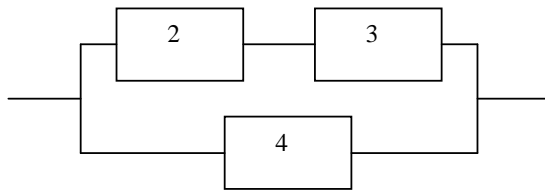
Figure 4.11b RBD equivalent to block B in Figure 4.11

The block A is same as item 1, where block B is equivalent to the RBD shown in Figure 4.11b.
The block B is equivalent to RBD shown in Figure 4.12.c.
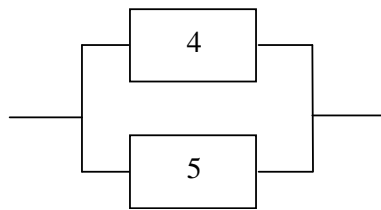
Figure 4.11c. RBD equivalent to block C in Figure 4.11

The expression for reliability function of the system in Figure 4.11 is given by:

$$R_s(t) = R_A(t) \times R_B(t) \times R_C(t)$$

where

$$R_A(t) = R_1(t)$$

$$R_B(t) = 1 - [1 - (1 - R_2(t) \times R_3(t)) \times (1 - R_4(t))]$$

$$R_C(t) = 1 - [1 - (1 - R_5(t)) \times (1 - R_6(t))]$$

For some systems, the reliability block diagram may have more complex configuration than the series/parallel structure as discussed so far. The well-known '*Wheatstone Bridge*' (see Figure 4.12) is an example of such configuration. To find the reliability of such systems one may have to use special tools such as *cut-set, path-set, enumeration* or *the conditional probability approach.* In this Section we illustrate the cut-set approach for evaluating reliability of complex structures.

4.27.   CUT SET APPROACH FOR RELIABILITY EVALUATION

Cut-set approach is one of the most popular and widely used methods for predicting reliability of complex structure. The main advantage of cut-set approach is that it is easy to program and most of the commercial software for reliability prediction use cut-set approach to evaluate the reliability of complex structures. *A cut-set is defined as the set of items that, when failed, will cause the system failure. A cut-set with minimum number of items is called minimal cut set.* That is if any item of the minimal cut set has not failed, then the system will not fail. Mathematically, if the set C is a cut set of the system. Then, the set C will be a minimal cut set if for all $c_i \in C$, $C - c_i$ is not a cut set. Here $C - c_i$ represents the set C without the element c_i. The cut set approach to reliability prediction involves identifying all the minimal cut sets of the system.
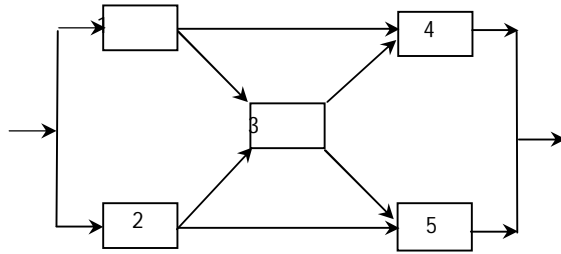
Figure 4.12 Bridge network

In Figure 4.12, the set of items $C = \{1, 2, 3\}$ forms a cut set, since the failure of the items 1, 2 and 3 will cause system failure. However, the set $C = \{1, 2, 3\}$ is not a minimal-cut set since $C - 3 = \{1, 2\}$ still forms a cut set. For the structure shown in Figure 4.12, the minimal cut sets are given by:

$C_1 = \{1, 2\}$, $C_2 = \{1, 3, 5\}$, $C_3 = \{2, 3, 4\}$ and $C_4 = \{4, 5\}$

Since all the elements of the minimal cut set should fail to cause the system failure, each cut set can be considered as a parallel configuration. Thus, the cut sets C_1, C_2, C_3 and C_4 represent the following structures shown in Figure 4.13.
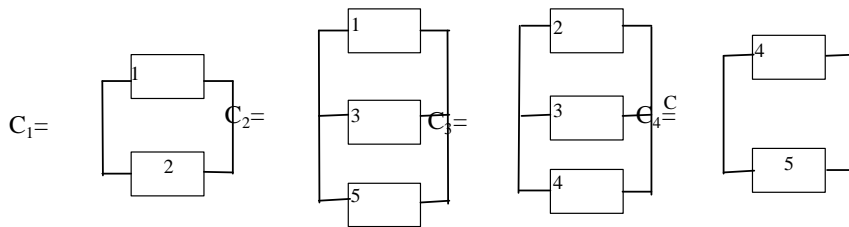
Figure 4.13. Equivalent RBD for minimal cut sets of the system shown in Figure 4.12

Since the system will fail when at least one minimal cut sets fail, the reliability function of the system can be written as:

$$R_S(t) = RC_1(t) \times RC_2(t) \times RC_3(t) \times RC_4(t)$$
(4.33)

where $RC_1(t)$, $RC_2(t)$, $RC_3(t)$ and $RC_4(t)$ are the reliability function of the structures represented by the cut sets C_1, C_2, C_3 and C_4 respectively. If $R_i(t)$ denote the failure function of the items 1, 2, 3, 4 and 5, then we have:

$$RC_1(t) = 1 - F_1(t)F_2(t), \qquad RC_2(t) = 1 - F_1(t)F_3(t)F_5(t)$$
$$RC_3(t) = 1 - F_2(t)F_3(t)F_4(t), \qquad RC_4(t) = 1 - F_4(t)F_5(t)$$

Substituting the above expressions in equation (4.34), we get the failure function for the complex structure shown in Figure 4.12.

In general, cut set approach involves the following steps:

1. Identify all the minimal cut sets of the system.
2. Since all the elements of the minimal cut set should fail to cause the system failure, each cut set can be treated as a parallel configuration.
3. Since failure of any one minimal cut set can cause system failure, different minimal cut sets can be treated as a series configuration.

4.

4.28.   CASE STUDY ON AIRCRAFT ENGINES

Aircraft engine is one of the most critical items used in today's aviation industry. In this section, we try to address several reliability measures one may like to know about an engine. There are totally eleven items including the external gearbox, oil tank and filter. The time-to-failure of these items are given in Table 4.8.

Table 4.8. Time-to-failure distribution of various items of the engine

Item no.	Item	Distribution	Parameter Values
01	LP compressor	Weibull	$\eta = 15\,000, \beta = 3$
02	LP stage 2 stator	Weibull	$\eta = 5\,000, \beta = 2.8$
03	Intermediate casing	Weibull	$\eta = 11\,000, \beta = 3$
04	HP compressor	Weibull	$\eta = 12\,000, \beta = 3.5$
05	HP NGV	Weibull	$\eta = 8\,000, \beta = 3$
06	HP turbine	Weibull	$\eta = 25\,000, \beta = 4$
07	LP NGV	Weibull	$\eta = 7\,000, \beta = 2.2$
08	LP turbine	Weibull	$\eta = 20\,000, \beta = 2.8$
09	Exhaust mixer	Weibull	$\eta = 7\,000, \beta = 3$
10	External gear box	Weibull	$\eta = 6\,500, \beta = 3$
11	Oil tank and filter	Weibull	$\eta = 5\,000, \beta = 3.8$

We are interested in carrying out the following tasks

1. Draw the reliability block diagram of the engine.
2. Find reliability of the engine for 3000 hours of operation.

3. Find the hazard rate of the engine at $t = 3000$ and $t = 7000$ hours.
4. Find the MTTF of different items of the engine and estimate the MTTF of the engine from the MTTF values of the items.
5. Find the MTTF of the engine if all the items are subject to preventive maintenance after every 1000 hours of operation (assume that after maintenance all the items behave as good as new).
6. For an engine of age 5000 hours, find the mission reliability for 1000 hours of operation.
7. Find the MFOPS of the engine for 500 hours of operation for different cycles.

SOLUTION:

1. Since all the item of the engine must maintain their function, the system will have a series configuration as shown below:

| LP compressor | LP stage 2 stator | Intermediate casing | HP compressor |

| HP NGV | HP turbine | LP NGV | LP Turbine |

| Exhaust mixer | External gearbox | Oil tank and filter |
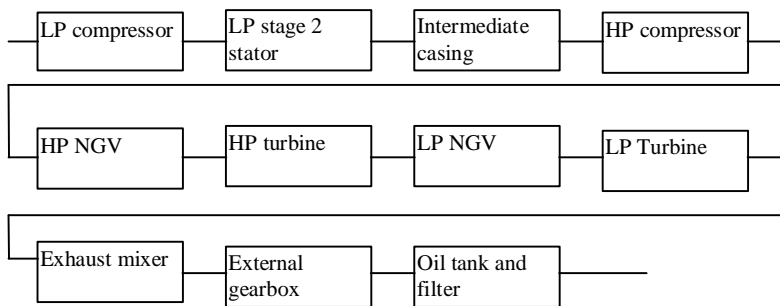
Figure 4.14 Reliability block diagram of the engine

2. Since all the items of the system follow Weibull distribution, the reliability function for each of these items is given by:

$$R(t) = \exp(-(\frac{t}{\eta})^{\beta})$$

Substituting the values of η and β for various items in the above equation, the reliability of various items for 3000 hours of operation is given by:

1. Reliability of LP compressor for 3000 hours of operation is given by:

$$R_1(3000) = \exp(-(\frac{3000}{15000})^3) = 0.9920$$

2. Reliability of LP stage 2 stator for 3000 hours of operation is given by:

$$R_2(3000) = \exp(-(\frac{3000}{5000})^{2.8}) = 0.7872$$

3. Reliability of intermediate casing for 3000 hours of operation is given by:

$$R_3(3000) = \exp(-(\frac{3000}{11000})^3) = 0.9799$$

4. Reliability of HP compressor for 3000 hours of operation is given by:

$$R_4(3000) = \exp(-(\frac{3000}{12000})^{3.5}) = 0.9922$$

5. Reliability of HP NGV for 3000 hours of operation is given by:

$$R_5(3000) = \exp(-(\frac{3000}{8000})^3) = 0.9486$$

6. Reliability of HP turbine for 3000 hours of operation is given by:

$$R_6(3000) = \exp(-(\frac{3000}{25000})^4) = 0.9997$$

7. Reliability of LP NGV for 3000 hours of operation is given by:

$$R_7(3000) = \exp(-(\frac{3000}{7000})^{2.2}) = 0.8563$$

8. Reliability of LP turbine for 3000 hours of operation is given by:

$$R_8(3000) = \exp(-(\frac{3000}{20000})^{2.8}) = 0.9950$$

9. Reliability of exhaust mixer for 3000 hours of operation is given by:

$$R_9(3000) = \exp(-(\frac{3000}{7000})^3) = 0.9243$$

10. Reliability of external gearbox for 3000 hours of operation is given by:

$$R_{10}(3000) = \exp(-(\frac{3000}{6500})^3) = 0.9063$$

11. Reliability of oil tank and filter for 3000 hours of operation is given by:

$$R_{11}(3000) = \exp(-(\frac{3000}{5000})^{3.8}) = 0.8662$$

Using the above values of individual reliabilities, the reliability of the system is given by
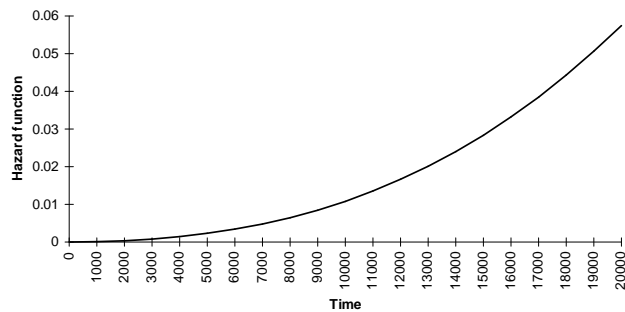
$$R_S(3000) = \prod_{i=1}^{11} R_i(3000) = 0.4451$$

Figure 4.15 hazard function for the engine.

3. Hazard function of the system.

Since all the items of the system follow Weibull time-to-failure, the hazard function is given by:

$$h(t) = (\frac{\beta}{\eta})(\frac{t}{\eta})^{\beta-1}$$

The system hazard function is given by:

$$h_S(t) = \sum_{i=1}^{11} h_i(t)$$

It is easy to verify that the hazard function of the system at t = 3000 and t= 7000 is given by:

$$h_S(3000) = 0.000791 \text{ and } h_S(7000) = 0.004796$$

Figure 4.15 depicts the hazard function for the engine.

4. The expression for MTTF is given by:

$$MTTF = \eta \times \Gamma(1 + \frac{1}{\beta})$$

By substituting the values of η and β, one can find the *MTTF* of different items. Table 4.9 gives the *MTTF* of different items.

Table 4.9 MTTF of different item of the engine

Item	*MTTF* (in hours)
LP compressor	13 395
LP stage 2 stator	4 450
Intermediate casing	9 823
HP compressor	10 800
HP NGV	7 144
HP turbine	22 650

LP NGV	6 202
LP turbine	17 800
Exhaust mixer	6 251
External gear box	5 804
Oil tank and filter	4 525

Since the lowest MTTF is 4 450 (LP stage 2 stator), the MTTF of engine will be less than 4 450.

5. Mean time to failure of a system subject to preventive maintenance is given by:

$$MTTF_{pm} = \frac{\int_0^{T_P} R_S(t)\,dt}{1 - R_S(T_P)}$$

It is given that the engine is subject to preventive maintenance every 1000 hours of operation. Thus, T_P = 1000 hours. The above expression can be evaluated using numerical integration. The approximate values of $MTTF_{pm}$ is:

$$MTTF_{pm} = \frac{\int_0^{1000} R_S(t)\,dt}{1 - R_S(1000)} \approx \frac{999.06}{0.0369} \approx 27,075$$

6. The mission reliability of the engine is given by:

$$MR(t_b, t_m) = \frac{R(t_b, t_m)}{R(t_b)}$$

where t_b is the age of the item at the beginning of the mission and t_m is the mission duration. Substituting t_b = 5000 and t_m = 1000, we have

$$MR(t_b, t_m) = \frac{R(5000+1000)}{R(5000)} = \frac{\prod\limits_{i=1}^{11} R_i(6000)}{\prod\limits_{i=1}^{11} R_i(5000)} = \frac{0.0013}{0.02369} = 0.0548$$

7. The maintenance free operating period survivability, MFOPS, for the engine described is given by:

$$MFOPS(t_{mf}) = \frac{R_S(i \times t_{mf})}{R_S([i-1] \times t_{mf})} = \frac{\prod\limits_{i=1}^{11} R_i(i \times t_{mf})}{\prod\limits_{i=1}^{11} R_i([i-1] \times t_{mf})}$$

The above equation can be evaluated for $t_{mf} = 500$ and for i = 1, 2, ... etc. Figure 4.16 shows the MFOPS values for different cycles (note that these values are derived without considering maintenance recovery period MRP).
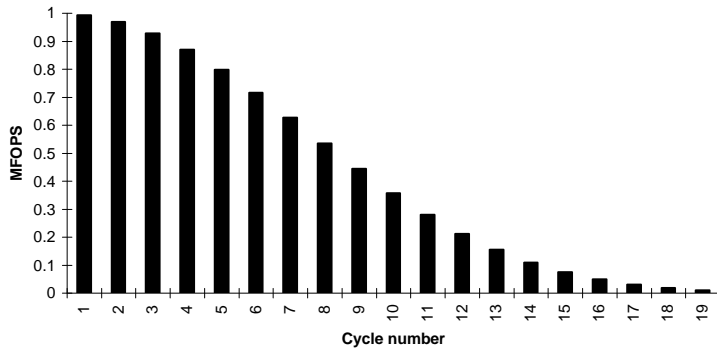
Figure 4.16 MFOPS value for different cycles for the engine

Chapter 5

Maintainability and Maintenance

*Maintenance is the management of failures and
the assurance of availability*

J Hessburg

Maintainability and maintenance has always been important to the industry as it affects the performance as well as the finance. For commercial airlines, maintenance costs around 10% of the airlines total cost, as much as fuel and travel agents' commission (M Lam, 1995). Operators/users would like their system to be available and safe to operate when required. One should be lucky to find a smiling customer when the system fails and it takes a long time to recover the functionality.

There are several ways that designers can provide maximum utility of their product. One way is to build items/systems that are extremely reliable (and consequently will, almost certainly, have a higher acquisition cost). Another is to design systems that are quick and easy to repair when they fail. Obviously, the main objective of the designer is to provide a reliable and safe item at an affordable price.

Maintenance is the action necessary to sustain and restore the performance, reliability and safety of the item. The main objective of maintenance is to assure the availability of the system for use when required. For aircraft, maintenance forms an essential part of airworthiness. The common objective of aircraft maintenance, civil or military, is to provide a fully serviceable aircraft when it is required by the operator at minimum cost (Knotts, 1996). However, maintenance costs money. The annual maintenance cost of production assets in the United Kingdom is estimated in excess of $13 billion, with $2 billion wasted through inefficient maintenance management practices (Knotts, 1999).

Maintenance also accounts for approximately 10% of the organisations' employees and at least 10-15% of its operating costs.

29. CONCEPT OF MAINTAINABILITY

In the previous chapters, we showed that it is important for the operator/user to know the reliability characteristics of the item. We also recognised that it is almost impossible for any item to maintain its function forever, as failure and the degradation of performance is inevitable. Thus, for the user it is equally, or even more important to know:

- When and how often maintenance tasks should be performed
- How they should be performed
- How many people will be needed
- What skills they will need and how much training
- How much the restoration will cost
- How long the system will be down
- What facilities and equipment (special and general) will be required.

All the above information is important as it affects the availability and the life cycle cost of the system. One has to apply a scientific discipline to find answers to these questions.

Maintainability is the scientific discipline that studies complexity, factors and resources related to the maintenance tasks needed to be performed by the user in order to maintain the functionality of a system, and works out methods for their quantification, assessment, prediction and improvement.

Maintainability Engineering is rapidly growing in importance because it provides a very powerful tool to engineers for the quantitative description of the inherent ability of their system/product to be restored by performing specified maintenance tasks. It also contributes towards the reduction of maintenance costs of a system during its utilisation to achieve optimum life cycle cost.

The maintainability engineering function involves the formulation of an acceptable combination of design features, which directly affect maintenance and system support requirements, repair policies, and maintenance resources. Some physical design features such as accessibility, visibility, testability, complexity and interchangeability affect the speed and ease with which maintenance can be performed.

Maintainability studies have the following objectives (R Knotts 1996):

- To guide and direct design decisions
- To predict quantitative maintainability characteristics of a system
- To identify changes to a system's design needed to meet operational requirements

In the technical literature, several definitions for maintainability can be found. For example, the US Department of Defence's MIL-STD-721C (1966) defines maintainability as:

The measure of the ability of an item to be retained in or restored to specified condition when maintenance is performed by personnel having specified skill levels, using prescribed procedures and resources, at each prescribed level of maintenance and repair.

Maintainability can be expressed in terms of maintenance frequency factors, maintenance elapsed times and maintenance cost. Maintainability therefore is an inherent design characteristic dealing with the ease, accuracy, safety, and economy in the performance of maintenance functions. Maintainability requirements are defined in conceptual design as part of system operational requirements and the maintenance concept. Anon (1992) describes maintainability as:

The characteristic of material design and installation that determines the requirements for maintenance expenditures including time, manpower, personnel skill, test equipment, technical data and facilities to accomplish operational objectives in the user's operational environment.

One of the common misperceptions is that maintainability is simply the ability to reach a component to perform the required maintenance task (accessibility). Of course, accessibility is one of the main concerns for many maintenance engineers. Figure 5.1 illustrates an accessibility problem in one of the older twin-engine fighter aircraft, Gloster Javelin. Before an engine could be changed, the jet pipe had to be disconnected and removed. To remove the jet pipe it was necessary for a technician to gain access through a hatch and then be suspended upside down to reach the clamps and pipes which had to be disconnected. The job

could only be achieved by touch; the items were outside of the technician's field of view. The technician had to work his way down between the engine and the aircraft's skin, with tools in his hand. For safety reasons, he was held by his ankles, as shown in figure 5.1 (source: R Knotts).

However, there are many other aspects to be considered other than accessibility. Maintainability should also consider factors such as visibility, that is the ability to see a component that requires maintenance action, testability (ability to detect system faults and fault isolation), simplicity and interchangeability. Additionally decision-makers have to be aware of the environment in which maintainers operate.   It is much easier to maintain an item on the bench, than at the airport gate, in a war, amongst busy morning traffic, or in any other result-oriented and schedule-driven environment.
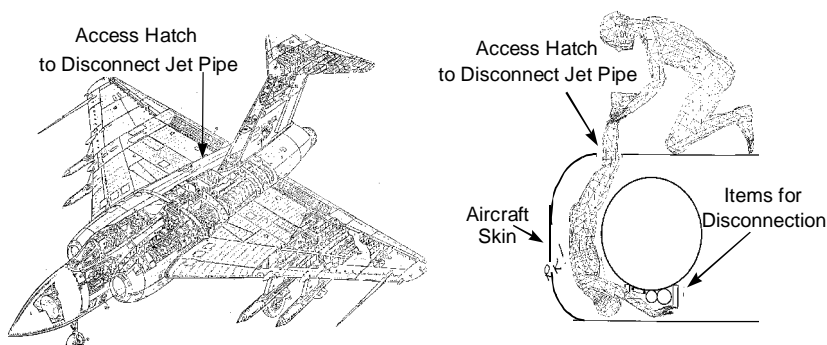
Figure 5.1 Accessibility concern in the Javelin fighter aircraft

Another area to be considered under maintainability is troubleshooting the various modules within the allowed time, i.e. determining whether the system is safe to operate and, if not, what action is needed. For the commercial airlines, there is usually less than an hour at the gate prior to the aircraft's departure to the next destination, whereas for a racing car or weapon system every second could be vital.

To meet these requirements, an easily manageable device is needed which can diagnose with a high degree of accuracy, which modules within the system are at fault. It is now widely accepted that false removals (often referred to as *No Fault Found – NFF)* cost about the same as an actual failure when the component under investigation is removed and replaced.

Reducing the number of false removals, therefore, would be a big cost saver.

Devices with these capabilities have been developed in the aerospace, Formula 1 racing car and luxury car industries. For example, the Boeing 777 includes an 'on-board maintenance system' with the objective to assist the airlines to avoid expensive gate delays and flight cancellations. For similar purposes the Flight Control Division of the Wright Laboratory in the USA has developed a fault detection/isolation system for F-16 aircraft, which allows maintainers, novice as well as expert, to find failed components.

In the next section, we discuss the maintainability measures and how these measures can be used for effective maintenance management.

30.   MEASURES OF MAINTAINABILITY

It is extremely important for the user to have information about the functionality, cost, safety, and other characteristics of the product under consideration at the beginning of its operating life. However, it is equally, or even more important to have information about the characteristics with which to define the maintenance time. Measures of maintainability are related to the ease and economy of maintenance such as; elapsed time that an item spends in the state of failure, man-hours required completing a maintenance task, frequency of maintenance, and the cost of maintenance. As the elapsed time has a significant influence on the availability of the system, operators would like to know the maintenance times; not just the mean time but also the probability that a maintenance task will be completed within a given time. Maintenance elapsed times are even advertised as a marketing strategy.

30.1   Maintenance Elapsed-Time

The length of the elapsed time, required for the restoration of functionality, called t*ime to restore,* is largely determined at an early stage of the design phase. The maintenance elapsed time is influenced by the complexity of the maintenance task, accessibility of the items, safety of the restoration, testability, physical location of the item, as well as the decisions related to the requirements for the maintenance support resources (facilities, spares, tools, trained personnel, etc). It is therefore a function of the maintainability and supportability of the system. It will, of course, also be influenced by other factors during the various stages of the life of the system but any bad decision made (either explicitly or by default) during

the design stage will be costly to rectify at a later stage and will significantly affect both the operational costs and system availability.

1. Personnel factors which represent the influence of the skill, motivation, experience, attitude, physical ability, self-discipline, training, responsibility and other similar characteristics related to the personnel involved;
2. Conditional factors which represent the influence of the operating environment and the consequences of failure with the physical condition, geometry, and shape of the item under restoration;
3. Environmental factors which represent the influence of factors such as temperature, humidity, noise, lighting, vibration, time of the day, time of the year, wind, noise, and others such as those similar to the maintenance personnel factors during restoration.

This maintainability measure can be represented using the probability that the maintenance task considered will be completed by a stated time. Since the maintenance elapsed time is a random variable, one can use the cumulative distribution function of the elapsed time to find the percentage of maintenance tasks that will be completed within a specified time.

Mean Time to Repair

One approach for measuring maintainability is through Mean Time to Repair (MTTR). MTTR is the expected value of the item's repair time. With the knowledge of the reliability and maintainability of the sub-systems one can evaluate the maintainability of the system, that is, mean time to repair of the system, $MTTR_s$ (Birolini, 1994).

Assume that the reliability block diagram of the system has a series structure with n items *with no redundancy.* Let $MTTF_i$ and $MTTR_i$ be the mean time to failure and mean time to repair of sub-system *i* in the system. Consider an arbitrarily large operating time *T*. Assuming that the failure rate of the unit is constant, the expected number of failures of unit *i* in during *T* is given by:

$$\frac{T}{MTTF_i} \tag{5.1}$$

The mean of total repair time to repair unit *i* during *T* is given by:

$$MTTR_i \frac{T}{MTTF_i} \qquad\qquad (5.2)$$

For the whole system, the mean number of failures is given by:

$$\sum_{i=1}^{n} \frac{T}{MTTF_i} \qquad\qquad (5.3)$$

For the whole system, the mean of total repair time is given by:

$$\sum_{i=1}^{n} MTTR_i \times \frac{T}{MTTF_i} \qquad\qquad (5.4)$$

Combining equation (5.3) and (5.4), we get the mean time to repair at the system level, $MTTR_s$, as:

$$MTTR_s = \frac{\sum\limits_{i=1}^{n} \dfrac{MTTR_i}{MTTF_i}}{\sum\limits_{i=1}^{n} \dfrac{1}{MTTF_i}} \qquad\qquad (5.5)$$

Assuming constant failure rate, that is,

$$\lambda_i = \frac{1}{MTTF_i} \text{ and } \lambda_s = \sum_{i=1}^{n} \lambda_i \text{, equation (5.5) can be written as:}$$

$$MTTR_s = \sum_{i=1}^{n} \frac{\lambda_i}{\lambda_s} MTTR_i \qquad\qquad (5.6)$$

Example 5.1

The MTTF and MTTR of four sub-systems in a system are given in Table 5.1. Estimate the system level mean time to repair, $MTTR_s$.

Sub-system	MTTF	MTTR
1	200	24
2	500	36

| 3 | 340 | 12 |
| 4 | 420 | 8 |

SOLUTION:

Applying equation (5.5), we get:

$$MTTR_s = \frac{\dfrac{24}{200} + \dfrac{36}{500} + \dfrac{12}{340} + \dfrac{8}{420}}{\dfrac{1}{200} + \dfrac{1}{500} + \dfrac{1}{340} + \dfrac{1}{420}} \approx 20 \text{ hours.}$$

Mean Time to Repair – Multi-Indenture Case

Many complex systems are broken down into a number of levels of indenture (LoI). For these systems, recovery of an LoI_i unit is usually achieved by the removal and replacement of LoI_{i+1} items. In many cases, the replacement LoI_{i+1} item will not be the item that was removed. It may be a new (i.e. unused) one or it may be one that was removed from another LoI_i unit and subsequently recovered and put into stock for such an occasion.

Now, for such a system, the time to repair will be the time to remove and refit the units at the next lower level of indenture. The elapsed time will need to take into account logistic delays (i.e. waiting for equipment, personnel, spares and any transport to and from the site at which the maintenance work is to be done). This is discussed in more detail in Chapter 10.

Suppose a system is made up of n levels of indenture and a unit at LoI_i is made up of m_i LoI_{i+1} items. Suppose also that to recovery an LoI_i unit, one of the m_i items is removed and replaced with average times, $MTTRM_{i,j}$ and $MTTRP_{i,j}$ respectively. Let us assume that the probability that item j is rejected given that unit i has been removed is $P_{i,j}$ then over an arbitrarily long operating time T, the expected number of system failures is:

$$\frac{T}{MTTF_1}$$

where $MTTF_1$ is the mean time between failures of the system (over time T).

Now, the probability that the failure was due to sub-system j is $P_{i,j}$ so the mean time between failures due to sub-system j is

$$MTTF_{1,j} = \frac{1}{\lambda_{1,j}} = \frac{1}{P_{1,j}\lambda_1} = \frac{MTTF_1}{P_{1,j}}$$

Assuming the system reliability block diagram is series and is series and there are no redundancies.

The expected number of failures of sub-system j is

$$P_{1,j}\frac{T}{MTTF_1} = \frac{T}{MTTF_{1,j}} = \lambda_{1,j}T$$

The expected time to recover the system given that sub-system j is the cause of its failure is

$$MTTR_{1,j} = MTTRM_{1,j} + MTTRP_{1,j}$$

The expected total time spent recovery the system due to sub-system j failures over time T is then

$$P_{1,j}\frac{MTTR_{1,j}}{MTTF_1}T = \frac{MTTR_{1,j}}{MTTF_{1,j}}T = \lambda_{1,j}MTTR_{1,j}T$$

So, the expected total time spent recovering the system by sub-system exchange is

$$\sum_{j=1}^{m_1}P_{1,j}\frac{MTTR_{1,j}}{MTTF_1}T = \sum_{j=1}^{m_1}\frac{MTTR_{1,j}}{MTTF_{1,j}}T = \sum_{j=1}^{m_1}\lambda_{1,j}MTTR_{1,j}T$$

Where m_1 is the number of sub-systems. Then the mean time to recover the system (by sub-system exchange per system failure) is

$$MTTR_{1,E} = \sum_{j=1}^{m_1}\frac{\lambda_{1,j}}{\lambda_1}MTTR_{1,j}$$

To determine the total maintenance time, we would have to look at the time spent recovering the sub-systems, by sub-sub-system exchange and so on down to the lowest level components that are recovered in this way and

then add on any time spent repairing the lowest level components (parts) if they can be repaired but we will leave this exercise until our next book.

30.2   **Maintenance Man Hour (MMH)**

Although elapsed time is an extremely important maintenance measure, one must also consider the maintenance man-hours, *MMH* (also known as maintenance labour hours). The MMH is an estimate of the expected "spanner-in-hand" time and takes into account all of the maintenance tasks and actions required for each system, sub-system or component recovery. It should be noted that the MMH can be considerably greater than the elapsed time as it is often possible and sometimes even necessary to employ more than one person on any given activity or task.

"Work study" and "time and motion" exercises have generated tables of times for every conceivable maintenance action, from releasing the catches that are used on access panels to inspecting the blades on a turbine using a boroscope to drilling out a stud that has sheared after too much torque has been applied to it, to disconnecting and reconnecting all of the pipes and leads when removing and replacing an engine.

In most cases, these times are based on carrying out these tasks and actions in ideal conditions, i.e. in a properly lit workshop, which is heated and provides shelter from the elements. They are generally done when the components are in pristine condition free from contamination, corrosion or damage. It is also generally assumed that the mechanic carrying out each action will have been properly trained and familiar with the correct procedures. In practice, however, it is very rare for all of these ideal conditions to be met so, the actual times will inevitably be longer than those used in the MMH prediction.

Maintenance man-hours are useful in their own right but very often they are given as a "rate" such as (*MMH/*operating hour), (*MMH/*cycle), (*MMH/*month), and (*MMH/*maintenance task). For example, elapsed times can be reduced (sometimes) by increasing the number of people involved in accomplishing the specific task. However, this may turn out to be an expensive trade-off, particularly when high skill levels are required to perform the tasks. Also, unless it actually requires more than one person to do the job, there is likely to be an "interference factor" which means that the efficiency of each person is reduced. Therefore, a proper balance among elapsed time, labour time, and personnel skills at a minimum maintenance cost is required.

Commercial airlines and air forces use the measure *Maintenance Man-Hour per Flight Hour* (MMH / FH) as an indicator of the maintainability of the

aircraft for comparison with other similar aircraft either of an older generation or made by another manufacturer. This measure may be used to decide between alternatives although, in many cases, it will be used to exert pressure on the manufacturer to make improvements. The following expression can be used to evaluate the MMH/FH:

$$MMH/FH = \frac{N_1(t) \times MPMT \times MNC_{pm} + N_2(t) \times MCMT \times MNC_{cm}}{Total\ flying\ hours}$$

(5.7)

Where:

$N_1(t)$ is the total number of preventive maintenance tasks during t hours, and $N_2(t)$ is the total number of corrective maintenance tasks. The value t should be equal to the operational life of the aircraft.
$MPMT$ = Mean preventive maintenance time.
$MCMT$ = Mean corrective maintenance time.
MNC_{pm} = Mean number of crew for preventive maintenance.
MNC_{cm} = Mean number of crew for corrective maintenance.
Note that these estimated mean values should be weighted according to the expected frequency of each maintenance task as we did when calculating $MTTR_s$ above.
A problem with estimating the MMH/FH metric is that it relies on the reliability of the various components of the system, which may be age-related and will, inevitably, depend on the maintenance and support policies. For these reasons, the MMH/FH may not remain constant with aircraft age. The implication of using such a metric is that it is preferential for it to be minimised, however, it may actually be both cheaper and yield a higher level of availability if more time is spent on maintenance, particularly preventative maintenance.

30.3 Maintenance Frequency Factors

Maintainability engineering is primarily concerned with designing a system so that it spends a minimum time in maintenance, given that it needs maintaining. Another characteristic of system design pertaining to maintainability is in optimising the mix between preventative and corrective maintenance.
The ideal system design would allow the operators to use the system until just before it fails but, with enough notice of the impending failure so that

the operator can choose to perform the necessary maintenance at the most opportune moment. In all but a few cases, prognostics have, as yet, not reached this level of sophistication. An alternative approach is built-in redundancy and fault-tolerant systems. These allow the operators to defer maintenance for a limited period or, in certain circumstances until the backup system fails.

Corrective maintenance can be expensive if the failure causes damage to other parts of the system or if it stops the system from earning its keep. However, redundant components will also add to the cost of the system and may reduce its load-carrying capacity. The spare wheel in cars takes up space that could otherwise be used for carrying luggage, it also increases the gross weight, which will reduce the performance of the car both by reducing its rate of acceleration and increasing the fuel consumption.

It is common practice for motorists to replace tyres before the tread has been completely worn away because it is unsafe to drive on bald tyres. It is also illegal and the penalties can be both expensive and inconvenient. It is also very easy to inspect tyres for wear so it is possible to leave them until the "last minute" or get them replaced when the car is not needed thus minimising the inconvenience or lack of availability.

Brake pads are more difficult to inspect by the owner. As a result, many cars are now fitted with pads that have an in-built electrode, which causes a warning light to be illuminated on the dashboard when it comes into contact with the metallic disc (due to the non-conductive part of the pad being worn away). this generally gives the driver a sufficient warning for him or her to find out what the warning light means and take the necessary corrective action before the brakes become dangerous.

Most motorists have their cam or timing belts replaced within about 1000 miles of the manufacturer's recommended mileage possibly during a routine service (scheduled maintenance) or at the driver/owner's convenience. In this case, the owner has almost certainly no way of knowing how much longer the belt will last and, indeed, it is likely to cost them almost as much to have the belt inspected as it would to have it replaced because of the amount of work involved. In this case, the extent of the damage to the engine if the belt breaks is likely to cost a great deal more than that of replacing the belt early. It would no doubt be possible to devise a monitor that could indicate when the belt was starting to wear but, whether it would be practical in terms of its size, reliability, cost and extra weight is very much open to debate.

Here we have seen four different solutions to the same problem of avoiding failures and hence the need for corrective maintenance. One of the tasks

of the maintainability engineer is to determine which, if any of these, or other similar approaches is appropriate taking into consideration the costs and practicalities in each circumstance.

There is clearly a need to strike a balance. Preventative maintenance may cause components to be replaced unnecessarily (or at least prematurely). Allowing a system to run until it fails may maximise the times between maintenance but failures can be expensive to rectify both because of the extent of the damage caused and because of the loss of availability of the system whilst it is being maintained. Prognostics can help but these too have their own problems of reliability and the need for maintenance as well as possibly adding to the weight, complexity and cost of the system.

30.4 Maintenance cost factors

For many systems/products, maintenance costs constitute a major segment of the total life-cycle cost. Further, experience has indicated that maintenance costs are significantly affected by design decisions made throughout the early stages of system development. Maintainability is directly concerned with the characteristics of system design that will ultimately result in the accomplishment of maintenance at minimum cost. Thus, one way of measuring maintenance cost is cost per maintenance task, which is the sum of all costs related to elements of logistics support which are required to perform the considered maintenance task.

In addition to the above factors, the frequency with which each maintenance action must be performed is a major factor in both corrective and preventive maintenance. Obviously this is greatly influenced by the reliability of the components but it can also be related to the type and frequency of the maintenance performed. If a component is repaired then it is likely that the time to failure for that component will be less than if it had been replaced by a new one. We will return to the question of repair effectiveness in Chapter 6.

Personnel and human factor considerations are also of prime importance. These considerations include the experience of the technician, training, skill level and number of technicians.

Support considerations cover the logistics system and maintenance organisation required to support the system. They include the availability of spare parts, technical data (manuals), test equipment and required special and general tools.

If a maintenance task requires highly skilled personnel, a clean environment equipped with expensive, special tools then it is unlikely, that it will prove economical to perform this task at first line or, possibly, even at second line.

However, if the maintainability engineer had designed the system so that this task could be done by personnel with lower skill levels using standard tooling then it might have allowed the task to be done in the field with a possible reduction in the turnaround (or out-of-service) time. If, the task is only likely to be done once in the system's life during a major overhaul when it would be at a central maintenance unit or returned to the manufacturer then such considerations may be less relevant. For example, there is little to be gained by making it easy to replace a broken cam belt by the side of the road. The damage done to the engine, as a result of a failed cam belt, will mean that the engine will either have to be replaced or overhauled/reconditioned before it is likely to function again.

31.   MAINTAINABILITY DEMONSTRATION

The objective of the maintainability demonstration is to show that the various maintenance tasks can be accomplished in the times allotted to them. Generally, the most important issue is whether the system can be recovered by sub-system (or line replaceable unit – LRU) exchange within the specified times. It is a common requirement that each LRU can be removed and replaced without interfering with any other LRU. Some of the early jet fighters were virtually built around the engine so that, in order to replace the engine, it was not so much a question of removing the engine from the aircraft as removing he aircraft from the engine.

A recent innovation on commercial aircraft is to use autonomics, which signal ahead to the destination any detected faults in the mission critical components (i.e. those not on the minimum equipment list). This allows the mechanics to prepare to replace these items as soon as the aircraft has reached the gate. If such replacements can be performed within the 50 min, or so, turnaround time then it will not be necessary to find a replacement aircraft or delay the departure. Anyone who has seen the film *Battle of Britain* or *Reach for the Sky* will recognise the importance of turning fighter aircraft around in minimum time when the airfield may be under attack from enemy bombers and fighters. An aircraft not in the air is bit like a duck out of water, it is particularly vulnerable and do very little to defend itself.

The demonstration is also expected to generate results that can contribute to the whole development process, identifying any remaining deficiencies such as the design of the system and the test equipment, compilation of maintenance manuals, etc. Any maintainability demonstration would involve the following steps:

1. Identify the operation and environmental condition in which the system is likely to be used.
2. Simulate the system failures and perform corrective maintenance action.   One should also record the maintenance man-hours required to complete the repair task successfully.

Further, it is an important to take care of the following issues during the demonstration:

1. The test must be on a sample of fixed final build standard.
2. The test conditions must be representative, the equipment/tools, maintenance manuals, lighting and similar factors must be carefully considered.
3. A mix of repairers representative in skills, training, and experience of those who would do the actual repair in service must conduct the repair.

Once we have the recorded repair time data from the above procedure, then it is easy to verify whether the maintainability target has been achieved using the following procedure.

Let t_1, t_2, ..., t_n denote the observed repair times to complete the repair tasks for a sample of n units. For n > 30, the (1 - α) 100 percent confidence limit is given by:

$$MTTR + z_\alpha \frac{s}{\sqrt{n}} \tag{5.8}$$

Where z_α is the *z* value (standard normal statistic) that locates an area of α to its right and can be found from the normal table. For example, for a 95% confidence limit, the z_α is given by 1.645. MTTR and 's' are given by:

$$MTTR = \frac{1}{n} \sum_{i=1}^{n} t_i \text{ , and } s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (t_i - MTTR)^2$$

If the target maintainability is $MTTR^*$, then to demonstrate that the system has achieved this, we have to show that:

$$MTTR^* \leq MTTR + z_\alpha \frac{s}{\sqrt{n}} \tag{5.9}$$

Whenever the number of repair time data is less than 30, we use t-distribution; in that case, the condition for acceptance is given by:

$$MTTR^* \leq MTTR + t_{\alpha,n-1} \frac{s}{\sqrt{n}} \qquad (5.10)$$

The value of $t_{\alpha,n-1}$ can be obtained from the t-distribution table shown given in appendix.

Example 5.2

A maintainability demonstration test is carried out on 20 parts and the accomplished repair times are shown in Table 5.2. If the target MTTR is 20 hours, check whether the system has achieved the target maintainability using 95% confidence level.

Table 5.2. Recorded repair times form a sample of 20 parts in hours

8	6	12	20	24
12	9	17	4	40
32	26	30	19	10
10	14	32	26	18

SOLUTION:

Since the observed number of data, n is less than 30, we use t-statistic. The MTTR and standard deviation, s, are given by:

$$MTTR = \frac{1}{20} \sum_{i=1}^{20} t_i = 18.45 \text{ hours}, \quad s = \sqrt{\frac{1}{19} \sum_{i=1}^{n} (t_i - MTTR)^2} = 10.06 \text{ hours}$$

From the t-distribution table (see appendix) we get, $t_{\alpha,n-1}$ = 1.729 (α = 0.05, n-1 = 19).

95% upper limit for MTTR is given by:

$$MTTR + t_{\alpha} \frac{s}{\sqrt{n}} = 18.45 + 1.729 \times \frac{10.06}{4.472} = 22.33$$

Which is greater than 20 hours, which is the target MTTR. Thus the achieved MTTR is significantly greater than the required MTTR and is therefore not acceptable.

32.　　MAINTENANCE

According to BS 4778, maintenance can be defined as:

The combination of all technical and administrative actions, including supervision actions, intended to retain an item in, or restore it to, a state in which it can perform a required function.

In other words, all actions, which keep the system running and ensure that it is maintained to an acceptable standard in which it is able to operate at the required levels efficiently and effectively.　The objectives of maintenance are to:

1. Reduce the consequences of failure.
2. Extend the life of the system, by keeping the system in a proper condition for a longer time. In other words, to increase the "up" time of the system.
3. Ensure that the system is fit and safe to use.
4. Ensure that the condition of the system meets all authorised requirements.
5. Maintain the value of the system.
6. Maintain reliability and achieve a high level of safety.
7. Maintain the system's availability and therefore minimise production and quality losses.
8. Reduce overall maintenance costs and therefore minimise the life cycle cost.

The purpose of maintenance is to keep systems in a state of functioning in accordance with their design and to restore them to a similar state as and when required.

33.　　MAINTENANCE CONCEPT

The maintenance concept begins with a series of statements defining the input criteria to which the system should be designed.　These statements relate to the maintenance tasks that should be performed at each level of

maintenance (organisational, intermediate and depot), the test equipment and tools that should be used in maintaining the system, the skill levels of the maintenance personnel that perform the identified tasks, maintenance time constraints, and anticipated maintenance environmental requirements (Knezevic, 1997). A preliminary maintenance concept is developed during the conceptual design stage, is continually updated, and is a prerequisite to system design and development. Maintenance concept at the design phase tends to ensure that all functions of design and support are integrated with each other. The maintenance concept evolved from the definition of system operational requirements delineates [Blanchard et. al., 1995]

- The anticipated level of maintenance
- Overall repair policies
- Elements of maintenance resources
- The organisational responsibilities for maintenance

The maintenance concept serves the following purposes:

1. It provides the basis for the establishment of maintainability and supportability requirements in the system design.
2. It provides the basis for the establishment of requirements for total support which include maintenance tasks, task frequencies and time, personnel quantities and skill levels, spare parts, facilities, and other resources.
3. It provides a basis for detailing the maintenance plan and impacts upon the elements of logistic support.

34.   LEVELS OF MAINTENANCE

Complex systems can be considered as made up of several levels of indenture. A combat aircraft that may be considered as the Level 0 (LoI-0), may be thought of as consisting five subsystems: airframe, armament, avionics, propulsion and general. The propulsion system then becomes a LoI-1 item that may consist of the engines, the auxiliary power unit (APU) and various accessories including control units and pumps, each of which may be considered as LoI-2 items. An engine is typically an assembly of a number of modules or LoI-3 items which, in turn, may be made up of sub-assemblies and parts, LoI-4 and 5 respectively.

At the same time, the military typically divides its maintenance and support infrastructure into 3, 4 or 5 echelons, lines or [maintenance] levels. "First

Line", or "O-Level" is from where the systems are operated. "Second Line" or "I-Level" is typically the main operational bases from which the squadrons are deployed. These are usually supported by a depot or maintenance unit at "Third Line" or "D-Level". The contractor, supplier or original equipment manufacturer (OEM) often provides a shadow facility at "Fourth Line" effectively duplicating the Third Line facility's capabilities.

Maintenance levels are concerned with grouping the tasks for each location where maintenance activities are performed. The criteria in which the maintenance tasks selected at each level are; task complexity, personnel skill-level requirements, special maintenance equipment and resources and economic measures. Within the scope of the identified level of maintenance, the manufacturer and the user should define a basic repair policy that may vary from repair/replace a part (LoI-5, say) to replace the entire system. The hierarchies of achieving maintenance tasks are divided into three or four levels.

34.1   User level (organisational)

This type of maintenance level is related to all maintenance tasks which are performed on the system whilst it is on deployment or at its operating site. This would include replenishment tasks, e.g. re-fuelling, re-arming, maintaining oil levels, simple condition and performance monitoring activities, external adjustments and replacement of line replaceable units (LRU). Some minor repairs and routine servicing may also come under this category.

34.2   Intermediate level

Intermediate maintenance level is related to all maintenance tasks, which are performed at workshops (mobile, semi-mobile and/or fixed) where the systems would normally be based. Common maintenance tasks accomplished at this level are detailed condition and performance monitoring activities, repair and replacement of major items in a system, major overhaul, system modification, etc. Performing maintenance tasks at this level require higher personnel skills than those at organisational level and additional maintenance resources. Traditionally, a removed LRU would be recovered, generally by module (or shop-replaceable unit – SRU) exchange, at this level.

34.3   **Depot level**

Depot maintenance level is related to all maintenance tasks, which are accomplished beyond the capabilities of intermediate level at remote sites. In the UK system, "Third Line" refers specifically to an operator-owned facility whereas in US nomenclature "D-Level" also includes manufacturer/contractor facilities.   Maintenance tasks at this level are carried out by highly skilled specialists at a specialised repair facility or the equipment producer's facility.   Maintenance tasks at depot level include complete overhauling and rebuilding of the system, highly complex maintenance actions, etc. They would also include tasks which may only be performed rarely, particularly if they require expensive equipment or are likely to take a long time.

34.4   **Hole-in-the-Wall**

With the move to ever greater efficiency and/or minimal costs, the perceived need to reduce manning levels and the desire of OEMs to increase their revenue by entering the "after-market", the "hole-in-the-wall" concept is gaining in popularity.   This is where the only intrusive maintenance task the operator performs is to remove the LRU (at first line). This is then passed through this mythical hole in the wall to the OEM or maintenance contractor in exchange for a replacement (serviceable) LRU. The contractor then takes the LRU away to a convenient location where it is recovered.   Such contracts are often funded by fleet hour arrangements such as "power-by-the-hour", see chapter 12.

The advantage to the operators is that they can get on with what they are in business for; putting "bums on seats" or "bombs on target".   It is also argued, perhaps more strongly by the OEM than the operator, that having designed and built the LRU, they (the OEM) are the best people to take it apart and repair it.   A secondary advantage to the OEM, and again, hopefully to the operator, is that because all of the maintenance is done in one place, the people doing it should become more efficient (as they see the same job more often) and the in-service data (time to failure, cause of failure, items repaired or replaced, etc.) should be consistent and more accurate.

Better data should lead to improved forecasting, reduced logistic delays, more appropriate maintenance policies and, ultimately, to improved designs.

35.   MAINTENANCE TASK CLASSIFICATION

All users would like their systems to stay in a state of functioning as long as possible or, at least, as long as they are needed. In order to achieve this, it is necessary to maintain the system's functionality during operation, by performing appropriate maintenance tasks. Thus, maintenance task can be defined as a set of activities that need to be performed, in a specified manner, in order to maintain the functionality of the item/system.

Figure 5.2 shows the process of maintenance task, which is initiated by the need for maintenance due to a reduction, or termination of the item/system functionality. The execution of a maintenance task requires resources such as the right number and skills of personnel, material, equipment, etc. It also requires an appropriate environment in which the maintenance activities can be carried out.
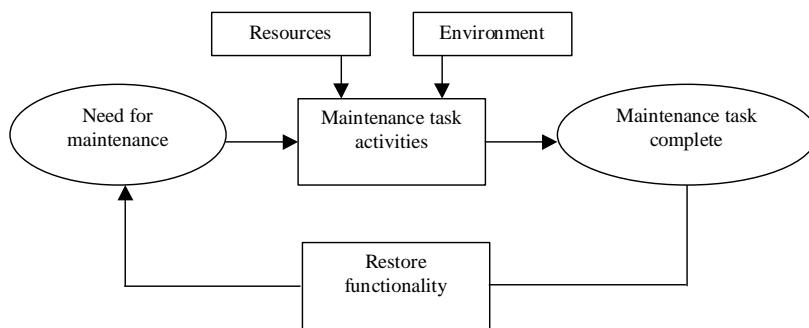
Figure 5.2 Process of maintenance task

Maintenance tasks can be classified into the following three categories:

1. corrective maintenance task
2. preventive (predictive) maintenance task
3. conditional maintenance task

Each maintenance task is briefly discussed in the following sections.

35.1   Corrective Maintenance Task

Corrective maintenance task, CRT, is a set of activities, which is performed with the intention of restoring the functionality of the item or system, after

the loss of the functionality or performance (i.e. after failure). Figure 5.3 illustrates typical corrective maintenance task activities. The duration of corrective maintenance task, DMT^c, represents the elapsed time needed for the successful completion of the task. Corrective maintenance task is also referred to as an unscheduled or unplanned maintenance task.
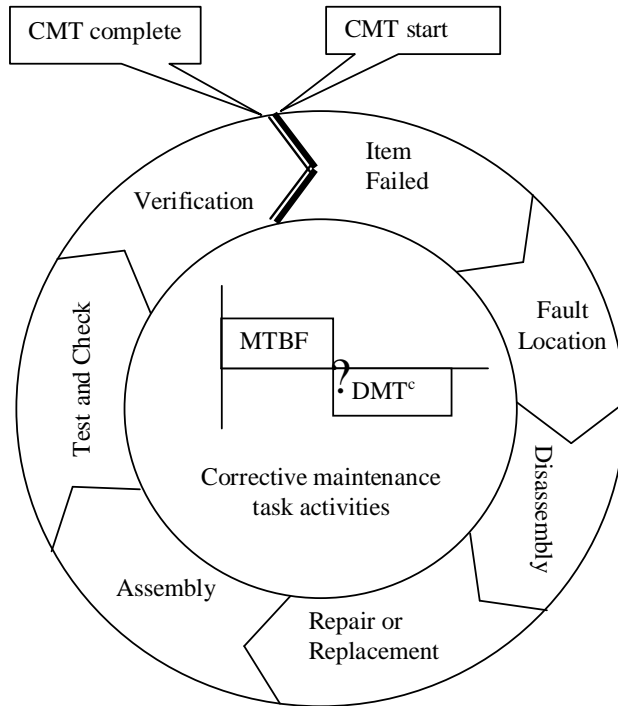
Figure 5.3 Activities of typical corrective maintenance task

35.2   **Preventive Maintenance Task**

Preventive maintenance task, *PMT*, is a maintenance activity that is performed in order to reduce the probability of failure of an item/system or to maximise the operational benefit. Figure 5.4 illustrates the activities of a typical preventive maintenance task. The duration of the preventive maintenance task, DMT^p, represents the elapsed time needed for the successful completion of the task.
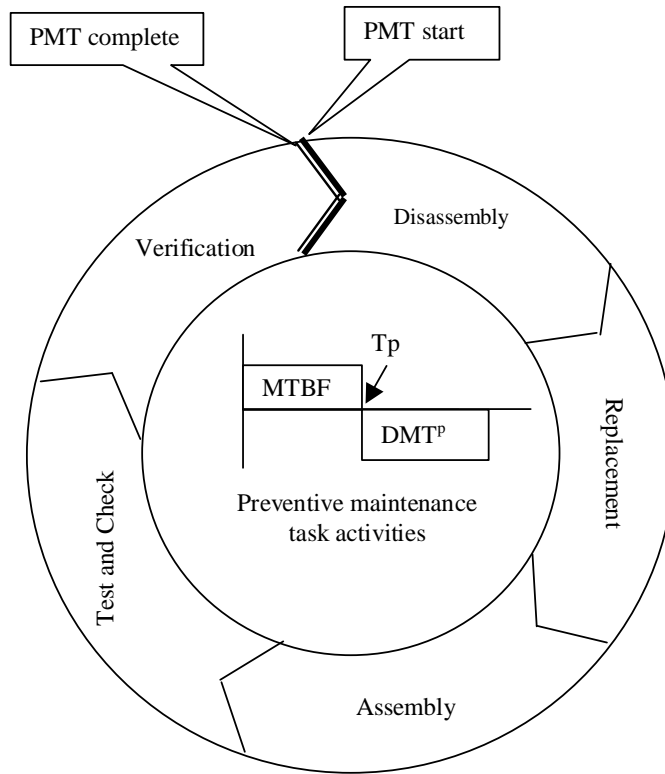
Figure 5.4 Activities of a typical preventive maintenance task

Preventive maintenance task is performed before the transition to the state of failure occurs with the main objective of reducing:

- The probability of the occurrence of a failure
- The consequences of failure

Common preventive maintenance tasks are replacements, renewal and overhaul. These tasks are performed, at fixed intervals based on operating time (e.g. hours), distance (e.g. miles) or number of actions (e.g. landings), regardless of the actual condition of the items/systems.

35.3 Conditional (Predictive) Maintenance Task

Conditional maintenance task, *COT*, recognises that a change in condition and/or performance is likely to precede a failure so the maintenance task should be based on the actual condition of the item/system. *COT* does not normally involve an intrusion into the system and actual preventive action is taken only when it is believed that an incipient failure has been detected. Thus, through monitoring of some condition parameter(s) it would be possible to identify the most suitable instant of time at which preventive maintenance tasks should take place.
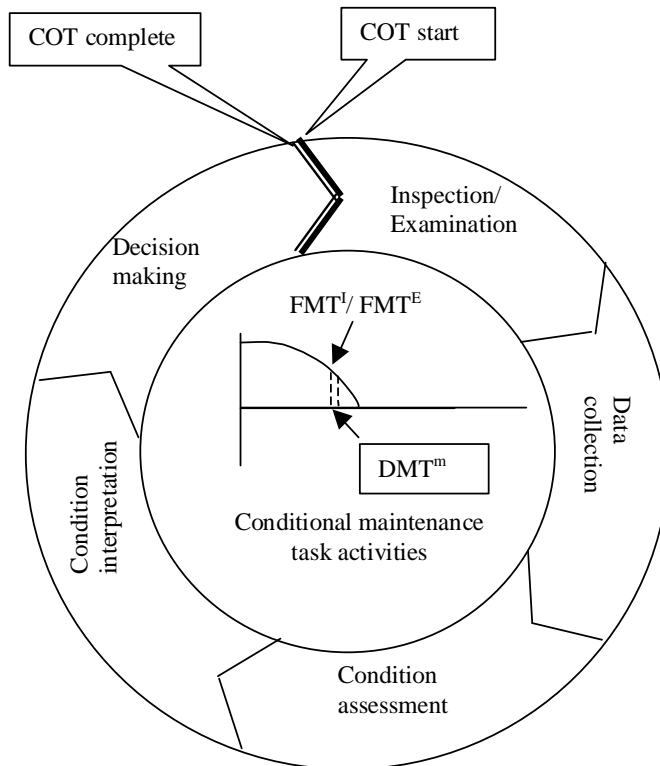
Figure 5.5 Activities of a typical conditional maintenance task.

Figure 5.5 illustrates the activities of a typical conditional maintenance. The duration of conditional maintenance task, DMT^m, represents the elapsed time needed for the successful completion of the task.

In the past, corrective maintenance and preventive maintenance tasks have been popular among maintenance managers. However, in recent years, the

disadvantages of these tasks have been recognised by many maintenance management organisations. The need for the provision of safety, and reduction of the maintenance cost have led to an increasing interest in using conditional maintenance task. Waiting until a component fails may maximise the life obtained from that component but, its failure may cause significant damage to other parts of the system and will often occur at inopportune times causing a disruption to the operation and inconvenience to the users. Routine or scheduled preventive maintenance, on the other hand, may be very convenient but is likely to result in an increase in the amount of maintenance needed because parts will be replaced when they have achieved a fraction of their expected life.

36.   MAINTENANCE POLICIES

The maintenance policy defines which type of maintenance will (normally) be performed on the various components of the system. It is determined by maintenance engineers, system producers and /or users to achieve high safety, reliability and availability at minimum cost. With respect to the relation of the instant of occurrence of failure and the instant of performing the maintenance task the following maintenance policies exist:

1) *Failure-Based maintenance policy, FBM*, where corrective maintenance tasks are initiated by the occurrence of failure, i.e., loss of function or performance,
2) *Time-Based maintenance policy, LBM*, where preventive maintenance tasks are performed at predetermined times during operation, at fixed length of operational life,
3) *Inspection-Based maintenance policy, IBM*, where conditional maintenance tasks in the form of inspections are performed at fixed intervals of operation, until the performance of a preventive maintenance task is required or until a failure occurs requiring corrective maintenance. Note that the failure could be due to a component of the system that was not being subjected to IBM or it could have happened as a result of some unpredictable external event such as foreign object damage or because the inspection interval was too long or the inspection was ineffective.
4) *Examination-Based maintenance policy, EBM*, where conditional maintenance tasks in the form of examinations are performed in accordance with the monitored condition of the item/system, until the execution of a preventive maintenance task is needed or a failure occurs.

The principal difference between the above maintenance policies occurs at the time when the maintenance task is performed. The advantages and disadvantages of each maintenance policy are briefly described below.

36.1 Failure-Based Maintenance Policy

Failure-Based maintenance policy, *FBM*, represents an approach where corrective maintenance tasks are carried out after a failure has occurred, in order to restore the functionality of the item/system considered. Consequently, this approach to maintenance is known as breakdown, post-failure, fire fighting, reactive, or unscheduled maintenance. According to this policy, maintenance tasks often take place in ad hoc manner in response to breakdown of an item following a report from the system user.
A schematic presentation of the maintenance procedure for the failure-based maintenance policy is presented in Figure 5.6. Corrective maintenance task priorities can range from "normal", "urgent" to "emergency". These categories reflect the nature of the response rather than the actual actions done. Failure based maintenance could be the most applicable and effective maintenance policy in situations where:

Items for which the loss of functionality does not compromise the safety of the user and/or the environment or the failure has little or no economic consequences (i.e. categories "major" and "minor" see "FMECA" in Chapter 11)

systems have built-in redundancy or have been designed to be fault-tolerant
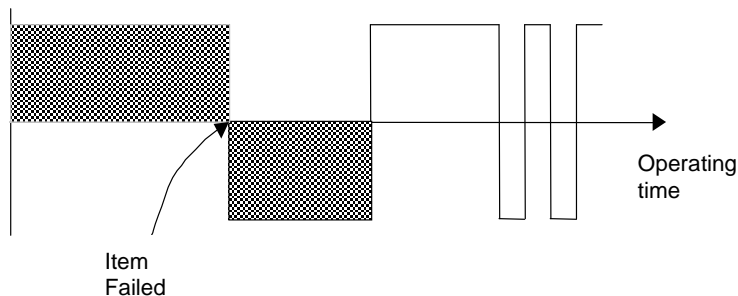
Figure 5.6 Failure-Based Maintenance Policy

Advantages of failure based maintenance

Implementation of FBM to the above situations could lead to full utilisation of the operating life of the item. This means that the non-critical items will have the ability to perform their function(s) for the stated period of time when they operate under stated conditions. This means that coefficient of utilisation, CU, which is the ratio of the Mean Duration of Utilised Life of the item *(MDUL^F)* to the expected operating life (MTTF), of items considered will have value of 1. The user will get maximum value out the component when the FB maintenance policy is applied.

Disadvantages of failure based maintenance

Despite the advantages of implementing this policy, it has some disadvantages when it is not correctly selected.

- The failure of an item will generally occur at an inconvenient time.

- Maintenance activities cannot be planned.

- It demands a lot of maintenance resources.

- The failure of an item can cause a large amount of consequential damage to other items in the system.

Analysis of maintenance costs have shown that a repair made after failure will normally be three to four times more expensive than the same maintenance activity when it is well planned [Mobley (1990)].

36.2   Time-Based Maintenance Policy

Some failures can lead to economical consequences such as loss of production and therefore a reduction in profit. Some failures may have an impact on the safety of the user, passengers, third parties and environment. Therefore, it is desirable to prevent these failures, if possible, by carrying out maintenance actions before failure occurs.
As the main aim is to reduce the probability of occurrence of failure and avoid the system breakdown, a time-based maintenance policy is performed at fix intervals, which is a function of the time-to-failure distribution of the item considered and in some cases it may be adjusted by the system's user. This policy is very often called age-based, life-based, planned or scheduled maintenance. The reason for that is the fact that the

maintenance task is performed at a predetermined frequency, which may be based on, for example, operating times such as, hours, years, miles, number of actions or any other units of use, that make it is possible to plan all tasks and fully support them in advance. A schematic presentation of time-based maintenance procedure is presented in Figure 5.7. The frequency of maintenance task, FMT^L, is determined even before the item has started functioning. Thus, at the predetermined length of operational life specified, preventive maintenance tasks take place. The time-based maintenance policy could be effectively applied to items/systems that meet some of the following requirements:

1. the probability of occurrence of failure is reduced
2. the likely consequences of failure is "catastrophic" (e.g. loss of life or serious injury)
3. the total costs of applying this policy are substantially lower than the alternatives
4. the condition of the system, or its consisting items, cannot be monitored or is impractical or uneconomical.

Advantages of time-based maintenance policy

One of the main advantages of this maintenance policy is the fact that preventive maintenance tasks are performed at a predetermined instant of time when all maintenance support resources could be planned and provided in advance, and potential costly outages avoided. For failures, which could have catastrophic consequences to the user/operator and environment (Chernobyl, Bhopal, Piper Alpha and similar) it may be the only feasible option. Time-based maintenance has many advantages over failure-based maintenance, which are summarised in the following list:

1. Maintenance can be planned ahead and performed when it is convenient from the operational and logistics point of view.
2. The cost of lost production and of consequential damage can be reduced.
3. Downtime, the time that the system is out of service, can be minimised.
4. Safety can be improved.

FMT^L

System in use

Preventive Task

Operating time

Predetermined Time Tp
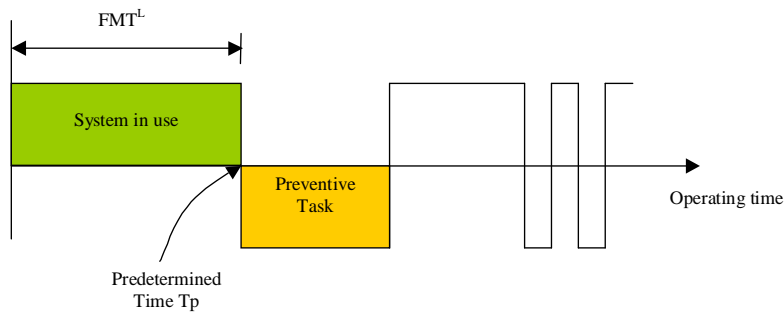
Figure 5.7 Time Based Maintenance Policy

Disadvantages of time -based maintenance policy

In spite of the advantages given above, the time-based maintenance policy has several disadvantages that must be minimised. This policy could be uneconomical because the majority of items are prematurely replaced, irrespective of their condition. In many industries this policy is now only used under special conditions because it is very costly, and also because its efficiency in reducing failures is not always supported by experience. A summary of the disadvantages of time-based maintenance policy is listed below.

1. Time-based maintenance is performed irrespective of the condition of the system. Consequently, a large number of unnecessary tasks will be carried out on a system that could have been operated safely for a much longer time.
2. The tasks may require higher numbers of skilled mechanics.
3. If the time to perform the maintenance is greater than the time the system would normally be idle (eg overnight) then because of the frequency, it could cause higher levels of unavilability.
4. It cannot guarantee the elimination of all failures and will do nothing to reduce non-age-related failures.
5. Increasing the frequency of maintenance tasks may lead to an increase in the probability of human errors in the form of maintenance-induced failures.
6. Reducing the probability of failure by prematurely replacing components means that the coefficient of utilisation of the item/system, CU^L, will have a value much less than one.

36.3   Condition Based Maintenance (Predictive Maintenance)

The need for the provision of safety, increased system availability, and reduced maintenance costs have led to an increasing interest in development of alternative maintenance policies.   A policy which overcomes many of the disadvantages of the previous maintenance policies (failure-based and time-based), and has proved its ability to extend the operating life of a system without increasing the risk of failure is condition-based maintenance, CBM. CBM is also known as predictive maintenance. Condition-based maintenance can be defined as: "*Maintenance carried out in response to a significant deterioration in a unit as indicated by a change in the monitored parameters of the unit's condition or performance*" [Kelly & Harris (1978)].   This means that the principle reason for carrying out maintenance activities is the change or deterioration in condition and/or performance, and the time to perform maintenance actions is determined by monitoring the actual state of the system, its performance and/or other condition parameters. This should mean the system is operated in its most efficient state and that maintenance is only performed when it is cost-effective. A schematic presentation of condition-based maintenance procedure is presented in Figure 5.8.   This policy is worth applying in situations where:

1- The state of the system is described by one or more condition parameters.
2- The cost of the condition monitoring technique is lower than the expected reduction in overall maintenance costs.
3- There is a high probability of detecting potentially catastrophic failures (before they happen).
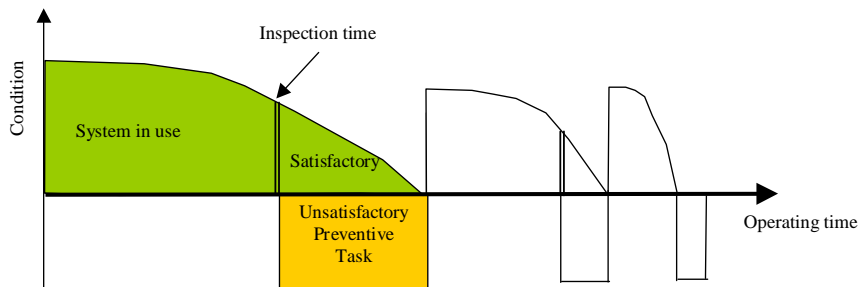
Figure 5.8 Condition based maintenance policy

The condition-based maintenance is a condition or performance-driven preventive maintenance. This means that the timing of the maintenance task is not simply a function of the mean-time-to-failure. The principle of condition-based maintenance therefore is based on the way of monitoring the condition parameters of a system giving three different types of condition monitoring:

    1- Inspection

Inspection is generally performed at regular intervals using any of a number of non-destructive test (NDT) procedures which are designed to determine whether the condition of the (inspected) item is satisfactory or unsatisfactory and hence whether further action is required.

    2- Examination

This is a condition-monitoring task, which presents a numerical description of the condition of the item at that moment through relevant condition predictors. The results directly affect the scheduling of the next examination. This is possible because of the unique properties and characteristics of the relevant condition predictor.

3 - Performance Trend Monitoring

For propulsion or energy producing systems, in particular, the "performance" may be expressed as a ratio of the output to input, e.g. miles per gallon, kilometres per litre, thrust per kilogram or (mega)watts per tonne. As the system deteriorates, usually through wear but also through damage, these ratios may show signs of decreasing. For systems operating in relatively constant conditions (e.g. constant ambient temperature, pressure and output), consistent changes in the specific fuel consumption (SFC) will almost certainly be indicative of a deterioration in the system which will need some form of maintenance to restore it to an acceptable level. For systems that are operated in an inconsistent manor for which the environmental conditions may be in a constant state of change, the SFC may be subject to considerable noise and hence any deterioration will only be apparent by using sophisticated trending algorithms, such as Kalman Filtering.

36.3.1 Setting up condition-based maintenance policy

In order to implement CBM policy, it is necessary to use the following management steps that are shown in Figure 5.9

Identification and selection of maintenance significant items

The first requirement of implementing CBM is to decide which items of the system should be monitored, since it is likely to be both uneconomical and impractical to monitor them all. Therefore, the first step of the condition-based maintenance decision process is a comprehensive review of all items in a system, in order to identify the maintenance significant items, MSIs. These are items whose failures could be safety-critical, environmentally damaging or revenue sensitive. Thus, each item within the system should be analysed from the point of view of failure, especially the consequences of failure. The most frequently used engineering tools for performing this task is a Failure Mode Effect and Criticality Analysis, FMECA and Reliability Centred Maintenance, RCM (see also Chapters 6 and 11). Care should be taken to ensure that all of the maintenance significant items are identified and listed.

Identification and selection of condition parameters

Once the maintenance significant items are identified it is necessary to determine all monitorable parameters which describe their condition or performance. The condition parameter can be defined as a measurable variable able to display directly or reflect indirectly information about the condition of an item at any instance of operating time. Ideally, maintenance engineers would like to find many condition/parameters which can be monitored and which accurately reflects the condition /performance of the system. In practice there are two distinguishable types of condition parameters which are able to achieve this (Knezevic *et al*, 1995):

A. Relevant Condition Indicator, RCI

The Relevant Condition Indictor, RCI, is a parameter that describes the condition of an item during its operating time and it indicates the condition

of the item at the instant of inspection. The numerical value of RCI represents the local value of the condition of an item/system at the time of inspection. This type of condition parameter is usually related to the performance. However, RCI is not able to predict the future development of the condition of the considered item/system. Typical examples of the RCI are performance, the level of vibration, level of oil, pressure, temperature, etc. It is necessary to stress that the RCI could have an identical value at different instances of operating time.
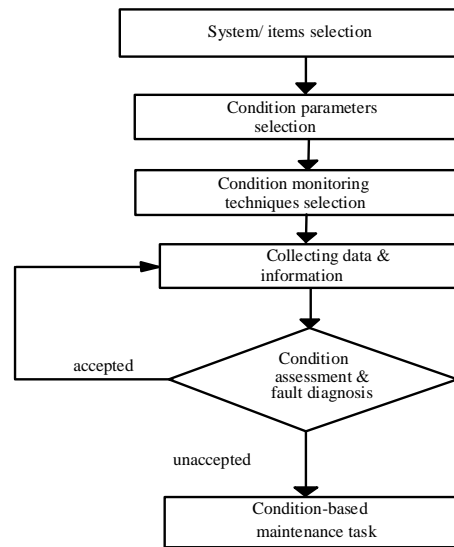
Figure 5.9 Flow of condition based maintenance

Relevant Condition Predictor (RCP)

The Relevant Condition Predictor, RCP, is a parameter, which describes the condition of an item at every instant of operating time. Usually this parameter is directly related to the shape, geometry, weight, and other characteristics, which describe the condition of the item under consideration. The RCP represents the condition of the item/system which is most likely to be affected by a gradual deterioration failure such as wear, corrosion fatigue crack growth. The general principles of the RCP are discussed by Knezevic (1987). Typical examples of RCP are: thickness of an

item, crack length, depth of tyre treads, etc. The RCP cannot have identical values at two or more instance of time. The numerical value of the relevant condition predictor at any instant of operating time quantifies the cumulative value of the condition of an item/system at the time of examination.

Selection of condition monitoring technique

Having identified the maintenance significant item and the associated condition parameter(s), the next step is to select the suitable monitoring technique, which will be used to inspect and examine each condition parameter.

The condition monitoring technique is a device used to inspect or examine an item in order to provide data and information about its condition at any instance of operating time. Numerous condition monitoring techniques, for instance, NDT techniques, performance, vibration, etc are available for use by maintenance engineers in order to determine measurable value of condition parameter. It is important to understand the behaviour of the failure that the item exhibits so that the most effective monitoring techniques can be chosen.

The decision as to which condition-monitoring techniques are selected depends greatly on the type of system, the type of condition parameter and, in the end, on cost and safety. Once the decision is made as to which techniques are to be used, it is possible to define the equipment or instrument that will be needed to carry out condition monitoring.

Collecting data and information

The philosophy of condition monitoring is to assess the condition of an item/system by the use of techniques which can range from human sensing to sophisticated instrumentation, in order to determine the need for performing preventive maintenance tasks. With the increased interest in condition monitoring in recent years there have been a number of developments in the techniques that are used to collect data and provide information, which helps maintenance engineers assessing the condition of an item or a system. These developments have made it possible to obtain more reliable information on the condition of the system. In many instances such information is used to insure that the status of the system will continue to be in a functioning state without significant risk of

breakdown, and in some instances to make a decision on the timing of when maintenance tasks should be performed. The method of data collection can be classified into the following categories:

On-line data collection and monitoring

On-line data collection and monitoring uses instrumentation fitted to the system which takes continuous measurements of the condition parameters. These may then be analysed by an on-board computer to determine whether there has been a change in the condition of the item/system and whether that change requires any action. The benefit of using on-line monitoring is to reduce the need for human intervention and minimise the probability of a failure occurring between inspections.

Off-line collection and monitoring

Off-line collection and monitoring is periodic measurement of a condition of an item/system or continuous data collection which is analysed remotely. This type of method involves either the collection of data using a portable data collector, or taking a physical sample, for example, lubrication oil samples for analysis of contamination and debris content. Periodic monitoring therefore provides a way of detecting progressive faults in a way that may be cheaper than the on-line system.
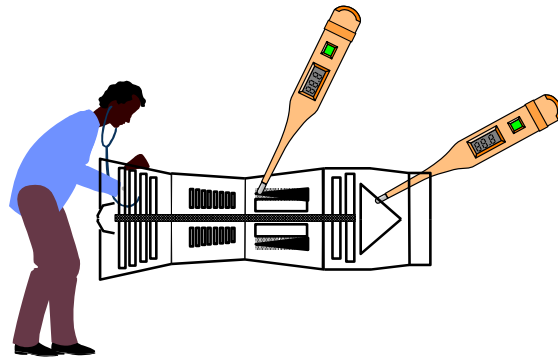
Figure 5.10. Condition monitoring and condition assessment

Condition assessment

The assessment of the condition of an item/system (Figure 5.10) can range from human experience to sophisticated instrumentation. The last few decades have seen a number of developments in the methods which are used to help the maintenance engineers assess and diagnose the condition of an item/system and provide them with information on which to base their decision. Once condition monitoring sensors have been installed and data are being collected, it is necessary to have reliable methods of interpreting the data to identify whether the considered item is undergoing a transition from the normal to abnormal condition and in many cases to identify the causes of the changes.

Effective condition-based maintenance requires a large number of measurements taken continuously or at intervals that assure recognition of change in the condition of the item/system in sufficient time to avoid the need for any corrective action. The volume of data necessary to accurately determine the condition of the item/system can require an excessive amount of time to process and analyse. Consequently, the demand to manipulate and process large amounts of data very quickly has lead to the development of tools such as Artificial Intelligence, AI, to assist engineers to gain maximum value from the data.

In recent years, Artificial Intelligence techniques such as Expert System, Neural Networks and Fuzzy Logic have been applied to the discipline of monitoring and diagnostic systems [Mann et al (1995)]. These techniques extend the power of the computer beyond the usual mathematical and statistical functions by using dialogue and logic to determine various possible courses of action or outcome. By processing information much faster (than humans) the time to assess the condition and diagnose the causes of failures can be reduced. It can analyse situations objectively and will not forget any relevant facts (given that it has been supplied them), therefore the probability of making a wrong assessment or diagnosis may be reduced. Furthermore, it can detect incipient failures through its on-line monitoring of the condition parameters of the system [Lavalle et al (1993)].

Implementation of condition based maintenance

Having identified and listed all the condition parameters of the maintenance significant items, the aim of this step is to implement condition based maintenance. According to the classifications of condition parameter, condition based maintenance could be divided in two policies:

Inspection Based Maintenance Policy

The suitable maintenance policy for items for which their conditions are described by the relevant condition indicator, *RCI* is inspection-based maintenance. The algorithm, which presents the maintenance procedure in this case, is shown in Figure 5.11

Inspection is carried at fixed intervals to determine whether the condition of the item, is satisfactory or unsatisfactory according to the *RCI*. Before the item/system is introduced into service the most suitable frequency of the inspection, FMT^1, and critical value of relevant condition indicator RCI_{cr} has to be determined. Once the critical level is reached, $RCI(FMT^1) > RCI_{cr}$, the prescribed preventive maintenance tasks take place. If the item fails between inspections, corrective maintenance takes place.

Advantages of inspection based maintenance

CBM has the potential to produce large savings simply by allowing items in the system to be run to the end of their useful life. This reduces the equipment down time and minimises both scheduled and unscheduled breakdown situations. By eliminating all unscheduled interruptions to operation and production and only carrying out required maintenance in a carefully controlled manner, it is possible to reduce the maintenance cost, to improve safety, improve the efficiency of the operation and increase the system's availability.

```
┌─────────────────────────────────────┐
│         MAINTENANCE PROCEDURE        │
├─────────────────────────────────────┤
│   Condition-based Maintenance policy │
├───────────┬─────────────────────────┤
│   Type    │       Inspection        │
└───────────┴─────────────────────────┘
                     │
                     ▼
         ┌─────────────────────┐
         │   Determination of   │
         │  FMT and RCI         │
         └─────────────────────┘
                     │
                     ▼
         ┌─────────────────────┐
    ┌───▶│     System in use    │◀───┐
    │    └─────────────────────┘    │
    │              │                │
    │              ▼                │
    │    ┌─────────────────────┐    │
    │    │   Inspection of RCI  │    │
    │    └─────────────────────┘    │
    │              │                │
    │              ▼                │
    │          ╱────────╲     No    │
    │         ╱ RCI > RCI ╲─────────┘
    │         ╲    cr     ╱
    │          ╲────────╱
    │           Yes │
    │               ▼
    │    ┌─────────────────────┐
    └────│   Maintenance task   │
         └─────────────────────┘
```
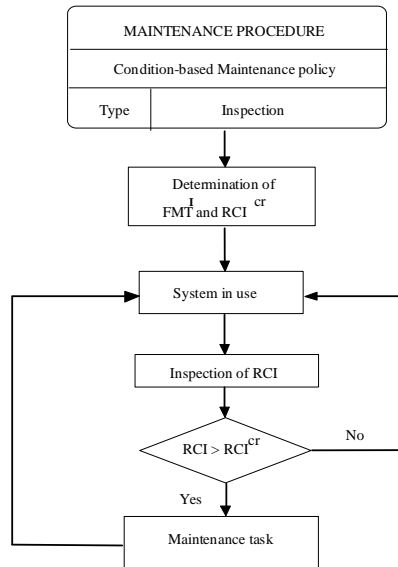
Figure 5.11 Algorithm for inspection based maintenance task

The benefits of inspection based maintenance policy can be summarised as follows:

1. Reduce unplanned downtime, since maintenance engineers can determine optimal maintenance intervals through the condition of constituent items in the system. This allows for better maintenance planning and more efficient use of resources.
2. Improve safety, since monitoring and detection of the deterioration in condition and/or performance of an item/system will enable the user to stop the system (just) before a failure occurs.
3. Extending the operating life of each individual items and therefore the coefficient of life utilisation will be increased compared to time based maintenance
4. Improve availability by being able to keep the system running longer and reducing the repair time.
5. Reduce maintenance resources due to reduction in unnecessary maintenance activities
6. The above benefits will lead to a reduction in maintenance costs

Examination Based Maintenance Policy

The decision for performing the condition-based maintenance tasks is based on the information related to the condition of an item/system established through condition checks during its operational life. This indicates that inspection-based maintenance strategy has achieved the demand for increasing the level of utilisation of an item/system. However, the system availability may not increase, due to an increased number of interruptions of the operation caused by increasing the number of inspections. Therefore, as an alternative, examination based maintenance approach is proposed by Knezevic (1987b) for the determination of maintenance tasks based on relevant condition predictors.

Examination based maintenance provides additional information about the change in condition of the items considered during its operational life. Consequently, examination based maintenance was developed for the control of maintenance procedures [El-Haram 1995]. With more information about the process of change in condition, a higher level of utilisation of the items can be achieved whilst maintaining a low probability of failure during the operation.

It is a dynamic process because the time of the next examination is fully determined by the real condition of the system at the time of examination. Dynamic control of maintenance tasks allows each individual item to perform the requested function with the required probability of failure, as in the case of time-based preventive maintenance but with fuller utilisation of operating life, hence with a reduction of total cost of operation and production.

The critical level of the relevant condition predictor RCP_{cr}, sets the limit above which appropriate maintenance tasks should be performed. The interval between the limit ($RCP_{\lim}$) and critical values depends on the ability of the operator to measure the condition of the item through the RCP. The item under consideration could be in one of the following three states, according to the numerical value of the RCP_l:

1. $RCP_{initial} < RCP(l) < RCP_{cr}$: continue with examinations;
2. $RCP_{cr} < RCP(l) < RCP_{\lim}$: preventive maintenance task required;
3. $RCP_{\lim} < RCP(l)$: corrective maintenance task, because the failure has already occurred.

In order to minimise interruptions to the operation and maximise the availability of the system, no stoppages occur until the time to the first examination of the condition of the item, FMT_1^E. The result of the examination is given as a numerical value of the relevant condition predictor, $MRCP(FMT_1^E)$, and it presents the real condition of the item at this instant of time. The following two conditions are possible, dependent on the value recorded:

1. $MRCP(FMT_1^E) > RCP_{cr}$, which means that a prescribed maintenance task should take place.
2. $MRCP(FMT_1^E) < RCP_{cr}$, the item can continue to be used.

The question, which immediately arises here, is: when will the next examination have to be done, preserving the required reliability level? The time to the next examination depends on the difference between the RCP_{cr} and $MRCP(FMT_1^E)$. The greater the difference, the longer the (operational) time to the next examination, FMT_2^E. At the predetermined time of the next examination, FMT_n^E, either of the two conditions is possible, and the same procedure should be followed, as shown in Figure 5.12

Advantages of Examination Based Policy

The advantages of the examination-based maintenance policy are:

1. Fuller utilisation of the functional life of each individual system than in case of time -based maintenance;
2. Provision of the required reliability level of each individual system as in case of time-based maintenance;
3. Reduction of the total maintenance cost as a result of extending the realisable operating life of the system and provision of a plan for maintenance tasks from the point of view of logistic support;
4. Increased availability of the item by a reduction of the number of inspections in comparison with inspection-based maintenance.

5. Applicability to all engineering systems. The main difficulties are the selection of a relevant condition predictor and the determination of the mathematical description of the $RCP(l)$.
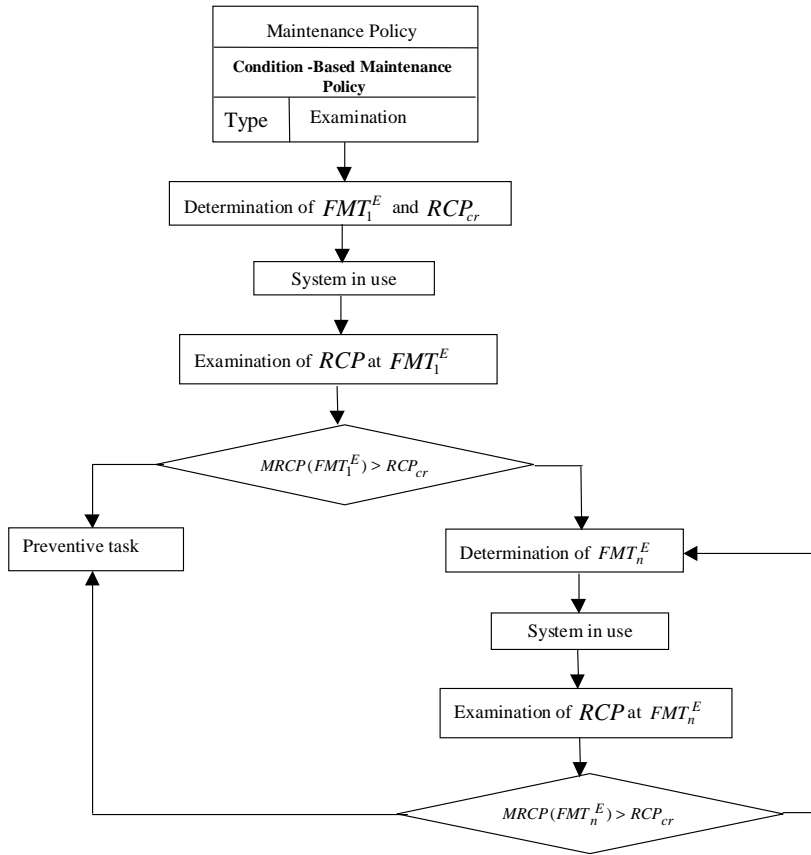
Figure 5.12 Maintenance procedure for examination based maintenance

In practice, it is impossible to eliminate all breakdowns. In some cases, it may not be economical or practical to use examination-based maintenance. Sometimes it is not physically possible to monitor the condition of all maintenance significant items. For these reasons, condition-based maintenance should not be considered to be a stand-alone policy. It should be integrated as a part of the overall maintenance policy. Thus, the optimal selection of maintenance policy for a system should include failure-based,

time-based, inspection -based and examination-based maintenance strategies. The reasons for this are summarised below:

1. Not all items in the system are significant; the suitable maintenance policy is therefore, failure-based maintenance.
2. It may not be possible or practical to monitor the condition / performance of the significant items, so the suitable maintenance policy is therefore, time-based maintenance.
3. If the condition parameters of a significant item cannot be described by a relevant condition predictor, then the suitable maintenance policy is inspection-based maintenance
4. For significant items with relevant condition predictors, the most suitable policy is examination-based maintenance.

A maintenance management approach such as reliability centred maintenance could be used to select the most applicable and effective maintenance task for each item in the system

37.   MAINTENANCE RESOURCES

It is important to stress that the number of activities, their sequence and the type and quantity of resource required mainly depends on the decisions taken during the design phase of the item/system. The time required to perform a maintenance task will also depend on decisions made during this phase, such as the complexity, testability, accessibility and any special facilities, equipment, tools or resources needed.

Resources required primarily to facilitate the maintenance process will be called *Maintenance Resources*, *MR.*. The resources needed for the successful completion of every maintenance task, could be grouped into the following categories (Knezevic 1997):

1. *Maintenance Supply Support, MSS*: is generic name which includes all spares, repair items, consumables, special supplies, and related inventories needed to support the maintenance process
2. *Maintenance Test and Support Equipment*, *MTE*: includes all tools, special condition monitoring equipment, diagnostic and check-out equipment, metrology and calibration equipment, maintenance stands and servicing and handling equipment required to support maintenance tasks associated with the item/system. Typically, MTE can be divided into two groups: *special to type equipment (STTE)* and *general (to type) equipment (GTTE).*
3. *Maintenance Personnel*, *MP*: required for the installation, check-out, handling, and sustaining maintenance of the item/system and its

associated test and support equipment are included in this category. Formal training for maintenance personnel required for each maintenance task should be considered

4. *Maintenance Facilities*, *MFC*: refers to all special facilities needed for completion of maintenance tasks. Physical plant, real estate, portable buildings, inspection pits, dry dock, housing, maintenance shops, calibration laboratories, and special repair and overhaul facilities must be considered related to each maintenance task

5. *Maintenance Technical Data*, *MTD*: necessary for check-out procedures, maintenance instructions, inspection and calibration procedures, overhaul procedures, modification instructions, facilities information, drawings and specifications that are necessary in the performance of system maintenance functions. Such data not only cover the system but test and support equipment, transportation and handling equipment, training equipment and facilities

6. *Maintenance Computer Resources*, *MCR*: refers to all computer equipment and accessories, software, program tapes/disks, data bases and so on, necessary in the performance of maintenance functions. This includes both condition monitoring and diagnostics.

On the other hand, it is important to remember that each task is performed in a specific work environment that could make a significant impact on the safety, accuracy and ease of task completion. The main environmental factors could be grouped as follows:

- space impediment (which reflects the obstructions imposed on maintenance personnel during the task execution which requires them to operate in awkward positions)
- Climatic conditions such as rain/snow, solar radiation, humidity, temperature, and similar situations, which could make significant impact on the safety, accuracy and ease of task completion.
- Platform on which maintenance task is performed (on operational site, on board a ship/submarine, space vehicle, workshops, and similar).

38. MAINTENANCE INDUCED FAILURES

Whenever the cause of failure is related to the maintenance performed on the system, we call it maintenance-induced failure *MIF*. The root cause of MIF is poor workmanship, which might lead to poor spares or material selection, improper use of test equipment, training, working environment

etc. A few examples of maintenance-induced failure are discussed in this section.

In 1991, Nigel Mansell lost his chance of becoming the Formula 1 World Champion in Portugal when one of the mechanics during a routine tyre change cross-threaded the retaining nut on the rear offside wheel. The result was that the wheel overtook the car as Nigel was exiting from the pit lane and his chance of victory and of the championship ended at that moment.

An airline pilot had a very lucky escape when he was nearly sucked through a window in the cockpit. The window was removed and replaced during a recently completed maintenance activity. When the cabin was pressurised as the aircraft climbed to cruising altitude, the window blew out. The rapid loss of pressure caused the pilot sat next to the window to be sucked through the hole. A combination of his size and the quick reactions of other members of the crew were all that saved him from a certain death. The cause of the window being blown out was that it had been refitted using under-sized screws.

In 1983, a new Air Canada Boeing 767 flying from Montreal to Edmonton ran out of fuel half way between the two at Gimli near Winnipeg. Although this was not entirely the fault of the refuellers, their miscalculations in converting between imperial and metric units was the final straw in an unfortunate sequence of events. A number of recommendations followed this incident which should mean that it never happens again (provided everyone follows the procedures correctly).

A few years ago, a team of "experienced" mechanics thought they knew how to do a particular maintenance task so did not follow the instructions in the maintenance manuals. The result was a cost of several million pounds sterling and a number of aircraft being out of service for considerably longer than they should have been.

These are extreme examples of what may be considered as "maintenance induced failures". They are also ones where it was relatively easy to determine the cause(s).

One of the major causes for accidental damage to components (from line replaceable units to parts) is the need to remove them in order to access other components. Using CATIA and EPIC (or similar systems) can do a great deal to aid the task of making components accessible and removing interference provided, of course, the design team are aware of these needs and their importance to the operational effectiveness of the aircraft.

Fasteners not properly tightened and locked (where appropriate) can work loose. Similarly, if they are not "captured" then there is a danger of them being "lost" when undone. If they are inside the engine or engine nacelle

they may be sucked into the delicate machinery almost certainly causing extensive and expensive damage. Fasteners over-tightened may cause distortion resulting in leaks or damage, which may again have serious consequences. Consistent and sensible use of fasteners can not only reduce such problems but will also reduce the parts list and hence improve the supportability of the aircraft.

Some spare parts may be expensive or difficult to obtain. There may be a temptation to use alternative sources (other than those authorised). In many cases these may be made from inferior materials or to less demanding tolerances and quality standards. The use of such rogue parts may result in premature component failure and, possible, serious damage. Configuration control and full traceability of parts is an essential element of aircraft safety but, until practical electronic tagging of all parts becomes available, it will remain difficult to police effectively.

39.   MAINTENANCE COST

The world's airlines spend around $21 billion on maintenance, out of which 21% is spent on line maintenance, 27% on heavy maintenance, 31% on engine overhaul, 16% on component overhaul and the remainder on modifications and conversions (M Lam 1995). Repair and maintenance of building stock in the UK represents over 5% of Gross Domestic Product, or £36 billion at 1996 [Building maintenance information report 254,1996]. Maintenance and repair costs can be two to three times the initial capital costs, over the life of many types of buildings.

If one recognises that maintenance is essentially the management of failure then clearly, this expenditure is primarily the result of poor quality and unreliability. However, since it is impossible to produce a system which will never fail if operated for long enough we must consider ways in which the costs of maintenance can be kept to a minimum whilst ensuring system availability, safety and integrity.

We have already seen that there are many factors which can affect the costs of maintaining a system. Whilst the original design will be a major influencing factor on these costs, the operators and maintainers of the system can, nonetheless, do much to minimise the cost of ownership by adopting the most suitable maintenance policies for the conditions prevailing.

39.1   Cost of Maintenance Task

The cost of the maintenance task is the cost associated with each corrective or preventive task, whether time-based or condition-based. The expected corrective maintenance cost is the total cost of maintenance resources needed to repair or replace failed items. Similarly, the expected preventive maintenance cost is the total cost of maintenance resources needed to inspect and/or examine an item before failure takes place and to replace any items rejected. Thus, the total maintenance cost throughout the life of a systems/product is the sum of the corrective and preventive maintenance costs and the overhead costs, which consist of all costs other than direct material, labour and plant equipment. The cost of maintenance task can be divided into two categories:

39.2   Direct cost of maintenance task

The direct cost associated with each maintenance task, CMT, is related to the cost of maintenance resources, CMR, which are mentioned in Section 9. This is the cost of the maintenance resources directly used during the execution of the maintenance task, which is defined as:

$$CMT = C_s + C_m + C_p + C_{te} + C_f + C_d \tag{5.11}$$

Where: C_s = cost of spare parts, C_m = cost of material, C_p = cost of personnel, C_{te} = cost of tools and support equipment, C_f = cost of facilities and C_d = cost of technical data.

39.3   Indirect cost of maintenance task

Indirect costs includes as management and administration staff needed for the successful completion of the task and the cost of the consequences of not having the system available which is related to a complete or partial loss of production and/or revenue. It also includes the overhead costs, i.e. salaries of employers, heating, insurance, taxes, facilities, electricity, telephone, IT, training and similar which are incurred while the item is in state of failure (and, of course, not included in the direct costs). These costs should not be neglected, because they could be even higher than the other cost elements.

Cost of lost production and/or revenue, CLR, is directly proportional to the product of the length of the time which the system spends in the state of failure (down time) and the income hourly rate, IHR, which is the money the system would earn whilst in operation. Thus, the cost of lost revenue could be determined according to the following expression:

$$CLR = (DMT + DST) \times IHR = DT \times IHR \qquad (5.12)$$

Where DMT is duration of maintenance task, DST is duration of support task and DT is total down time. Note for systems that are not normally in continuous operation, the downtime should take account of the proportion of the time the system would normally be expected to be operational. In particular, preventative, planned or scheduled maintenance would normally be done when the system would be expected to be idle and would only count as "downtime" for any period that the system would be expected to be operational. Thus, for example, if an airliner is not permitted to fly between the hours of 21:00 and 07:00 then any maintenance tasks undertaken and completed during those 10 hours would not affect the revenue-earning capacity of the aircraft.

39.4 Total cost of maintenance task

The total cost of maintenance task is the sum cost of direct and indirect costs, thus:

$$CMT = CMR + CLR \qquad (5.13)$$

Making use of the above equations the expression for the cost of the completion of each maintenance task is defined as:

$$CMT = C_s + C_m + C_p + C_{te} + C_f + C_d + (DMT + DST) \times IHR \quad (5.14)$$

It is necessary to underline that the cost defined by the above expression could differ considerably, due to:

1. Adoption of different maintenance policies
2. The direct cost of each maintenance task
3. Consumption of maintenance resources
4. Duration of maintenance task, DMT^c, DMT^p, DMT^I and DMT^E

5. Frequency of preventive maintenance task, FMT^L, the frequency of inspection, FMT^I and frequency of examination FMT^E
6. Duration of support task, DST^c, DST^p, DST^I and DST^E
7. The expected number of maintenance tasks $NMT(T_{st})$ performed during the stated operational length, L_{st}. For example, in the case of FBM, $NMT(T_{st}) = \dfrac{T_{st}}{MTTF}$
8. Different probability distributions and different values which random variables
$DMT^c, DMT^p, DMT^I, DMT^E, DST^c, DST^p, DST^I$ and DST^E
can take.
9. Indirect costs of maintenance tasks.

Thus, the general expression for the cost of each maintenance task will have different data input for different maintenance policies, as shown below:

$$CMT^c = C_{sp}^c + C_m^c + C_p^c + C_{te}^c + C_f^c + C_d^c + (DMT^c + DST^c) \times IHR^c$$

$$CMT^p = C_{sp}^p + C_m^p + C_p^p + C_{te}^p + C_f^p + C_d^p + (DMT^p + DST^p) \times IHR^p$$

$$CMT^I = C_{sp}^I + C_m^I + C_p^I + C_{te}^I + C_f^I + C_d^I + (DMT^I + DST^I) \times IHR^I$$

$$CMT^E = C_{sp}^E + C_m^E + C_p^E + C_{te}^E + C_f^E + C_d^E + (DMT^E + DST^E) \times IHR^E$$

Where: CMT^c is related to the cost of each maintenance task performed after the failure, CMT^p is cost in the case of time based maintenance CMT^I is cost of inspection based maintenance and CMT^I is cost of examination based maintenance.
The expected total maintenance cost for a stated time, $CMT(T_{st})$, is equal to the product of the maintenance cost for each maintenance task and the expected number of maintenance tasks performed during the stated time, $NMT(T_{st})$, thus:

$$
\begin{aligned}
CMT(T_{st}) = CMT^c \times NMT^c(T_{st}) + CMT^p \times NMT^p(T_{st}) + \\
CMT^I \times NMT^I(T_{st}) + CMT^E \times NMT^E(T_{st})
\end{aligned}
\tag{5.15}
$$

39.5   Factor Affecting Maintenance Costs

Maintenance cost could be affected by the following factors:

1. Supply responsiveness or the probability of having a spare part available when needed, supply lead times for given items, levels of inventory, and so on.
2. Test and support equipment effectiveness, which is the reliability and availability of test equipment, test equipment utilisation, system test thoroughness, and so on.
3. Maintenance facility availability and utilisation.
4. Transportation times between maintenance facilities.
5. Maintenance organisational effectiveness and personnel efficiency.
6. Durability and reliability of items in the system
7. Life expectancy of system
8. Expected number of maintenance tasks
9. Duration of maintenance and support task
10. Maintenance task resources

In order to reduce maintenance costs, it is necessary that the impact of the above factors should be reduced and/or controlled.

In calculating the various cost elements of maintenance, it is important to recognise that facilities, equipment, and personnel may be used for other tasks. For example, mechanics in the armed forces may be put on guard duty or provide a defence role when not performing maintenance tasks. Thus eliminating all maintenance tasks at first line (or O-Level) may not necessarily lead to a significant reduction in the personnel deployed or, indeed, in the operational costs.

40.   AIRCRAFT MAINTENANCE - CASE STUDY

For every commercial airline, maintenance is one of the most important functions to assure safe operation. Federal Aviation Regulation (FAR) require that, no person may operate an aircraft unless the mandatory replacement times, inspection intervals and related procedures or alternative inspection intervals and related procedures set forth in the operations specifications or inspection program has been complied with. All aircraft must follow a maintenance program that is approved by a regulatory authority such as FAA (Federal Aviation Administration, USA) and CAA (Civil Aviation Authority, UK).   Each airline develops its own maintenance plan, based on the manufacturer's recommendations and by

considering its own operation. Thus, two different airlines may have slightly different maintenance program for same aircraft model used under similar operating conditions. Aircraft maintenance is reliability centred. It is claimed that each aircraft receives approximately 14 hours of maintenance for every hour it flies (R Baker, 1995). Maintenance accounts for approximately 10% of an airline's total costs. On average a typical Boeing 747 will generate a total aircraft maintenance cost of approximately $1,700 per block hour.

Aircraft maintenance can be categorised as:

1. Routine scheduled maintenance.
2. Non-routine maintenance.
3. Refurbishment.
4. Modifications.

Routine Scheduled Maintenance

Scheduled maintenance tasks are required at determinant recurring intervals or due to Airworthiness Directives (AD). The most common routine maintenance is visual inspection of the aircraft prior to a scheduled departure (known as *walk around*) by pilots and mechanics to ensure that there are no obvious problems. Routine maintenance can be classified as:

1. Overnight maintenance.
2. Hard time maintenance.
3. Progressive Inspection.

Overnight maintenance normally includes low level maintenance checks, minor servicing and special inspections done at the end of the working for about one to two hours to ensure that the plane is operating in accordance with Minimum Equipment List. Overnight maintenance provides an opportunity to remedy passenger and crew complaints (M Lam, 1995).
Hard time is the oldest primary maintenance process. Hard time requires periodic overhaul or replacement of affected systems/components and structures and is flight, cycle and calendar limited. That is, as soon as the

component age reaches it hard time it is replaced with a new component. Most of the rotating engine units are hard timed. The purpose of hard time maintenance is to assure operating safety of component or system, which have a limited redundancy.

Progressive inspection groups like time related maintenance tasks into convenient 'blocks' so that maintenance workload becomes balanced with time and maintenance can be accomplished in small 'bites' making equipment more available. Grouping maintenance tasks also helps better utilisation of the maintenance facilities. These maintenance task groups are (detailed information can be found in M Lam (1995) and L R Crawford, 1995):

1. *Pre-flight* – Visual inspections carried out by the mechanic and the pilots to ensure that there are no obvious problems.

2. *A Check* – Carried out approximately every 150 flight hours, which includes selected operational checks (general inspection of the interior/exterior of the aircraft), fluid servicing, extended visual inspection of fuselage exterior, power supply and certain operational tasks. During A check, the aircraft is on ground for approximately 8 to 10 hours and requires approximately 60 labour hours.

3. *B Check* – Occurs about every 750 flight hours and includes some preventive maintenance such as engine oil spectro-analysis, oil-filter are removed and checked, lubrication of parts as required and examination of airframe. Also incorporates A-check. The aircraft could be on ground for 10 hours and will require approximately 200 labour hours.

4. *C Check* – Occurs every 3, 000 flight hour (approximately 15 months) and includes detailed inspection of airframe, engines, and accessories. In addition, components are repaired, flight controls are calibrated, and major internal mechanisms are tested. Functional and operational checks are also performed during C-check. It also includes both A and B checks. The aircraft will be on ground for 72 hours and will require approximately 3,000 labour hours.

5. *D Check* – This is the most intensive form of routine maintenance occurs about 20,000 flight hours (six to eight years). It is an overhaul that returns the aircraft to its original condition, as far as possible. Cabin interiors including seats, galleys, furnishings etc are removed to allow careful structural inspections. The aircraft is on ground for about 30 days and will require approximately 20,000 labour hours.

A and B checks and overnight maintenance are instances of line maintenance (performed upon the aircraft incidental to its scheduled revenue operations), often carried out an airport. C and D checks, however

are heavy maintenance that requires special facilities and extensive labour. The task intervals for various checks mentioned above could vary significantly. The recommended time intervals for different aircraft models are given in Table 5.3 (*Aircraft Economics).*

Table 5.3 Different scheduled checks in a commercial aircraft

Aircraft Type	A check Flight hours	B Check Flight hours	C Check Flight hours	D Check Flight hours
Boeing 707	90		450	14,000
Boeing 727	80	400	1,600	16,000
Boeing 737-100	125	750	3,000	20,000
Boeing 747-100	300		3,600	25,000
DC-8	150	540	3,325	23,745
DC-9	130	680	3,380	12,600

Non-routine maintenance refers to the maintenance tasks that has to be performed on regular basis during checks, but which is not specified as routine maintenance task on the job cards of the maintenance schedule. Non-routine maintenance shouldn't be confused with unscheduled maintenance, which is repairs that have to be done as a result of an unexpected failure such as accidental damage (such as bird strike) to critical components or a response to airworthiness directives (AD). As the aircraft age, they require more maintenance due to fatigue and corrosion. The most significant of these aging aircraft airworthiness directives concerns Boeing 747. The fuselage of the Boeing 747 is built in sections as separate entities and then assembled during the aircraft production phase. The fuselage is built in five sections and the points at which these sections are joined are called the production breaks. Section 41 is the section from the nose to just aft of the forward passenger entry (*Maintaining the Boeing 747, Aircraft Economics, 1994).* The modification of Section 41, which is the area ahead of the forward passenger doors, requires approximately 60,000-70,000 man-hours to complete and requires replacement of most of the structural components (L Crawford, 1995).

Availability, Design for Reliability and Data Analysis

Course Material

Course Instructor: Professor U Dinesh Kumar
Indian Institute of Management Bangalore

Chapter 10

Availability

There is nothing in this world constant, but inconsistency

Jonathan Swift

Availability is used to measure the combined effect of reliability, maintenance and logistic support on the operational effectiveness of the system. A system, which is in a state of failure, is not beneficial to its owner; in fact, it is probably costing the owner money. If an aircraft breaks down, it cannot be used until it has been declared airworthy. This is likely to cause inconvenience to the customers who may then decide to switch to an alternative airline in future. It may disrupt the timetables and cause problems for several days.

As mentioned in Chapter 9, most large airliners have a very high utilisation rate with the only down time being to do a transit check, unload, clean the cabin, refuel, restock with the next flight's foods and other items, and reload with the next set of passengers and baggage. The whole operation generally takes about an hour. Any delay may cause it to miss its take off slot and more significantly its landing slot, since an aircraft cannot take-off until it has been cleared to land, even though this may be 12 hours later. Many airports close during the night to avoid unacceptable levels of noise pollution. If the particular flight was due to land just before the airport closes, missing its slot could mean a delay of several hours.

An operator of a system would like to make sure that the system will be in a state of functioning (*SoFu*) when it is required.   Designers and manufacturers know that they are unlikely to remain in business for very long if their systems do not satisfy the customers' requirements in terms of operational effectiveness. Many forms of availability are used to measure the effectiveness of the system.   Inherent availability, operational availability and achieved availability are some of the measures used to

quantify whether an item is in an operable state when required. Availability is defined as:

The probability that an item is in state of functioning at a given point in time (point availability) or over a stated period of time (interval availability) when operated, maintained and supported in a prescribed manner.

It is clear from the above definition that availability is a function of reliability, maintainability and supportability factors (Figure 10.1).

```
┌─────────────────┐
│   Reliability   │────────┐
└─────────────────┘        │
                           │        ┌─────────────────┐
┌─────────────────┐        │        │                 │
│ Maintainability │────────┼───────▶│  Availability   │
└─────────────────┘        │        │                 │
                           │        └─────────────────┘
┌─────────────────┐        │
│  Supportability │────────┘
└─────────────────┘
```
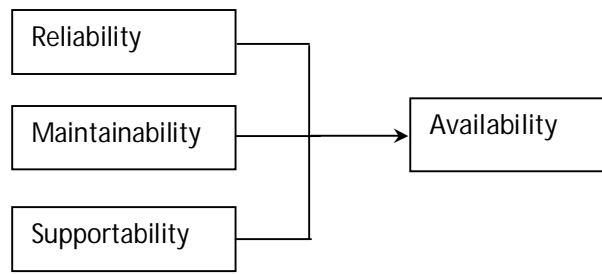
Figure 10.1 Availability as a function of reliability, maintainability and supportability

In this chapter, we look at few important availability measures such as point availability, interval availability, steady state inherent availability, operational availability and achieved availability.

10.41.   POINT AVAILABILITY

Point availability is defined as the probability that the system is in the state of functioning (*SoFu*) at the given instant of time *t*. We use the notation $A(t)$ to represent the point availability. Availability expressions for systems can be obtained by using stochastic processes. Depending on the time to failure and time to repair distributions, one can use Markov chain, renewal process, regenerative process, semi-Markov process and semi-regenerative process models to derive the expression for point availability. For example, consider an item with constant failure rate λ and constant repair rate μ. At any instant of time, the item can be in either the state of functioning (say, state 1) or in the state of failure (say, state 2). As both failure and repair

rates are constant (and thus follow exponential distribution), we can use a Markov chain to model the system to derive the availability expression.

Let $p_{ij}(h)$ denote the transition probability from state i to state j during the interval 'h' ($i,j = 1, 2$). Define, $P_i(t+h)$, as the probability that the system would be in state i at time $t+h$, for $i = 1, 2$. The expression for $P_1(t+h)$ can be derived using the following logic:

1. The system was in state 1 at time t and continues to remain in state 1 throughout the interval h.
2. The system was in state 2 at time t and it transits to state 1 during the interval h.

The corresponding expression can be written as:

$$P_1(t + h) = P_1(t) \times p_{11}(h) + P_2(t) \times p_{21}(t) \tag{10.1}$$

Using similar logic, the expression for $P_2(t+h)$ can be written as:

$$P_2(t + h) = P_1(t) \times p_{12}(h) + P_2(t) \times p_{22}(h) \tag{10.2}$$

$p_{11}(h)$ is the probability of remaining in state 1 during the interval h. The probability $p_{11}(h)$ is given by

$$p_{11}(h) = \exp(-\lambda h) \approx 1 - \lambda h \ \ \text{for} \ \lambda h \ll 1$$

$p_{21}(h)$ is the probability of entering state 1 from state 2 during the interval h. The corresponding expression is given by

$$p_{21}(h) = 1 - \exp(-\mu h) \approx \mu h \ \text{for} \ h\mu \ll 1$$

$p_{12}(h)$ is the probability of entering state 2 from state 1 during the interval h. The probability $p_{12}(h)$ is given by

$$p_{12}(h) = 1 - \exp(-\lambda h) \approx \lambda h \ \ \text{for} \ h\lambda \ll 1$$

$p_{22}(h)$ is the probability of remaining in state 2 during the interval h. The probability $p_{22}(h)$ is given by:

$$p_{22}(h) = \exp(-\mu h) \approx 1 - \mu h \ \ \text{for} \ h\mu \ll 1$$

Substituting the values of $p_{ij}(h)$ in equation (10.1) and (10.2), we get

$$P_1(t + h) = P_1(t) \times (1 - \lambda h) + P_2(t) \times \mu h$$

$$P_2(t + h) = P_1(t) \times (\lambda h) + P_2(t) \times (1 - \mu h)$$

By rearranging the terms and setting $h \to 0$, we have

$$\underset{h \to 0}{Lt} \frac{P_1(t + h) - P_1(t)}{h} = \frac{dP_1(t)}{dt} = -\lambda P_1(t) + \mu P_2(t)$$

$$\underset{h \to 0}{Lt} \frac{P_2(t + h) - P_2(t)}{h} = \frac{dP_2(t)}{dt} = \lambda P_1(t) - \mu P_2(t)$$

On solving the above two differential equations, we get

$$P_1(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} \times \exp(-(\lambda + \mu)t)$$

$P_1(t)$ is nothing but the availability of the item at time t, that is the probability that the item will be in state of functioning at time t. Thus, the point availability $A(t)$ is given by:

$$A(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} \times \exp(-(\lambda + \mu)t) \tag{10.3}$$

Substituting λ = 1/MTTF and μ = 1/MTTR in the above equation, we get

$$A(t) = \frac{MTTF}{MTTF + MTTR} + \frac{MTTR}{MTTF + MTTR} \times \exp(-(\frac{1}{MTTF} + \frac{1}{MTTR})t) \tag{10.4}$$

When the time to failure and time to repair are not exponential, we can use a *regenerative process* to derive the availability expression. If $f(t)$ and $g(t)$ represent the time-to-failure and time-to-repair distributions respectively, then the point availability $A(t)$ can be written as (Birolini, 1997):

$$A(t) = 1 - F(t) + \int_0^t \sum_{n=1}^{\infty} [f(x) * g(x)]^n [1 - F(t - x)]dx$$

where $[f(x)*g(x)]^n$ is the n-fold convolution of $f(x)*g(x)$. The summation $\sum_{n=1}^{\infty}[f(x)*g(x)]^n$ gives the renewal points $f(x)*g(x)$, $f(x)*g(x)*f(x)*g(x)$, ... lies in $[x, x+dx]$, and $1 - F(t-x)$ is the probability that no failures occur in the remaining interval $[x, t]$.

10.41.1   Average Availability

Interval availability, *AA*(t), is defined as the expected fractional duration of an interval (0, t] that the system is in state of functioning. Thus,

$$AA(t) = \frac{1}{t}\int_0^t A(x)dx \tag{10.5}$$

where *A(x)* is the point availability of the item as defined in equation (10.3) and (10.4). For an item with constant failure rate λ and constant repair rate μ, the average availability is given by:

$$AA(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{(\lambda+\mu)^2 t}[1-\exp(-(\lambda+\mu)t)] \tag{10.6}$$

10.41.2 Inherent Availability

Inherent availability (or steady-state availability), A_i, is defined as the steady state probability (that is, $t \to \infty$) that an item will be in a state of functioning, assuming that this probability depends only on the time-to-failure and time to repair distributions. It is assumed that any support resources that are required are available without any restriction. Thus, the inherent availability is given by:

$$A_i = \underset{t\to\infty}{Lt} A(t) = \frac{MTTF}{MTTF + MTTR} \tag{10.7}$$

The above result is valid for any time to failure function *F*(*t*) and any time to repair distribution G(*t*) (*Birolini*, 1997). Also, in the case of constant failure rate λ and constant repair rate μ, the following inequality is true.

$$|A(t) - A_i| \le \exp(-t/MTTR) \qquad (10.8)$$

Example 10.1
Time to failure distribution of a digital engine control unit (DECU) follows an exponential distribution with mean time between failures 1200 hours and the repair time also follows an exponential distribution with mean time to repair 400 hours.

1. Plot the point availability of the DECU.
2. Find the average availability of the DECU during first 5000 hours.
3. Find the inherent availability.

SOLUTION:

1. The point availability of the DECU is calculated using the equation (10.4). Figure 10.2 depicts the point availability of the system.
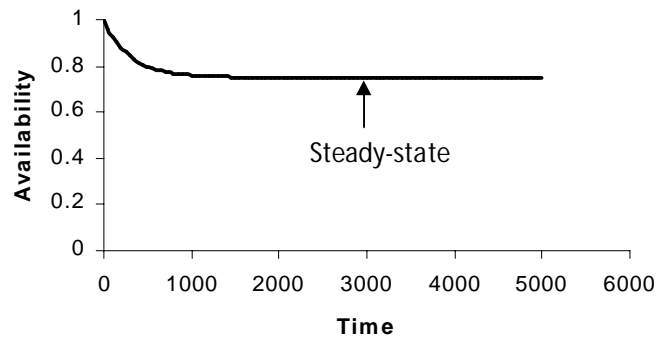
Figure 10.2 Point availability of DECU

2. The average availability of the system during 5000 hours of operation is given by:

$$AA(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{(\lambda + \mu)^2 t}[1 - \exp(-(\lambda + \mu)t)]$$

Substituting the values of λ (= 1/1200) and μ (=1/400), we get the value of the average availability during 5000 hours as 0.7649.

3. The inherent availability is given by

$$A_i = \frac{MTTF}{MTTF + MTTR} = \frac{1200}{1200 + 400} = 0.75$$

Thus, the steady state availability of the system is 0.75 or 75%.

10.41.3 System Availability of different reliability block diagrams

Availability of a system with series reliability block diagram with n items is given by

$$A_s(t) = \prod_{k=1}^{n} A_i(t) \tag{10.9}$$

where $A_i(t)$ is the point availability of ith item. The inherent availability of the system is given by

$$A_{i,s} = \prod_{k=1}^{n} \frac{MTTF_i}{MTTF_i + MTTR_i} \tag{10.10}$$

For a series system with all the elements having constant failure and repair rates, the system inherent availability

$$A_{i,s} = \frac{MTTF_s}{MTTF_s + MTTR_s} \tag{10.11}$$

$MTTF_s$ and $MTTR_s$ are system mean time to failure and system mean time to repair respectively. Let λ_i and μ_i represent the failure rate and repair rate of item i respectively. $MTTF_s$ and $MTTR_s$ are given by

$$MTTF_s = \frac{1}{\sum_{i=1}^{n} \lambda_i}$$

$$MTTR_s = \sum_{i=1}^{n} \frac{\lambda_i MTTR_i}{\lambda_s}, \text{ where } \lambda_s = \sum_{i=1}^{n} \lambda_i$$

Availability of a parallel system with n items is given by

$$A_s(t) = 1 - \prod_{i=1}^{n}[1 - A_i(t)] \tag{10.12}$$

Example 10.2

A series system consists of four items. The time to failure and the time to repair distributions of the different items are given as given in Tables 10.1 and 10.2. Find the inherent availability of the system.

Table 10.1. Time to failure distribution for different items.

Item Number	Distribution	Parameters
Item 1	Weibull	$\eta = 2200$ hours $\beta = 3.7$
Item 2	Exponential	$\lambda = 0.0008$ per hour
Item 3	Weibull	$\eta = 1800$ hours $\beta = 2.7$
Item 4	Normal	$\mu = 800$ hours $\sigma = 180$ hours

Table 10.2. Time to repair distribution for different items

Item number	Distribution	Parameters
Item 1	Lognormal	$\mu_l = 3.25$ and $\sigma_l = 1.25$
Item 2	Normal	$\mu = 48$ hours $\sigma = 12$ hours
Item 3	Lognormal	$\mu_l = 3.5$ and $\sigma_l = 0.75$
Item 4	Normal	$\mu = 72$ hours $\sigma = 24$ hours

SOLUTION:

First we calculate $MTTF_i$ and $MTTR_i$ for different items:

$$MTTF_1 = \eta \times \Gamma(1 + \frac{1}{\beta}) = 2200 \times \Gamma(1 + \frac{1}{3.7}) = 2200 \times 0.902 = 1984.4$$

$$MTTF_2 = 1/\lambda = 1/0.0008 = 1250, \quad MTTF_3 = 1600.2, \quad MTTF_4 \quad = 800$$

$$MTTR_1 = \exp(\mu_l + \sigma_l^2/2) = 56.33 \text{ hours}, \quad MTTR_2 \quad = 48 \quad \text{hours}$$

$$MTTR_3 = \exp(\mu_l + \sigma_l^2/2) = 43.87 \text{ hours}, MTTR_4 = 72 \text{ hours}$$

Inherent availability, A_i, for item i can be calculated using the equation (10.11). Substituting the values of $MTTF_i$ and $MTTR_i$ in equation (10.11), we have

$A_1 = 0.9723$, $A_2 = 0.9630$, $A_3 = 0.9733$, $A_4 = 0.9174$

The system availability is given by

$$A_s = \prod_{i=1}^{4} A_i = 0.8362$$

10.42.   ACHIEVED AVAILABILITY

Achieved availability is the probability that an item will be in a state of functioning (*SoFu*) when used as specified taking into account the scheduled and unscheduled maintenance; any support resources needed are available instantaneously. Achieved availability, A_a, is given by

$$A_a = \frac{MTBM}{MTBM + AMT} \tag{10.13}$$

MTBM is the mean time between maintenance and AMT is active maintenance time. The mean time between maintenance during the total operational life, T, is given by:

$$MTBM = \frac{T}{M(T) + T/T_{sm}} \tag{10.14}$$

M(T) is the renewal function, that is the expected number of failures during the total life T. T_{sm} is the scheduled maintenance interval (time between scheduled maintenance). The above expression is valid when after each scheduled maintenance, the item is 'as-bad-as-old' and after each

corrective maintenance the item is 'as-good-as-new'. The active maintenance time, AMT, is given by:

$$AMT = \frac{M(T) \times MTTR + (T/T_{sm})MSMT}{M(T) + T/T_{sm}} \qquad (10.15)$$

MTTR stands for the mean time to repair and MSMT is the mean scheduled maintenance time.

Example 10.3

Time to failure distribution of an engine monitoring system follows a normal distribution with mean 4200 hours and standard deviation 420 hours. The engine monitoring system is expected to last 20,000 hours (subject to corrective and preventive maintenance). A scheduled maintenance is carried out after every 2000 hours and takes about 72 hours to complete the task. The time to repair the item follows a lognormal distribution with mean time to repair 120 hours. Find the achieved availability for this system.

SOLUTION:

Mean time between maintenance, MTBM, is given by

$$MTBM = \frac{T}{M(T) + T/T_{sm}} = \frac{20000}{M(20000) + 20000/2000}$$

M(20000) for normal distribution with mean 4200 hours and standard deviation 420 hours is given by

$$M(20000) = \sum_{n=1}^{\infty} \Phi(\frac{20000 - n \times 4200}{\sqrt{n} \times 420}) = 4.1434$$

$$MTBM = \frac{20000}{4.1434 + 10} \approx 1414 \, \text{hours}$$

The active maintenance time is given by:

$$AMT = \frac{M(T) \times MTTR + (T/T_{sm})MSMT}{M(T) + T/T_{sm}}$$

$$= \frac{4.1434 \times 120 + 10 \times 72}{4.1434 + 10} \approx 86.06$$

The achieved availability of the system is given by:

$$A_a = \frac{MTBM}{MTBM + AMT} = \frac{1414}{1414 + 86.06} = 0.9426$$

10.43.   OPERATIONAL AVAILABILITY

Operational availability is the probability that the system will be in the state of functioning (*SoFu*) when used as specified taking into account maintenance and logistic delay times. Operational availability, A_o, is given by

$$A_o = \frac{MTBM}{MTBM + DT} \tag{10.16}$$

where, MTBM is the mean time between maintenance (including both scheduled and unscheduled maintenance) and DT is the Down time. The mean time between maintenance during the total operational life, T, is given by:

$$MTBM = \frac{T}{M(T) + T/T_{sm}} \tag{10.17}$$

M(T) is the renewal function, that is the expected number of failures during the total life T. T_{sm} is the scheduled maintenance interval (time between scheduled maintenance).   The system down time DT is given by:

$$DT = \frac{M(T) \times MTTRS + (T/T_{sm})MSMT}{M(T) + T/T_{sm}} \tag{10.18}$$

MTTRS stands for the mean time to restore the system and MSMT is the mean scheduled maintenance time. MTTRS is given by

MTTRS = MTTR + *MLDT*

where MLDT is the mean logistic delay time for supply resources. In the absence of any scheduled maintenance the operational availability can be calculated using the following simple formula

$$A_O = \frac{MTBF}{MTBF + MTTR + MLDT} \qquad (10.19)$$

Example 10.4

In the previous example, assume that whenever a system fails it takes about 48 hours before all the necessary support resources are available. Find the operational availability.

SOLUTION

MTBM is same as in the previous example and is equal to 1414 hours. The mean time to restore the system is given by

MTTRS = MTTR + MLDT = 120 + 48 = 168 hours

The system down time is given by

$$DT = \frac{M(T) \times MTTRS + (T/T_{sm})MSMT}{M(T) + T/T_{sm}}$$

$$= \frac{4.1434 \times 168 + 10 \times 72}{14.1434} = 100.12 \text{ hours}$$

The operational availability of the system is given by

$$A_O = \frac{MTBM}{MTBM + DT} = \frac{1414}{1414 + 100.12} = 0.9338$$

Chapter 11

Design for Reliability, Maintenance and Logistic Support

Reliability, maintenance and supportability should be designed into the product. Design phase is particularly important for any product as the decisions made during this stage can determine how reliable the product is going to be as well as the maintainability and supportability of that product. In this chapter, we would like to discuss few tool and techniques that can be used at the design stage to improve the RMS characteristics.

44.   RELIABILITY ALLOCATION

Reliability allocation is a process by which the system's reliability requirements is divided into sub-system and component reliability requirements.

45.   FAILURE MODES, EFFECTS AND CRITICALITY ANALYSIS (FMECA)

The failure modes, effects and criticality analysis (FMECA) is a systematic method for examining all modes through which a failure can occur,

potential effects of these failures on the system performance and their relative severity in terms of safety, extend of damage, and impact on mission success. FMECA is performed to identify reliability, maintenance, safety and supportability problems resulting from the effects of a product/process failure. It is an excellent methodology for identifying and investigating potential product weaknesses. FMECA establishes a detailed study of the product design, manufacturing operation or distribution to determine which features are critical to various modes of failure.   The FMECA concept was developed by US defence industries in the 1950s, to improve the reliability of military equipment.   Since then, FMECA has become an important tools applied by almost all industries around the world to improve the reliability, maintainability and supportability of their product. It is claimed that a more rigorous FMECA analysis would have avoided the disastrous explosion of the Challenger launch on 28th January 1986.

The three principal study areas in FMECA analysis are the failure mode, failure effect and failure criticality. Failure mode analysis lists all possible mode the failure would occur which include the condition, the components involved, location etc. The failure effect analysis includes the study of the likely impact of failure on the performance of the whole product and the process. The criticality analysis examines how critical a failure would be for the operation and safe use of the product. The criticality might range from minor failure through lowering of performance, shutdown of the product, safety and environmental hazard to a catastrophic failure. This analysis is best utilised during the early design and development phase of new systems, and in the evaluation of existing system (D Verma, 1993).

The actual FMECA performed could be both quantitative and qualitative based on the information available to the analyst. Input requirements for FMECA analysis include reliability data, their modes of failure, and the estimated criticality of the failures.   Additionally, the probabilities of detection for the various failure modes are also required. A prerequisite for the successful completion of FMECA is good knowledge of, and familiarity with the product/process being analysed and its design and functionality (D Verma, 1993).

45.1   Procedural Steps in the FMECA analysis

The procedural steps in FMECA analysis depend to a certain extend what product or process is being examined. The sequence of steps followed to

accomplish the failure modes, effect and criticality analysis is depicted in Figure 11.1. The following are the key steps involved in the FMECA analysis:

1. Identification of the system requirements, by defining the basic requirements for the system in terms of input criteria for design. During the system requirement definition, the following tasks should be addressed (Refer to Blanchard and Fabrycky 1999 for detailed discussion).
   – What is expected from the system in terms of operation and performance.
   – What is the customer requirements with respect to reliability, maintainability and supportability
   – How the system is used in terms of hours of operation/number of cycles per day etc.
   – What are the requirements for disposal after the system is withdrawn from service.
2. Accomplish functional analysis (Functional analysis is a systematic approach to system design and development, which employs functional approach as a basis for identification of design requirements for each hierarchical level of the system. Functional analysis is accomplished through functional flow diagram that portrays the system design requirements illustrating series and parallel relationships and functional interfaces).
3. Accomplish requirements allocation, that is for a specified requirement at system level, what should be specified at unit and assembly level. System effectiveness factors such as reliability, maintainability and supportability specified at system level are allocated to unit and assembly level.
4. Identification of all possible failure modes for the system as well as the subsystem, modules and components.
5. Determine cause of failures, which could be design and manufacturing deficiency, ageing and wear-out, accidental damage, transportation and handling, maintenance induced failures.
6. Identify the effects of failure. Effect of failure might range from catastrophic failure to minor performance degradation.
7. Assess the probability of failure. This can be achieved by analysing the failure data and identifying the time-to-failure distribution.
8. Identify the criticality of failure. Failure criticality can be classified in any one of four categories, depending upon the failure effects as follows

    a) *Minor failure* – Any failure that doesn't have any noticeable affect on the performance of the system.

    b) *Major failure* – Any failure that will degrade the system performance beyond an acceptable limit.

    c) *Critical failure* – Any failure that would affect safety and degrade the system beyond an acceptable limit.

    d) *Catastrophic failure* – Any failure that could result in significant system damage and may cause damage to property, serious injury or death.

9. Compute the *Risk Priority Number* (RPN) by multiplying the probability of failure, the severity of the effects and the likelihood of detecting a failure mode.

10. Initiate corrective action that will minimise the probability of failure or effect of failure that show high RPN.
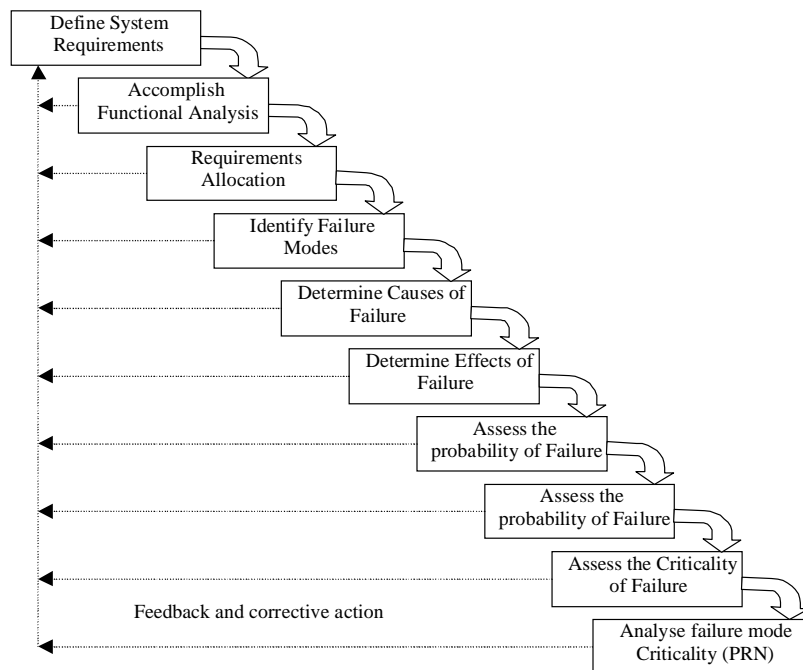
Figure 11.1 Sequence of steps involved in FMECA

45.2   **Risk Priority Number**

Risk Priority Numbers play a crucial role in selecting the most significant item that will minimise the failure or effect of failure. As mentioned earlier, RPN is calculated by multiplying the probability of failure, the severity of the effects of failure and likelihood of failure detection. That is:

$$RPN = FP \times FS \times FD \qquad\qquad (11.1)$$

Where, FP is the Failure probability, FS is the failure severity and FD denotes the failure detection probability. Tables 11.1 – 11.3 gives possible ratings for probability of failure, severity of failure and failure detection. Note that, the ratings given in the tables 11.1-11.3 are only suggested ratings.

Table 11.3. Rating scales for occurrence of failure

Description	Rating
Remote probability of occurrence	1
Low probability of occurrence	2 - 3
Moderate probability of occurrence	4 - 6
High probability of occurrence	7 – 8
Very High probability of occurrence	9 - 10

Table 11.2 Rating scales for severity of failure

Description	Rating
Minor failure	1 – 2
Major Failure	3 – 5
Critical Failure	6 – 9
Catastrophic Failure	10

Description	Rating

Table 11.3. Rating scales for detection of failure

Description	Rating	Probability of Detection
Remote probability of detection	1	0 – 0.05
Low probability of detection	2 - 3	0.06 – 0.15
Moderate probability of detection	4 - 5	0.16 – 0.35
High probability of detection	6 - 8	0.36 – 0.75
Very high probability of detection	9 - 10	0.76 – 1.00

Assume that a failure mode has following ratings for probability of failure, failure severity and failure detection:

Failure probability = 7

Failure severity = 4

Failure detection = 5

Then the risk priority number for this particular failure mode is given by $7 \times 4 \times 5 = 140$. Risk priority number for all the failure modes are calculated and priority is given to the one with highest RPN for eliminating the failure. This is usually achieved using Pareto analysis with a focus on failure mode, failure cause and failure criticality. Outputs from a properly conducted FMECA can be used in developing a cost effective maintenance analysis, system safety hazard analysis, and logistic support analysis.

46. FAULT TREE ANALYSIS (FTA)

Fault tree analysis is a deductive approach involving graphical enumeration and analysis of the different ways in which a particular system failure can occur, and the probability of its occurrence. It starts with a top-level event (failure) and works backward to identify all the possible causes and therefore the origins of that failure. During the very early stages of system design process, and in the absence of information required to complete a FMECA, fault tree analysis (FTA) is often conducted to gain insight into critical aspects of selected design concepts. Usually, a separate fault tree is developed for every critical failure mode or undesired Top-Level event. Attention is focused on this top-level event and the first-tier causes associated with it. Each first-tier cause is next investigated for its causes, and this process is continued. This '*Top-Down*' causal hierarchy and the associated probabilities, is called a Fault Tree.

One of the outputs from a fault tree analysis is the probability of occurrence of the top-level event or failure. If this probability is unacceptable, fault tree analysis provides the designers with an insight into aspects of the system to which redesign can be directed or compensatory provisions be provided such as redundancy. The FTA can have most impact if initiated during the conceptual and preliminary design phase when design and configuration changes can be most easily and cost effectively implemented.

The logic used in developing and analysing a fault tree has its foundations in Boolean Algebra. The following steps are used to carry out FTA (Figure 11.2).

1. *Identify the top-level event* – The most important step is to identify and define the top-level event. It is necessary to specific in defining the top-level event, a generic and non-specific definition is likely to result in a broad based fault tree which might be lacking in focus.
2. *Develop the initial fault tree* – Once the top-level event has been satisfactorily identified, the next step is to construct the initial causal hierarchy in the form of a fault tree. Techniques such as *Ishikawa's cause and effect diagram* can prove beneficial. While developing the fault tree all hidden failures must be considered and incorporated. For the sake of consistency, a standard symbol is used to develop fault trees. Table 11.4 depicts the symbols used to represent the causal hierarchy and interconnects associated with a particular top-level event. While

constructing a fault tree it is important to break every branch down to a reasonable and consistent level of detail.

3. *Analyse the Fault Tree* – The third step in FTA is to analyse the initial fault tree developed. The important steps in completing the analysis of a fault tree are 1. Delineate the minimum cut-sets, 2. Determine the reliability of the top-level event and 3. Review analysis output.
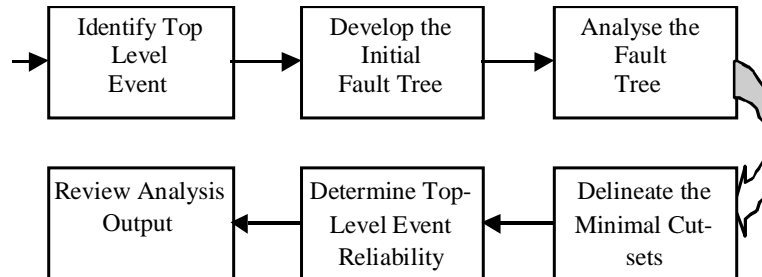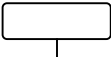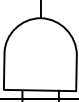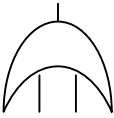
Figure 11.2 Steps involved in a fault tree analysis.

Table 11.4. Fault tree construction symbols

Symbol	Description
	The Ellipse represents the top-level event (thus always appears at the very top of the fault tree).
	The rectangle represents an intermediate fault event. A rectangle can appear anywhere in a fault tree except at the lowest level in the hierarchy.
	A circle represents the lowest level failure event, also called a basic event.
	The diamond represents an undeveloped event, which can be further broken. Very often, undeveloped events have a substantial amount of complexity below and can be analysed through a separate fault tree.
	This symbol represents the AND logic gate. In this case, the output is realised only after all the associated inputs have been received.

Symbol	Description
	This symbol represents the OR logic gate. In this case, any one or more of the inputs need to be received for the output to be realised.

47. FAULT TREE ANALYSIS CASE STUDY – PASSENGER ELEVATOR

In this section we discuss a case study on fault tree analysis of a passenger elevator (Main source, D Verma, 1993). Consider a passenger elevator depicted in Figure 11.3. We consider two major assemblies for FTA 1. Control assembly and 2. Drive/suspension assembly. All drive assembly failures are generalised as '*motor failures*' and '*other failures*' while control unit failures are generalised as '*hardware failures*' and '*software failures*' for the sake of simplicity.
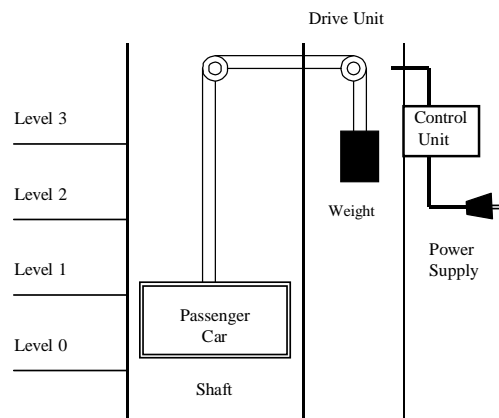
Figure 11.3 Schematic diagram of a passenger elevator

The control assembly consists of a microprocessor, which awaits an operator signal request to move the car to a certain level. The control unit activates drive unit that moves the car to that level and opens the elevator

door once the car comes to a stop. Switches exist at each level and inside the car allowing the controller to know where the car is at any time. Drive/suspension assembly holds the car suspended within the shaft and moves it to the correct level as indicated by the control unit. The Drive unit moves or stops the car only when prompted to do so by the control unit. The brake unit is designed to hod the car stationary when power is removed and to allow the motor shaft to turn when power is applied.

We define the top-level event in this case is '*passenger injury occurs*'. The following are the possible system operating conditions:

A.   Elevator operating properly.
B.   Car stops between levels.
C.   Car falls freely.
D.   Car entry door opens in the absence of car.
      In this case, operating conditions 'C' and 'D' are of concern. The initial fault tree is shown in Figure 11.4.
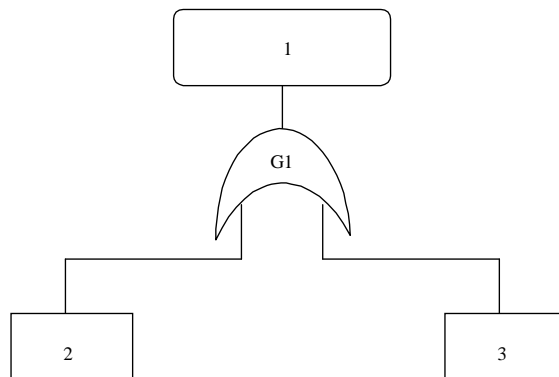
Figure 11.4 Initial fault Tree

In Figure 11.4, G1 is represents the OR logic gate and the events 1, 2 and 3 are as defined below:

Event 1 – Passenger injury occurs
Event 2 – Car free falls
Event 3 – Door opens without car present.

Thus, the top-level event (passenger injury occurs) can be either due to car free fall or door opens without the car present.

Now the event, car free fall, can further analysed by treating it as a top-level event, resulting in a fault tree depicted in Figure 11.5. In Figure 11.5, G2 is again a OR gate and the events 4, 5 and 6 are defined below:

Event 4 – Cable slips of pulley
Event 5 – Holding brake failure
Event 6 – Broken cable

Event 4 and 6 are *undeveloped event,* which can further broken, which can be further analysed using a separate fault tree. Event 5 is an *intermediate event.*
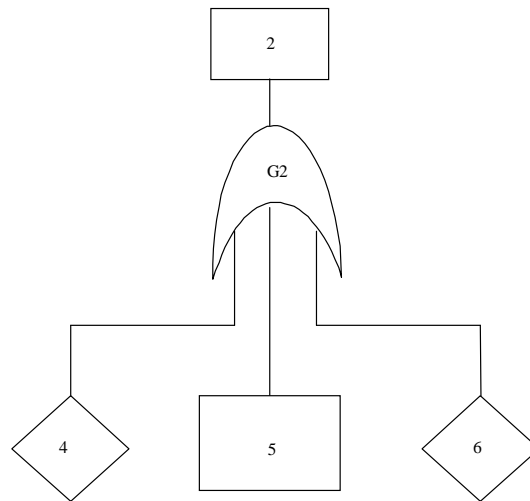
Figure 11.5 Further FTA analysis of the event car free fall
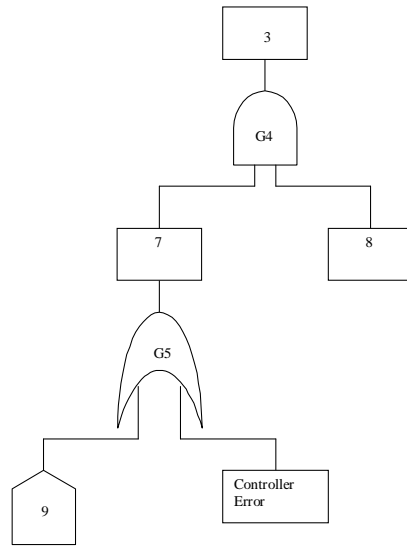
Figure 11.6 Fault tree for the event, the door opens erroneously

The event 3 can be further analysed to find the causes, Figure 11.6 depicts FTA for the event 3, door opens without the car present. This can be caused due to the following events:

Event 7 – Door close failure
Event 8 – Car not at level
Event 9 – Latch failure

For the event, *door opens erroneously*, to occur, events 7 and 8 must happen, thus we have a AND gate G3. The door close failure can be caused either due to the latch failure or due to controller error (denoted by OR gate, G4). Combining fault trees depicted in Figures 11.4-11.6, we can construct a complete (almost) for the event, passenger injury, as shown in Figure 11.7. Note that events 4, 5, 6, 8 and 9 can be further expanded to find the causes using fault tree analysis. The probability for the occurrence of the top-level event can be calculated once the time-to-failure and probability of occurrence of all the events are known. If the derived top-level probability is unacceptable, necessary redesign or compensation efforts should be identified and initiated.   As it is a simple mathematical calculation, it is not covered in this book.
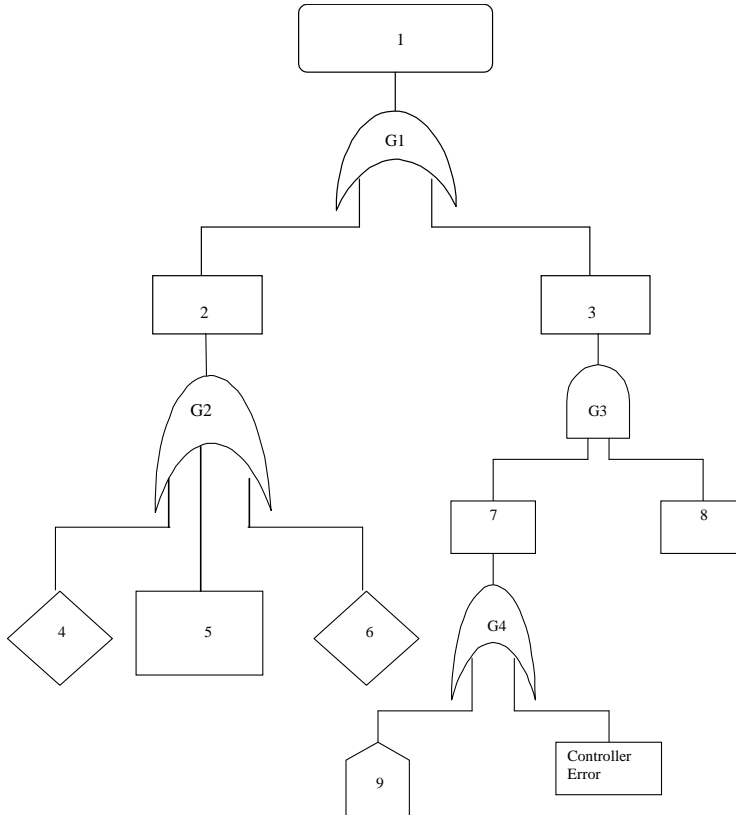
Figure 11.7 Fault tree for the event, passenger injury

Chapter 12

Analysis of Reliability, Maintenance and Supportability Data

Often statistics are used as a drunken man uses lamp posts... for support rather than illumination.

To predict various reliability characteristics of an item, as well as its maintainability and supportability function, it is essential that we have sufficient information on the time to failure, time to repair (maintain) and time to support characteristics of that item. In most cases these characteristics are expressed using theoretical probability distributions. Thus, the problem which every logistician face is the selection of the appropriate distribution function to describe the empirical data (obtained from data capturing sources) using theoretical probability distributions. Once the distribution is identified, then one can extract information about the type of the hazard function and other reliability characteristics such as mean time between failures and failure rate etc. In the case of maintenance and supportability data, we would identify the maintainability and supportability function as in the case of reliability data and then compute MTTR and MTTS.

To start with we look at ways of fitting probability distributions to in-service data, that is the data relating to the age of the components at the time they failed while they were in operation (in maintenance and logistic support we analyse the data corresponding to the maintenance and support task completion times). We look at three popular tools; 1. Probability papers, 2. Linear regression, and 3. Maximum likelihood estimates to identify the best distribution using which the data can be expressed and to estimate the corresponding parameters of the distribution. In the section on "censored data" we recognise that very often we do not have a complete set of failure data. We may wish to determine whether a new version of a component is more reliable than a previous version to decide whether we have cured the problem (of premature failures, say). Often, components will be replaced before they have actually failed, possibly because they have started to crack, they have been damaged or they are showing signs of excessive wear. We may have a number of systems undergoing testing to determine whether the product is likely to meet the various requirements but we need to go into production before they have all failed. There is useful data to be gleaned from the ones that have not failed as well as from the ones that have failed. If a component is being used in a number of different systems, it may be reasonable to assume that the failure mechanism in each of these instances will be similar. Even though the way the different systems operate may be different, it is still likely that the shape of the failure distribution will be same and that only the scale will be different.

Even relatively simple systems can fail in a number of different ways and for a number of different reasons. Suppose we wish to fasten two pieces of metal together using a nut and bolt. If we over-tighten the nut, we might strip the thread or we might shear the bolt. If we do not put the nut on squarely, we could cross the threads and hence weaken the joint. If the two pieces of metal are being forced apart then the stress on the nut and bolt may cause the thread to strip either inside the nut or on the outside of the bolt or it may cause the bolt to exceed its elastic and plastic limits until it eventually

breaks. If the joint is subject to excessive heat this could accelerate the process. Equally, if it is in very low temperatures then the bolt is likely to become more brittle and break under less stress than at normal temperatures. If the diameter of the bolt is towards the lower limit of its tolerance and the internal diameter of the nut is towards the upper limit then the amount of metal in contact may not be sufficient to take the strains imposed. As the two components age, corrosion may cause the amount of metal in contact to be even further reduced. It may also change the tensile strength of the metals and cause premature failure.

Components may therefore fail due to a number of failure modes. Each of these modes may be more or less related to the age. One would not expect corrosion to be the cause of failure during the early stages of the component's life, unless it was subjected to exceptionally corrosive chemicals. On the other hand, if the components have been badly made then one might expect to see them fail very soon after the unit has been assembled.

Very often, a possibly small, number of components may fail unexpectedly early.   On further investigation it may be found that they were all made at the same time, from the same ingot of metal or by a particular supplier.   Such a phenomenon is commonly referred to as a *batching* problem. Unfortunately, in practice, although it may be possible to recognise its presence, it may not always be possible to trace its origin or, more poignantly, the other members of the same batch or, indeed, how many there may be.

In deciding whether a new version of a component is more reliable than the old one, we need to determine how confident we are that the two distributions are different. If they both have the same (or nearly the same) shapes then it is a relatively straightforward task to determine if their scales are different. In some cases, the primary cause of failure of the origin version may have been eliminated or, at least, significantly improved but, another, hitherto rarely seen cause, may have become elevated in significance. This new primary cause may have a distinctly different shape than the first one that often makes it very difficult to decide between the two.

In this chapter, we first look at the empirical approaches for finding estimates for MTTF, MTTR and MTTS as well as failure function, maintainability and supportability functions. Rest of the chapter describes some of the well-known methods for selection of the most relevant theoretical distribution functions for the random variables under consideration.

12.48.   RELIABILITY, MAINTENANCE AND SUPPORTABILITY DATA

A very common problem in reliability engineering is the availability of failure data. In many cases getting sufficient data for extracting reliable information is the most difficult task. This may be due the fact that there is no good procedure employed by the operator (or supplier) to collect the data or the item may be highly reliable and the failure is very rare. However, even without any data, one should be able to predict the time-to-failure distribution if not the parameters. For example, if the failure mechanism is corrosion, then it cannot be an exponential distribution. Similarly if the failure cause is 'foreign object damage' then the only distribution that can be used is exponential. The main problem with insufficient failure data is getting an accurate estimate for the shape parameter. Fortunately, we don't have such problems with maintenance and supportability data. These are easily available from the people who maintain and support the item. The reliability data can be obtained from the following sources:

1.   Field data and the in-service data from the operator using standard data capturing techniques. There are standard *failure reporting* forms for the purpose of capturing desired information regarding the reliability of the item under consideration. Unfortunately, all these forms are

flawed, as they record only MTBF (or MTTR and MTTS in case of maintenance and support). Just the value of MTBF alone may not be enough for many analyses concerning reliability (similarly, in the case of maintenance (support), information on MTTR (MTTS) is not enough for complete analyses).

2. From *life testing* that involves testing a representative sample of the item under controlled conditions in a laboratory to record the required data. Sometimes, this might involve '*accelerated life testing*' (ALT) and '*highly accelerated life testing*' (HALT) depending on the information required.

As mentioned earlier, in some cases it is not possible to get a complete failure data from a sample. This is because some of the items may not fail during the life testing (also in the in-service data). These types of data are called '*censored data*'. If the life testing experiment is stopped before all the items have failed, in which cases only the lower bound is known for the items that have not failed. Such type of data is known as '*right censored data'.* In few cases only the upper bound of the failure time may be known, such type of data is called '*left censored data'.*

12.49. ESTIMATION OF PARAMETERS - EMPIRICAL APPROACH

The objective of empirical method is to estimate failure function, reliability function, hazard function, MTTF (or MTTR and MTTS) from the failure times (or repair and support times). Empirical approach is often referred as n*on-parametric approach* or *distribution free approach.* In the following sections we discuss methods for estimating various performance measures used in reliability, maintenance and support from different types of data.

Estimation of Performance Measures - Complete Ungrouped Data

Complete ungrouped data refers to a raw data (failure, repair or support) without any censored data. That is, the failure times of the whole sample under consideration are available. For example, let t_1, t_2, ..., t_n, represents n ordered failure times such that $t_i \leq t_{i+1}$. Then the possible estimate for failure function (cumulative failure distribution at time t_i) is given by:

$$\hat{F}(t) = \frac{i}{n} \qquad\qquad (12.1)$$

A total of i units fail by time t out of the total n in the sample. This will make $F(t_n) = n/n = 1$. That is, there is a zero probability for any item to survive beyond time t_n. This is very unlikely, as the times are drawn from a sample and it is extremely unlikely that any sample would include the longest survival time. Thus the equation (12.1) underestimates the component survival function. A number of mathematicians have tried to find a suitable alternative method of estimating the cumulative failure probability. These range from using n+1 in the denominator to using -0.5 in the numerator and +0.5 in the denominator. The one that gives the best approximation is based on *median rank*. Bernard's approximation to the median rank approach for cumulative failure probability is given by

$$\hat{F}(t_i) = \frac{i - 0.3}{n + 0.4} \qquad\qquad (12.2)$$

Throughout this chapter we use the above approximation to estimate the cumulative failure distribution or failure function. From equation (12.2), the estimate for reliability function can be obtained as

$$\hat{R}(t_i) = 1 - \hat{F}(t_i) = 1 - \frac{i - 0.3}{n + 0.4} = \frac{n - i + 0.7}{n + 0.4} \qquad (12.3)$$

The estimate for the failure density function $f(t)$ can be obtained using

$$\hat{f}(t) = \frac{\hat{F}(t_i) - \hat{F}(t_{i+1})}{t_i - t_{i+1}}, \qquad t_i \le t \le t_{i+1} \qquad (12.4)$$

Estimate for the hazard function can be obtained by using the relation between the reliability function $R(t)$ and the failure density function $f(t)$. Therefore,

$$\hat{h}(t) = \hat{f}(t) \Big/ \hat{R}(t) \quad \text{for} \quad t_i < t < t_{i+1} \qquad (12.5)$$

An estimate for the mean time to failure (or mean time to repair or mean time to support) can be directly obtained from the sample mean. That is,

$$\hat{MTTF} = \sum_{i=1}^{n} \frac{t_i}{n} \qquad (12.6)$$

Estimate for the variance of the failure distribution can be obtained from the sample variance, that is

$$s^2 = \sum_{i=1}^{n} \frac{(t_i - \hat{MTTF})^2}{n - 1} \qquad (12.7)$$

Estimate for MTTR (MTTS) and Variance of time to repair distribution (time to support distribution) can be obtained by replacing failure times by repair times (support times) in equation (12.6) and (12.7) respectively.

Confidence Interval

It is always of the interest to know the range in which the measures such as MTTF, MTTR and MTTS might lie with certain confidence. The resulting interval is called a *confidence interval* and the probability that it contains the estimated parameter is called its *confidence level* or *confidence coefficient*. For example, if a confidence interval has a confidence coefficient equal to 0.95, we call it a 95% confidence interval.

To derive a (1-α) 100% confidence interval for a *large sample* we use the following expression:

$$\hat{MTTF} \pm z_{\alpha/2} \left(\frac{\sigma}{\sqrt{n}} \right) \qquad (12.8)$$

Where $z_{\alpha/2}$ is the z value (standard normal statistic) that locates an area of $\alpha/2$ to its right and can be found from the normal table. σ is the standard deviation of the population from which the population was selected and n is the sample size. The above formula is valid whenever the sample size *n* is greater than or equal to 30. The 90%, 95% and 99% confidence interval for MTTF with sample size n $\geq$ 30 are given below:

$$90\% \text{ confidence } \hat{MTTF} \pm 1.645 \times \left(\frac{\sigma}{\sqrt{n}} \right) \tag{12.9}$$

$$95\% \text{ confidence } \hat{MTTF} \pm 1.96 \times \left(\frac{\sigma}{\sqrt{n}} \right) \tag{12.10}$$

$$99\% \text{ confidence } \hat{MTTF} \pm 2..58 \times \left(\frac{\sigma}{\sqrt{n}} \right) \tag{12.11}$$

When the number of data is small (that is when n is less than 30), the confidence interval is based on t distribution. We use the following expression to calculate $(1-\alpha)100\%$ confidence interval.

$$\hat{MTTF} \pm t_{\alpha/2} \left(\frac{s}{\sqrt{n}} \right) \tag{12.12}$$

where $t_{\alpha/2}$ is based on (*n*-1) degrees of freedom and can be obtained from *t* distribution table (refer appendix).

Example 12.1

Time to failure data for 20 car gearboxes of the model *M2000* is listed in Table 12.1. Find:

1. Estimate of failure function and reliability function.
2. Plot failure function and the reliability function.
3. Estimate of MTTF and 95% confidence interval.

Table 12.1. Failure data of gearboxes in miles

1022	1617	2513	3265	8445
9007	10505	11490	13086	14162
14363	15456	16736	16936	18012
19030	19365	19596	19822	20079

SOLUTION:

The failure function and reliability function can be estimated using equations 12.2 and 12.3. Table 12.2 shows the estimated values of failure function and reliability function.

Table 12.2. Estimate for failure and reliability function.

Failure data	$\hat{F}(t_i)$	$\hat{R}(t_i)$
1022	0.0343	0.9657
1617	0.0833	0.9167
2513	0.1324	0.8676
3265	0.1814	0.8186
8445	0.2304	0.7696
9007	0.2794	0.7206
10505	0.3284	0.6716
11490	0.3774	0.6225
13086	0.4264	0.5736
14162	0.4754	0.5246
14363	0.5245	0.4755
15456	0.5735	0.4265
16736	0.6225	0.3775
16936	0.6716	0.3284
18012	0.7206	0.2794
19030	0.7696	0.2304
19365	0.8186	0.1814
19596	0.8676	0.1324
19822	0.9167	0.0833
20079	0.9657	0.0343

The failure function and the reliability function graph are shown in Figure 12.1 and 12.2 respectively.
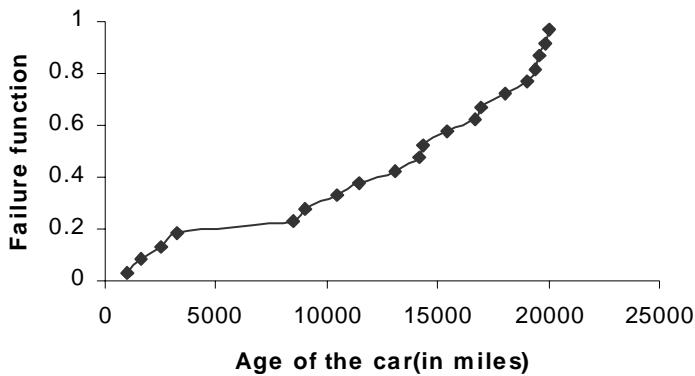
Figure 12.1 Estimate of failure function for the data shown in Table 12.1
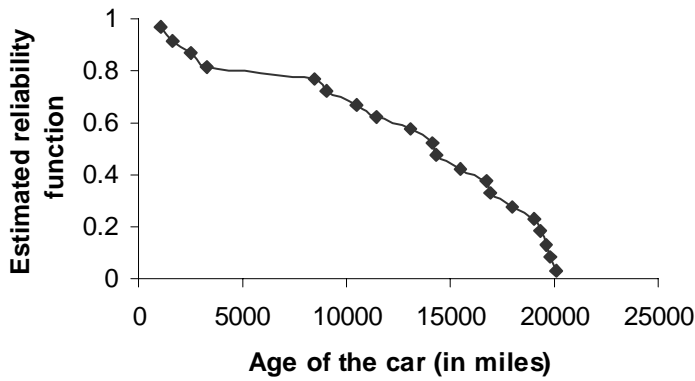
Figure 12.2 Estimated reliability function for the data given in Table 12.2

The estimate for mean time to failure is given by:

$$MTT\hat{F} = \sum_{i=1}^{20} \frac{t_i}{20} = 12725.5 \text{ miles.}$$

Estimate for the standard deviation is given by

$$s = \sqrt{\sum_{i=1}^{n} \frac{(t_i - MTT\hat{F})^2}{n-1}} = 14827.16 \text{ miles}$$

As the sample data is less than 30, we use equation (12.12) to find the 95% confidence level. From t-table the value of $t_{0.025}$ for (n-1) = 19 is given by 2.093. The 95% confidence level for MTTF is given by:

$$MTT\hat{F} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) = 12725.5 \pm 2.093(14827.16/\sqrt{19})$$

That is, the 95% confidence interval for MTTF is (5605.98, 19845.01).

Example 12.2

Time taken to complete repair tasks for an item is given in Table 12.3. Find the cumulative time to repair distribution and mean time to repair. Find 95% confidence level for MTTR.

Table 12.3. Time to repair data

28	53	71	90
30	56	72	92
31	58	74	94
33	59	75	95
35	61	79	97

40	65	81	99
41	67	82	100
44	68	84	103
49	69	85	108
51	70	89	110

Maintainability function can be estimated using following expression:

$$\hat{M}(t_i) = \frac{i - 0.3}{n + 0.4} = \frac{i - 0.3}{40.4}$$

Figure 12.3 shows the estimated maintainability function.
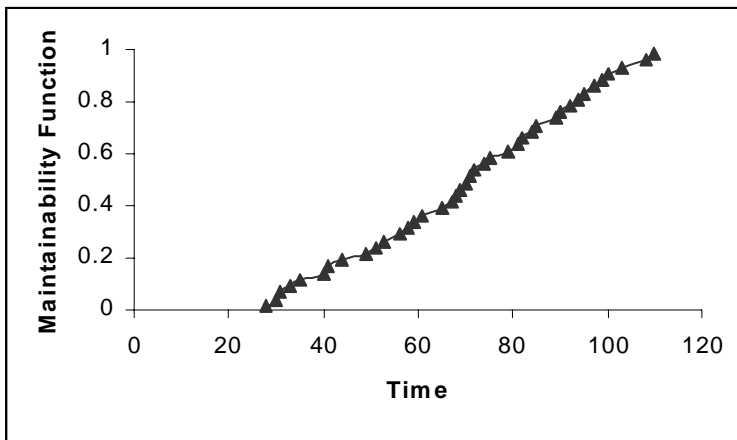
Figure 12.3.Maintainability function for the data given in Table 12.3.

Mean Time to Repair is given by:

$$\hat{MTTR} = \sum_{i=1}^{40} \frac{t_i}{40} = 69.7 \text{ hours}$$

Standard deviation for repair time is given by

$$s^2 = \sqrt{\sum_{i=1}^{n} \frac{(t_i - \hat{MTTR})^2}{n - 1}} = 23.43 \text{ hours}$$

Since $n > 30$, we use equation (12.10) to calculate 95% confidence interval for MTTR. 95% confidence level for MTTR is given by

$$\hat{MTTR} \pm 1.96 \left(\frac{s}{\sqrt{n}} \right) = 69.7 \pm (1.96) \left(\frac{23.43}{\sqrt{40}} \right) = (62.43, \ 76.96)$$

Analysis of Grouped Data

Often failure data is placed into time intervals when the sample size is large. The failure data are classified into several intervals. The number of intervals, *NI*, depends on the total number of data *n*. Following equation can be used as guidance for determining the suitable number of intervals:

$$\lfloor NI \rfloor = 1 + 3.3 \times \log_{10}(n) \qquad (12.13)$$

$\lfloor NI \rfloor$ denotes that the value is rounded down to the nearest integer.

The length of each interval, *LI*, is calculated using:

$$LI = \frac{(x_{max} - x_{min})}{\lfloor NI \rfloor} \qquad (12.14)$$

where x_{max} is the maximum recorded failure time and x_{min} is the minimum recorded failure time. The lower and upper bound of each interval is calculated as follows:

$$X_{min,i} = x_{min} + (i-1) \times LI$$
$$X_{max,i} = x_{min} + i \times LI$$

$X_{min,i}$ is the lower bound of the ith interval and $X_{max,i}$ is the upper bound value of the ith interval. Let n_1, n_2, ... n_h be the number of items that fail in the interval *i*. Then the estimate for cumulative failure distribution is given by

$$\hat{F}(X_{max,i}) = \frac{\sum_{k=1}^{i} n_k - 0.3}{n + 0.4} \qquad (12.15)$$

Estimate for the reliability function R(t) is given by:

$$\hat{R}(X_{max,i}) = 1 - \hat{F}(X_{max,i}) = \frac{\sum_{k=i+1}^{n} n_i + 0.7}{n + 0.4} \qquad (12.16)$$

Estimate for the failure density is given by:

For $X_{max,i+1} < t < X_{max,i}$

$$\hat{f}(t) = \frac{\hat{F}(X_{max,i+1}) - F(X_{max,i})}{X_{max,i+1} - X_{max,i}} = \frac{n_{i+1}}{(n+0.4) \times (X_{max,i+1} - X_{max,i})}$$

The MTTF is estimated using the expression:

$$\hat{MTTF} = \sum_{i=1}^{NI} \frac{X_{med,i} \times n_i}{n} \qquad (12.17)$$

where $X_{med,i}$ is the midpoint in the ith interval and n_k is the number of observed failures in that interval. Estimate for sample variance is given by

$$s^2 = \sum_{i=1}^{NI} (X_{med,i} - \hat{MTTF})^2 \times \frac{n_i}{n} \qquad (12.18)$$

Example 12.3

Results of 55 observed values of the duration of support tasks in hours are given in Table 12.4. Calculate the Mean Time to Support (MTTS).

Table 12.4. Time to support data

3	56	9	24	56	66	67	87	89	99	4
26	76	79	89	45	45	78	88	89	90	92
99	2	3	37	39	39	77	93	21	24	29
32	44	46	5	46	46	99	47	77	79	89
31	78	34	67	86	86	75	33	55	22	44

SOLUTION:

First we need to find the number of groups using equation (12.13). The number of intervals is given by:

$$\lfloor NI \rfloor = 1 + 3.3 \times \log_{10}(55) = \lfloor 6.74 \rfloor = 6$$

The length (range) if each interval (group) is given by:

$$LI = \frac{x_{max} - x_{min}}{\lfloor NI \rfloor} = \frac{99 - 2}{6} = 16.17$$

Table 12.5 shows the various calculations associated in computing the mean time to support.

Table 12.5. Analysis of grouped data given in example 12.3

i	$LI\ (x_{min,I} - x_{max,i})$	n_i	$x_{med,i}$	$X_{med,i} \times n_i$
1	2 - 18.17	6	10.08	60.51
2	18.17 – 34.34	10	26.25	262.55
3	34.34 - 50.51	11	42.42	466.67
4	50.51 - 66.68	5	58.59	292.97
5	66.68 - 82.85	9	74.76	672.88
6	82.85 - 99	14	90.92	1272.95

MTTS is given by:

$$\hat{MTTS} = \sum_{i=1}^{NI} \frac{X_{med,i} \times n_i}{n} = \sum_{i=1}^{6} \frac{X_{med,i} \times n_i}{55} = 55.06$$

12.50.   ANALYSIS OF CENSORED DATA

In many cases, the complete data may not be available due to the reasons such as all the items may not have failed or the manufacturer may wish to get interim estimates of the reliability etc. The mechanism for censoring may be based on a fixed age, on a fixed number of failures or at some arbitrary point in time. In practice, provided the times at the time of failure or, at the time of suspension (censor) are known, the reason for terminating the test is not important. We will assume that the times of failure are known precisely. We will look at cases in which we do not know the exact time, only that the failure occurred sometime between the last inspection and the current age later. In this section we derive estimates for failure function, reliability function when the data is multiple censored. We denote t_i to represent a complete data and t_i^* to denote a censored time.

The only difference between the estimation of parameters in complete data and the censored data is the calculation of median ranks. Now we will need to adjust the ranks in order to take account of the components that have not failed. The rank adjustment is done in the following two steps:

1. Sort all the times (failures and suspensions) in ascending order and allocate a sequence number i starting with 1 for the first (lowest) time and ending with n (the sample size for the highest recorded time). Now we discard the suspended times as it is only the (adjusted rank) of the failures with which we are concerned.
2. For each failure calculate the adjusted rank as follows:

$$R_i = R_{i-1} + \frac{n+1-R_{i-1}}{n+2-S_i} \qquad\qquad (12.19)$$

where, R_i is the adjusted rank of the ith failure, R_{i-1} is the adjusted rank of the $(i-1)^{th}$ failure, that is the previous failure. R_0 is zero and S_i is the sequence number of the ith failure.

As a quick check, the adjusted rank of the ith failure will always be less than or equal to the sequence number and at least 1 greater than the previous adjusted rank. If there is no suspensions, the adjusted rank will be equal to the sequence number as before. These adjusted ranks are then substituted into the Benard's approximation formula to give the median rank and the estimate for cumulative probability is given by:

$$\hat{F}(t_i) = \frac{R_i - 0.3}{n + 0.4}$$

Example 12.4

The following data were observed during the data capturing exercise on 12 compressors that are being used by different operators. Estimate the reliability and failure function (* indicates that the data is a censored data)

2041, 2173, 2248*, 2271, 2567*, 2665*, 3008, 3091, 3404*, 3424, 3490*, 3716

SOLUTION:

We need to calculate the adjusted rank of the failure times using equation (12.19), once this is done, then the failure and reliability function can be estimated using equations (12.2) and (12.3) respectively. The estimated failure and reliability functions are shown in Table 12.6.

Table 12.6 Estimated failure and reliability function

S_i	t_i	j	$R_j = R_{j-1} + [(n+1- R_{j-1}) / (n+2 - S_i)]$	$F(t_i)$	$R(t_i)$
1	2041	1	1	0.0565	0.9435
2	2173	2	2	0.1370	0.8630
3	2248*				
4	2271	3	3.1	0.2258	0.7742
5	2567*				
6	2665*				
7	3008	4	4.51	0.3395	0.6605
8	3091	5	5.92	0.4532	0.5468
9	3404*				
10	3424	6	7.69	0.5960	0.4040
11	3490*				
12	3716	7	10.34	0.8097	0.1903

12.51.   FITTING PROBABILITY DISTRIBUTIONS GRAPHICALLY

The traditional approach for measuring reliability, maintenance and supportability characteristics is using a theoretical probability distribution. It should however, be borne in mind that failures do not occur in accordance with a given distribution. These are merely convenient tools that can allow us to make inferences and comparisons in not just an easier way but also with known levels of confidence. In this section we will look at a graphical method that can be used to not only to fit distributions to given data but also help us determine how good the fit is. To illustrate the graphical approach we use the following failure data observed on 50 tyres.

Table 12.7. Failure data for 50 tyres

1022	14363	20208	26530	31507
1617	15456	20516	28060	33326
2513	16736	20978	28240	33457
3265	16936	21497	28757	35356
8445	18012	24199	28852	35747
9007	19030	24582	29092	36250
10505	19365	25512	29236	36359
11490	19596	25743	29333	36743
13086	19822	26102	30620	36959
14162	20079	26163	30924	38958

To draw a graph we obviously need a set 'x' and 'y' co-ordinates. Sorting the times-to-failure in ascending order will give us the 'x' values so all we need is to associate a cumulative probability to each value. This is done using the median rank approach discussed earlier, that is 'y' axis values are given by the cumulative failure probabilities calculated using the equation (12.2). Now, we can plot the values [t_i, $F(t_i)$]. In Figure 12.4 we can see the result of this for the 50 tyre time-to-failure.
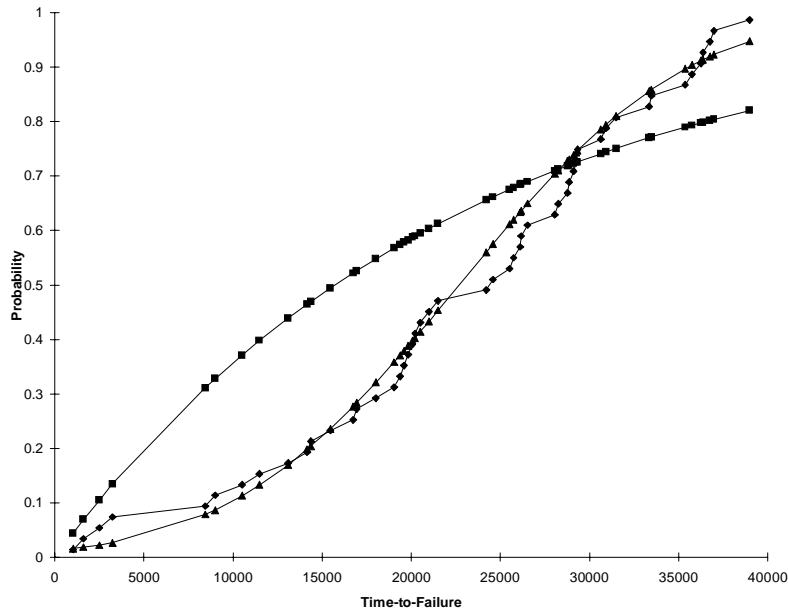
Figure 12.4 Tyre Data compared to Exponential and Normal Distributions

The two additional lines on this graph have been plotted to show what an exponential distribution (with the same mean as the sample) would look like and similarly for a normal distribution with the sample mean and standard deviation. This indicates that the exponential distribution is not a very good fit whereas the normal is certainly better. What it does not tell us, however, is how much better or, indeed, whether another distribution gives an even better fit.

A measure of how good the curve fits the data would be the *correlation coefficient* but, this only applies to straight line fits. Similarly we could use the *Kolmogorov-Smirnov* test but this really only tells us whether the there is a significant difference between the data and that which would be expected if the data were exponentially or normally distributed.

There are, in fact, two standard approaches to fit the data to a probability distribution graphically: to use "probability paper" or to transform either the "x" or "y" (or both) data so that the resulting graph would be a straight line if the data were from the given distribution. Actually both methods are essentially the same because to create probability paper the axes have been so constructed as to produce straight lines plot if the data is from the given distribution. If we can determine the necessary transforms then we can easily construct the probability paper.

Fitting an exponential distribution to data graphically

The cumulative probability density function for the exponential distribution is given by

224

$$F(t) = \begin{cases} 0, & t < 0 \\ 1 - \exp(-\lambda t), & t \geq 0 \end{cases}$$

Since we are only considering positive failure times, we can, without loss of generality, omit the expression for $t < 0$. If we replace $F(t)$ with p then we get

$$p = 1 - \exp(-\lambda t)$$

Rearranging and taking natural logarithm we get

$$\ln[\frac{1}{1-p}] = \lambda t \qquad\qquad (12.20)$$

This is a linear function in t such that the slope of the line is the reciprocal of the MTTF. Figure 12.5 is an example of "Exponential Graph Paper" (for the failure date from Table 12.7). The y-scale is given as percentages rather than probabilities. The x-scale is linear.
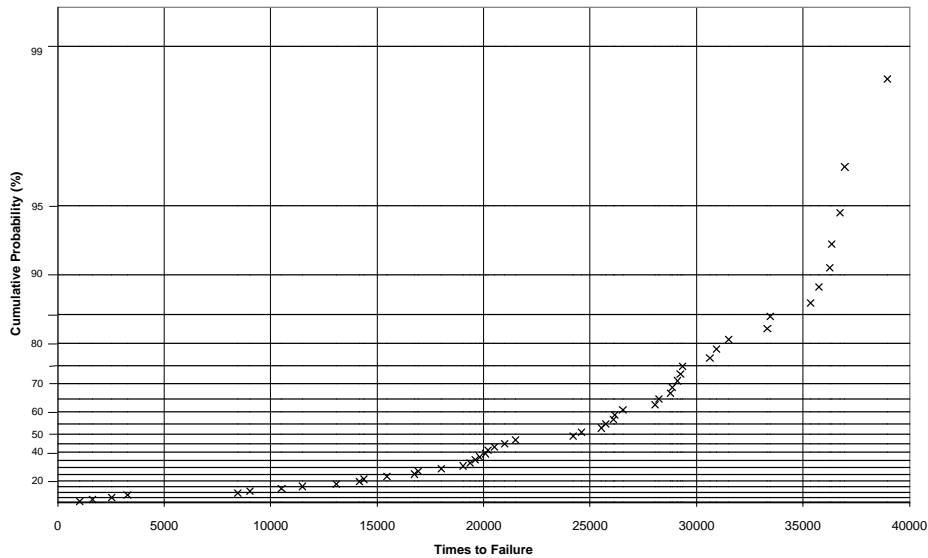
Figure 12.5 Data Plotted on Exponential Graph Paper

If the data forms a straight line in the exponential probability paper, then we can find the value of MTTF by using the relation F(MTTF) = 0.632. That is, we find the time to failure from the paper for which the percentage failures is 63.2.

Fitting a Normal Distribution Graphically

We will now see how good a fit the normal distribution gives. Again we can plot the times-to-failure on special normal (probability) paper. Such paper is becoming increasingly more difficult to obtain commercially. It can, however, be created using a proprietary spreadsheet package. Figure 12.6 shows how the tyre example failure times (and their respective median ranks) would appear on "normal paper".
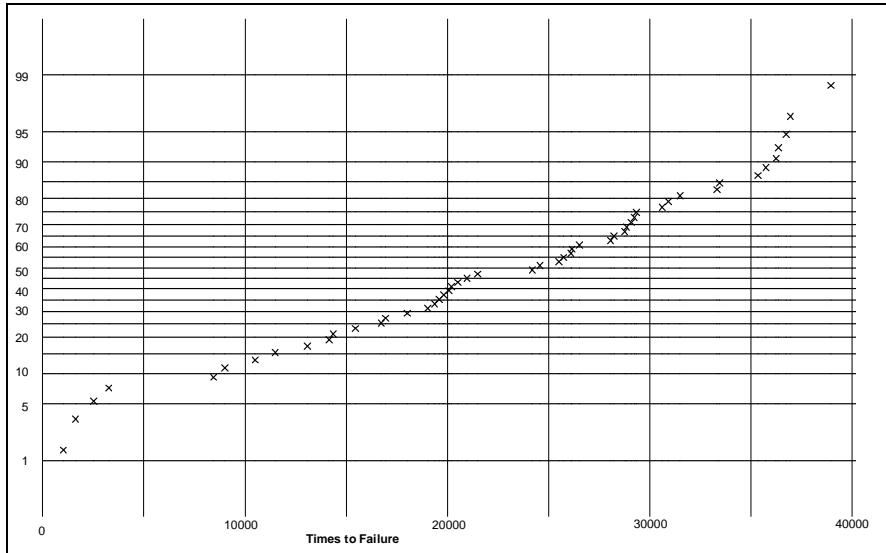
Figure 12.6 Times-to-Failure plotted on Normal Paper

The cumulative density function for the normal distribution is not as simple to transform to a linear form as the exponential.

$$F(t) = p = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

However, we can obtain the standardised normal variable $z = (\frac{x-\mu}{\sigma})$, for any given value of p ($F(t)$) either from tables or, using the *NORMSINV* function in MicroSoft™ Excel®, for example. Now we can plot this value(as the y co-ordinate) against the corresponding time-to-failure (as the x co-ordinate). The value of μ and σ can be found by using the relation, $F(\mu) = 0.5$ and $F(\mu+\sigma) = 0.84$.

Fitting a Log-Normal Distribution Graphically

Essentially the log-normal distribution is the same as a normal distribution excepting that the (natural) logarithm of the x-values are used in place of the actual values. Figures 12.7 and 12.8 show log-normal plot for the data given in Table 12.7.

226
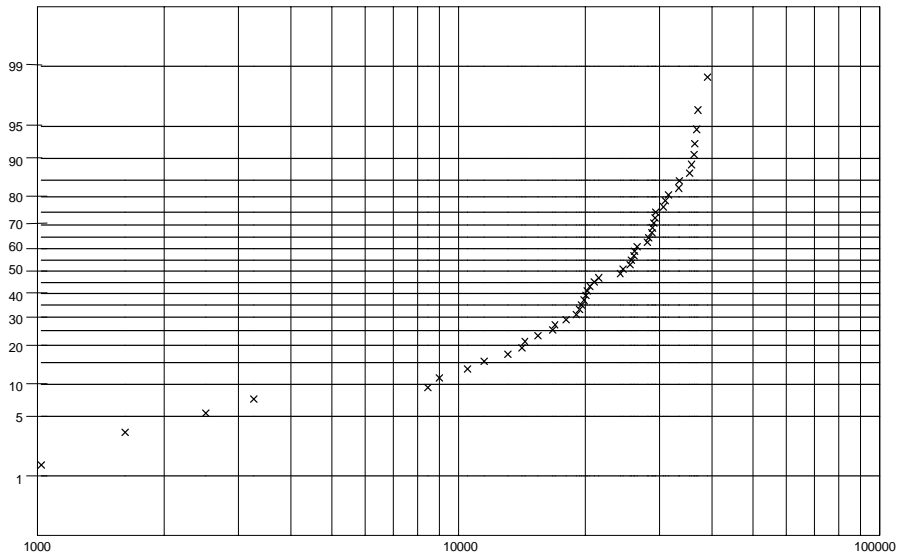
Figure 12.7 Times-to-Failure plotted on Log-Normal Paper
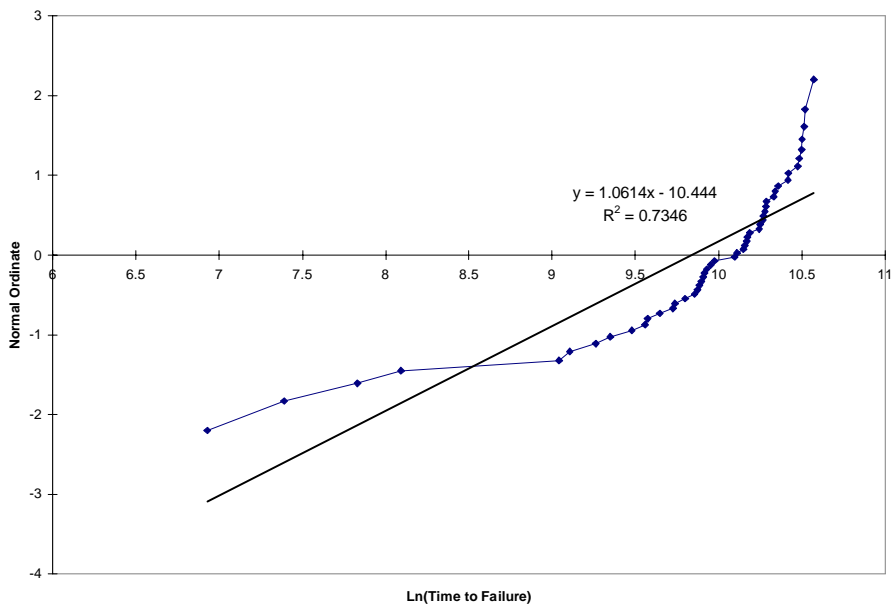
Figure 12.8 Fitting a Log-Normal Distribution Graphically

227

Here the plotted points form a concave curve to which the straight line is not a particularly good fit although it is still better than the exponential fit. The mean in this case is 18,776 which is considerably lower than the mean from the previous graphs but, this is because it is the geometric mean (the nth root of the product of the TTFs) and not the arithmetic mean with which we are more familiar.

Fitting a Weibull Distribution Graphically

The cumulative density function of the Weibull distribution is similar to that of the exponential, indeed the latter is the (mathematically) degenerative form of the former.

$$F(t) = p = \begin{cases} 0 & \text{for } t < 0 \\ 1 - e^{-\left(t/\eta\right)^{\beta}} & \text{for } t \geq 0 \end{cases}$$

By re-arranging and taking natural logarithms

$$-\ln(1-p) = \left(\frac{t}{\eta}\right)^{\beta}$$

which is still not in a linear form so we have to take logs again to give:

$$\ln(-\ln(1-p)) = \beta \ln(t) - \beta \ln(\eta)$$

So if we plot *ln(-ln(1-p))* against *ln(t)* an estimate of the shape parameter (β) of the Weibull will be given by the slope of the straight line drawn through the plotted points. To get an estimate of the scale parameter (η) we need to carryout a transform on the intercept:

$$\eta = e^{-c/\beta}$$

where *c* is the intercept of the regression line with the x-axis. Figures 12.9 and 12.10 shows Weibull plot for the data given in Table 12.7.
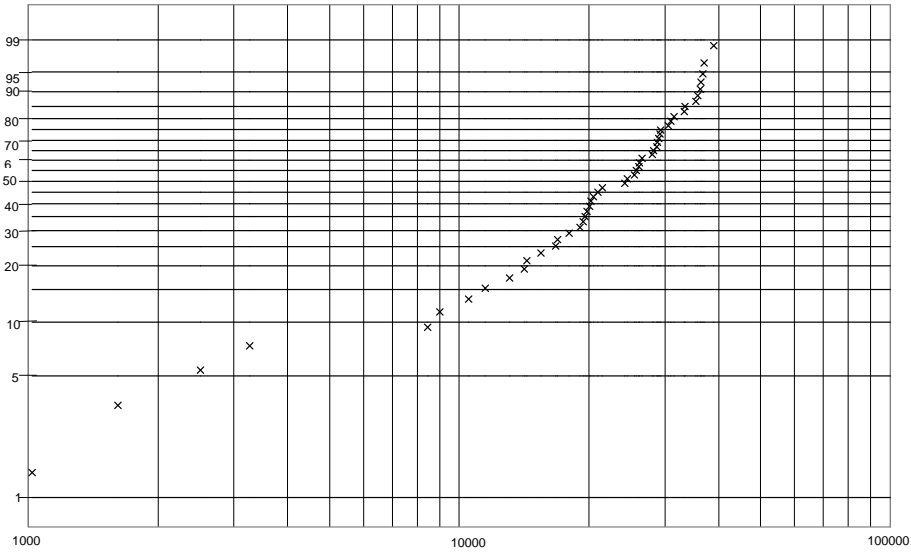
Figure 12.9 Times-to-Failure Fitted on Weibull Paper
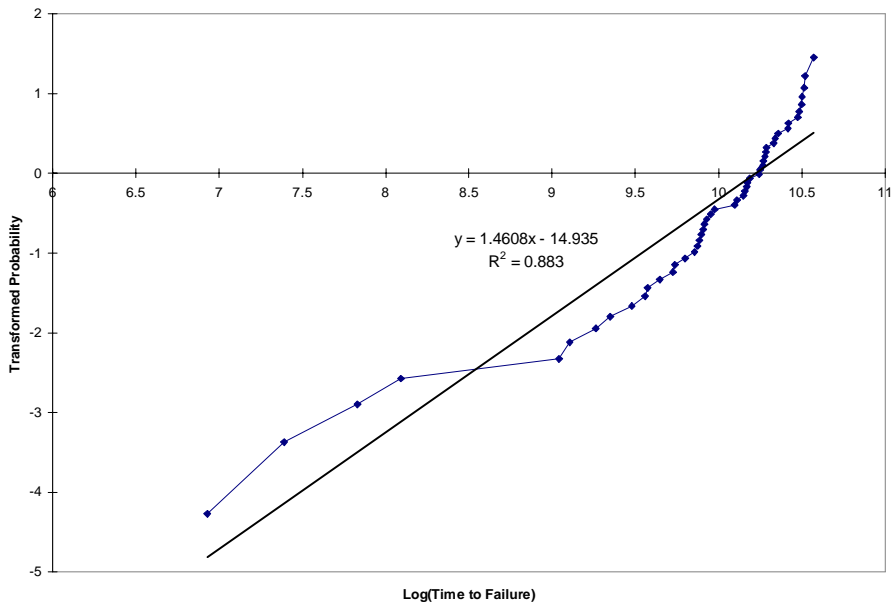
Figure 12.10 Fitting a Weibull distribution graphically

Again the Weibull distribution does not give as good a fit as the normal (distribution) but it is better than either the exponential or the log-normal. The slope (1.48) indicates that there could be a certain amount of age-relatedness to the failures.

12.52.   REGRESSION

The models used to relate a dependent variable *y* to the independent variables x are called regression models. The simplest regression model is the one that relates the variable *y* to a single independent variable *x* (*linear regression model*). Linear regression provides predicted values for the dependent variables (y) as a linear function of independent variable (*x*). That is, linear regression finds the best-fit straight line for the set of points (x, y). The objectives of linear regression are:

1. To check whether there is a linear relationship between the dependent variable and the independent variable.
2. To find the best fit straight line for a given set of data points.
3. To estimate the constants 'a' and 'b' of the best fit y = a + bx.
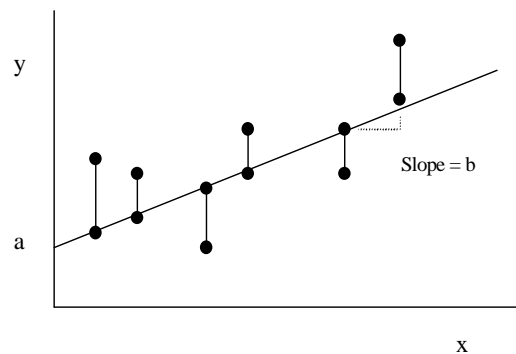
Figure 12.11 Least square regression.

The standard method for linear regression analysis (fitting a straight line to a single independent variable) is using the *method of least squares*. Least square regression is a procedure for estimating the coefficients 'a' and 'b' from a set of X, Y points that have been measured. In reliability analysis, the set X is the set of time to failures (or function of TTF) and set Y is their corresponding cumulative probability values (or function of cumulative distribution). Figure 12.11 illustrates the least square regression. The measure of how well this line fits the data is given by the correlation coefficient. If we construct a line such that it passes through the point $(\bar{x}, \bar{y})$ where $\bar{x}$ is the mean of the x values and $\bar{y}$ is the mean of the y values then the sum of the distances between each point and the point on the line vertically above (-ve) or below (+ve) will always be zero (provided the line is not parallel to the y-axis). The same holds for the horizontal distances provided that the line is not parallel to the x-axis. This means that any line passing through the means (in the way described) will be an unbiased estimator of the true line.
If we now assume that there is a linear relationship between the x's (x ∈ X) and y's (y ∈ Y), that the x's are known exactly and that the "errors" in the y values are normally distributed with mean 0 then it can be shown that the values of a and b which minimises the expression:

$$\sum_{i=1}^{n}(y_i - a - bx_i)^2 \qquad\qquad (12.21)$$

230

Will give the best fit. The expression ($y_i - a - bx_i$) gives the vertical distance between the point and the line. Cutting out lot of algebra, one can show that the values of a and b can be found by solving the following equations:

$$na + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \qquad\qquad (12.21)$$

$$a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \qquad\qquad (12.22)$$

'a' is the estimate of the intercept (of the line with the y-axis) and 'b' is the estimate of the slope – i.e. y = a + bx is the equation of the line giving:

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \qquad\qquad (12.23)$$

$$a = \sum_{i=1}^{n} \frac{y_i}{n} - b\sum_{i=1}^{n} \frac{x_i}{n} \qquad\qquad (12.24)$$

Note also that these expressions are not symmetrical in x and y. The formula quoted here gives

what is called *"y on x" regression* and it assumes the errors are in the y-values.

By replacing each x with a y and each y with an x we can perform "x on y" regression (which assumes

the errors are in the x-values). If *c* is the estimate of the intercept so obtained and d is the estimate

of the slope then to get estimates of *a* and *b* (the intercept and slope of the original graph):

$$b = \frac{1}{d} \text{ and } a = -\frac{c}{d}$$

Note: unless the points are collinear, the "x on y" estimates will not be the same as the "y on x" estimates. In the special case where you want to force the line through the origin (i.e. the intercept is zero), the least squares formula for the slope becomes:

$$b = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} \qquad\qquad (12.25)$$

Note this line does not pass through the means (unless it is a perfect fit).

Correlation Co-efficient

A measure of the dependence between two variables is given by the correlation coefficient. The correlation coefficient, r is given by:

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \times \sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}} \qquad (12.26)$$

The correlation coefficient always lies between –1 and +1. A value of +1 or –1 means that x and y are exactly linearly related. In the former case y increases as x increases but for r = -1, y decreases as x increases. Note that if x and y are independent then r = 0, but r = 0 does not mean that x and y are independent. The best fit distribution is the one with maximum r value (close to one). To find the best fit, regression analysis is carried out on the popular distribution such as exponential, Weibull, normal and log-normal. The one with highest correlation coefficient is selected as the best. The coordinates (x, y) and the corresponding parameters for different distributions are listed given in the following sections.

Linear Regression for Exponential Distribution

To fit a data to an exponential distribution, we transform the co-ordinates $(t_i, F(t_i))$ such a way that, when plotted, it gives a straight line. Here t_i is the observed failure times and $F(t_i)$ is the estimated cumulative distribution function. The cumulative distribution of exponential distribution is given by:

$$F(t) = 1 - \exp(-\lambda t)$$

that is,

$$\ln[\frac{1}{1 - F(t)}] = \lambda t \qquad (12.27)$$

Equation (12.27) is a linear function. Thus, for an exponential distribution, the plot of $(t, \ln[\frac{1}{1 - F(t)}])$ provides a straight line. Thus, if $t_1, t_2, \ldots, t_n$ are the observed failure times, then to fit this data into an exponential distribution, we set:

$$x_i = t_i \qquad (12.28)$$

$$y_i = \ln[\frac{1}{1 - F(t_i)}] \qquad (12.29)$$

Substituting (x_i, y_i) in equation (12.23) we get:

$$b = \frac{\sum\limits_{i=1}^{n} x_i y_i}{\sum\limits_{i=1}^{n} x_i^2} \qquad\qquad (12.30)$$

Note that, for exponential distribution b = 1/MTTF.

Example 12.5

The following failure data were observed on Actuators. Fit the data to an exponential distribution and find the MTTF and the correlation coefficient.

14, 27, 32, 34, 54, 57, 61, 66, 67, 102, 134, 152, 209, 230

SOLUTION:

First we carry out least square regression on $t_i, \ln[\frac{1}{1 - F(t_i)}]$, various calculations are tabulated in Table 12.8.

Table 12.8. Regression analysis for the data in example 12.5

i	$t_i (= x_i)$	$F(t_i)$	$y_i = \ln[1 / (1 - F(t_i))]$
1	14	0.0486	0.0498
2	27	0.1180	0.1256
3	32	0.1875	0.2076
4	34	0.2569	0.2969
5	54	0.3263	0.3951
6	57	0.3958	0.5039
7	61	0.4652	0.6260
8	66	0.5347	0.7651
9	67	0.6041	0.9267
10	102	0.6736	1.1196
11	134	0.7430	1.3588
12	152	0.8125	1.6739
13	209	0.8819	2.1366
14	230	0.9513	3.0239

The value of b is given by:

$$b = \frac{\sum\limits_{i=1}^{n} x_i y_i}{\sum\limits_{i=1}^{n} x_i^2} = \frac{\sum\limits_{i=1}^{n} t_i \times \ln[\frac{1}{1 - F(t_i)}]}{\sum\limits_{i=1}^{n} t_i^2} = 0.01126$$

MTTF is given by 1/b = 1/0.01126 = 88.73. The corresponding correlation coefficient is 0.9666.

Linear Regression for Weibull Distribution

Cumulative distribution of Weibull distribution is given by:

$$F(t) = 1 - \exp(-(\frac{t}{\eta})^\beta)$$

That is, $\ln[\ln(\frac{1}{1-F(t)})] = \beta \ln(t) - \beta \ln(\eta)$, which is a linear function. Thus to fit the data to a Weibull distribution, we set:

$$x_i = \ln(t_i) \hspace{4cm} (12.31)$$

$$y_i = \ln[\ln(\frac{1}{1-F(t_i)})] \hspace{3cm} (12.32)$$

From least square regression, it is evident that the shape and scale parameters of the distribution are given by:

$$\beta = b \hspace{4cm} (12.34)$$

$$\eta = \exp(-a/\beta) \hspace{4cm} (12.35)$$

Example 12.6

Construct a least square regression for the following failure data:

17, 21, 33, 37, 39, 42, 56, 98, 129, 132, 140

SOLUTION:
Making use of equations (12.31) and (12.32), we construct the least square regression, which are presented in Table 12.9.

Table 12.9. Weibull regression for the data in example 12.6

i	t_i	$F(t_i)$	$x_i = \ln(t_i)$	$Y_i = \ln\ln(1/1-F(t_i))$
1	17	0.0614	2.8332	- 2.7581
2	21	0.1491	3.0445	- 1.8233
3	33	0.2368	3.4965	- 1.3082
4	37	0.3245	3.6109	- 0.9354
5	39	0.4122	3.6635	- 0.6320
6	42	0.5	3.7376	- 0.3665
7	56	0.5877	4.0253	- 0.1209
8	98	0.6754	4.5849	0.1180
9	129	0.7631	4.8598	0.3648
10	132	0.8508	4.8828	0.6434
11	140	0.9385	4.9416	1.0261

Using equations (12.34) and (12.35), we get β = 1.4355, η = 76.54 and the correlation coefficient r = 0.9133.

Linear regression for Normal Distribution

For normal distribution,

$$F(t) = \Phi(\frac{t - \mu}{\sigma}) = \Phi(z)$$

Now z can be written as:

$$z_i = \Phi^{-1}[F(t)] = \frac{t_i - \mu}{\sigma} = \frac{t_i}{\sigma} - \frac{\mu}{\sigma} \qquad (12.36)$$

Which is a linear function. Now for regression, we set $x_i = t_i$ and $y_i = z_i = \Phi^{-1}[F(t_i)]$. The value of z can be obtained from standard normal distribution table. One can also use the following expression that gives polynomial approximation for z_i.

$$x_i = t_i \qquad (12.37)$$

$$P = \sqrt{\ln[\frac{1}{[1 - F(t_i)]^2}]}$$

$$y_i = P - \frac{C_0 + C_1 P + C_2 P^2}{1 + d_1 P + d_2 P^2 + d_3 P^3} \qquad (12.38)$$

where

$C_0 = 2.515517, C_1 = 0.802853, C_2 = 0.010328, d_1 = 1.432788,$
$d_2 = 0.189269, d_3 = 0.001308$

The estimate for μ and σ are given by

$$\mu = -\frac{a}{b} \text{ and } \sigma = \frac{1}{b}$$

Example 12.7

Fit the following data into a normal distribution

62, 75, 93, 112, 137, 170, 185

SOLUTION:

Table 12.10 gives various computations involved in regression.

Table 12.10. Normal regression for example 12.7

i	t_i	$F(t_i)$	$z_i = P - (c_0 + c_1 P + c_2 P^2 / 1 + d_1 P + d_2 P^2 + d_3 P^3)$
1	62	0.0945	- 1.2693
2	75	0.2297	- 0.7302
3	93	0.3648	- 0.3434
4	112	0.5	0
5	137	0.6351	0.3450
6	170	0.7702	0.7394
7	185	0.9054	1.3132

The estimate for μ = 118.71, σ = 54.05 and the correlation coefficient r = 0.9701.

Linear Regression for Log-normal Distribution

For log-normal distribution we set:

$$x_i = \ln(t_i) \tag{12.39}$$

$$P = \sqrt{\ln\left[\frac{1}{[1 - F(t_i)]^2}\right]}$$

$$y_i = P - \frac{C_0 + C_1 P + C_2 P^2}{1 + d_1 P + d_2 P^2 + d_3 P^3} \tag{12.40}$$

where
$C_0 = 2.515517$, $C_1 = 0.802853$, $C_2 = 0.010328$, $d_1 = 1.432788$, $d_2 = 0.189269$, $d_3 = 0.001308$

Total Cost of Ownership (TCO) and Life Cycle Cost Models

Prof. U Dinesh Kumar
Indian Institute of Management Bangalore

CONTENTS

1. Life Cycle Cost and Total Cost of Ownership

1.1 Introduction

"Value for Money" has become one of the important criteria in an increasingly competitive business environment. Life Cycle Cost (LCC) and the Total Cost of Ownership (TCO) are two important financial measures that are used for decision making in acquisitions. From its origins in defence equipment procurement in the US in early 1960s, the use of life cycle cost and cost of ownership has extended to other areas of the public and private sectors. LCC and TCO are being used to assist in decision-making, budget planning, cost control, and range of other activities that occur over the life of complex technological equipment.

It is important to consider the difference between LCC and TCO. LCC analysis is applied routinely to military projects. In the military sector the consumer, by funding the project and operating the related product, essentially bears the total life cycle cost covering the major cost elements in all stages of a product's life cycle. The term LCC analysis is rarely used in the commercial sector. Instead, the main focus is on TCO where related costs, covering acquisition (purchase or lease), operation, maintenance and support are borne by the customer. In addition, the customer can also incur costs when the product is not available for use, that is, '*down time costs*'.

The objectives of LCC/TCO are (Flanagan and Norman, 1983):

Ø   To enable investment options to be more effectively evaluated.

Ø   To consider the impact of all costs rather than only the initial capital costs.

Ø   To assist in the effective management of completed projects.

Ø   To facilitate choice between competing alternatives.

In the Defence industry the system's life cycle is divided into various phases, which allow proper planning and control of a project. The number of phases depend on the nature of the project, purpose and whether they are applied to commercial, military or space projects (Knotts, 1998). Commonly used phases are:

1. Requirements (Functional Specification).

2. Concept/Feasibility Studies.

3. Design and Development.

4. Production.

5. Testing and Certification.

6. Operation, Maintenance and Support.

7. Disposal

It is reported by the US Department of Defence that 70% of weapon system life cycle cost is committed by the end of concept studies, 85% by the end of system definition and 95% by the end of full scale development. The US Department of Defence has formally used the concept of life cycle cost in weapon system acquisition since the early 1960s through life cycle costing and life cycle cost analysis.

The cost of ownership approach identifies all future costs and reduces them to their present value by use of the discounting techniques through which the economic worth of a product or product options can be assessed. In order to achieve these objectives the following elements of cost of ownership have been identified (Woorward, 1997):

Ø   Initial capital costs

Ø   Life of the asset

Ø   The discount rate

  Ø Operating and maintenance costs

  Ø Disposal cost

  Ø Uncertainty and sensitivity analysis

1.2   Initial capital costs

The initial capital costs can be divided into three sub-categories of cost namely: (1) purchase costs, (2) acquisition/finance costs, and (3) installation/commissioning/training costs. Purchase costs will include assessment of items such as land, buildings, fees, and equipment. Finance costs include alternative sources of funds. Basically, the initial capital cost category includes all the costs of buying the physical asset and bringing it into operation.

1.3   Life of the Asset

The estimated life of an asset has a major influence on life cycle cost analysis. Ferry *et al* (1991) has defined the following five possible determinants of an asset's life expectancy:

Functional life – the period over which the need for the asset is anticipated.

Physical life – the period over which the asset may be expected to last physically, to when replacement or major rehabilitation is physically required.

Technological life – the period until technical obsolescence dictates replacement due to the development of a technologically superior alternative.

Economic life – the period until economic obsolescence dictates replacement with a lower cost alternative.

Social and legal life – the period until human desire or legal requirement dictates replacement.

1.4   The discount rate

As the cost of ownership is discounted to their present value, selection of a suitable discount rate is crucial for TCO analysis. A high discount rate will tend to favour options with low capital cost, short life and high recurring cost, whilst a low discount rate will have the opposite effect.

1.5   Operations and Maintenance Costs

Cost of ownership, in many cases, is about operation and maintenance cost.   Estimation of operation and maintenance costs is the essential to minimise the total cost of ownership of the asset. In the whole of TCO analysis, estimation of operation and maintenance is the most challenging task.

1.6   Disposal cost

This is the cost incurred at the end of as asset's working life in disposing of the asset. The disposal cost would include the cost of demolition, scrapping or selling the asset.

1.7   Uncertainties and Sensitivity Analysis

LCC/TCO is highly dependent on the assumptions and estimates made whist collecting data. Even though it is possible to improve the quality of these estimates, there is always an element of uncertainty associated with these estimates and assumptions.   Macedo *et al* (1978) identifies the following five major sources of uncertainty:

1. Differences between the actual and expected performance of the system could affect future operation and maintenance cost.

2. Changes in operational assumptions arising from modifications in user activity.

3. Future technological advances that could provide lower cost alternatives and hence shorten the economic life of any system/subsystem.

4. Changes in the price levels of major resources such as energy or manpower, relative to other resources can affect future alteration costs.

5. Error in estimating relationships, price rates for specific resources and the rate of inflation in overall costs from the time of estimation to the availability of the asset.

While undertaking a LCC/TCO analysis, there may be some key parameters about which uncertainty exists, usually because of the inadequacy of the input data. Blanchard (1972) suggested the following should be the subject of sensitivity analysis:

Ø   Frequency of the maintenance factor.

Ø   Variation of the asset's utilization or operating time.

Ø   Extent of the system's self-diagnostic capability.

Ø   Variation of corrective maintenance hours per operating hour.

Ø   Product demand rate.

Ø   The discount rate

1.8   Summary

In this chapter, we looked at the concept of Life Cycle Cost and Cost of Ownership and the factors that influence LCC and TCO as described in the literature. In the next chapter we survey the existing LCC/TCO models, methodologies, practices and techniques available in the literature and its applications and limitations. The models that have direct application for assessing the total cost of ownership of airborne military equipment are highlighted.

53. 2. SURVEY OF EXISTING LITERATURE

Although a considerable body of literature relating to life cycle cost and cost of ownership has been developed over the past four decades, much of the published material has emanated from practitioners (Nicholas, 1999). Publications by practitioners have tended to consist of general guidelines and a substantial amount of technical reports in the form of technical reports and conference papers, detailing the development and application of specific models and modeling techniques. Most of these papers lacked rigor that one can expect from academic publications. Much fewer in number, academic publications have taken the form of text books which present tools and techniques of analysis (Dhillon 1989, Fabrycky and Blanchard 1991) and Journal publications, which tend to consider very specific technical aspects of LCC and TCO. We have grouped the literature under different classification, namely, (1) Publications on LCC/TCO concepts, (2) Publications on LCC models, (3) Publications on TCO models, and (4) Publications on LCC/TCO applications.

2.1   Life Cycle Cost/Total Cost of Ownership Concept

Asiedu and Gu (1998) in their paper titled, 'Product life cycle cost analysis – state of the art review,' provides an in-depth analysis of several issues of the life cycle cost. The paper discusses issues such as (1) life cycle approach to design; (2) life cycle cost analysis, and (3) cost analysis models. They point out that LCC analysis should not be seen as an approach for determining the cost of the system but as an aid to design decision-making. The use of life cycle cost analysis and cost of ownership should

therefore be restricted to the cost that we can control. For designers, estimating the LCC of a proposed product during its development phase is required for a number of reasons including:

(1) Determining the most cost efficient design amongst a set of alternatives.

(2) Determining the cost of a design for budgetary purposes.

(3) Identifying cost drivers for design changes and optimisation.

Life cycle costing functions	Physical asset functions	Financial functions	Organizational functions

Specification

Costs

Time

Trade-off

Capacity
Utility
Performance

Reliability
Availability
Maintainability

Capital cost

Revenue Cost

Design
Development
Acquisition

Disposal

Operations
Maintenance
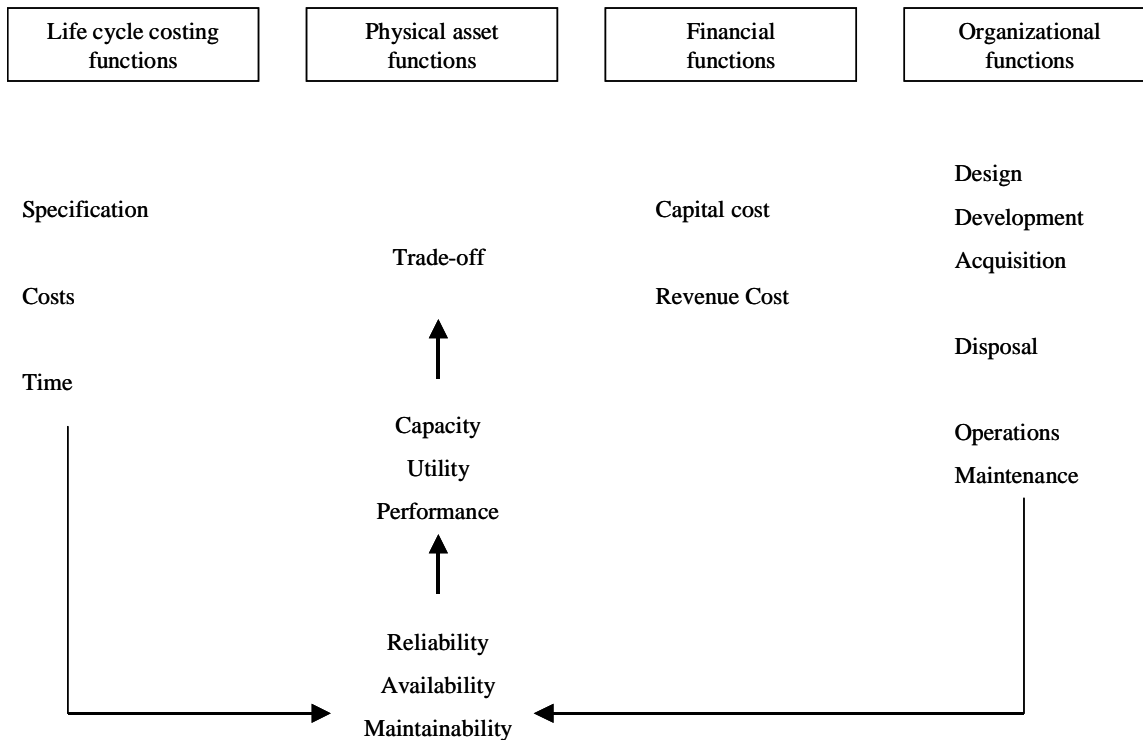
Figure 1. Key factors in Life Cycle Cost (Rose, 1984)

Rose (1984) in a short paper reviewing the status of life cycle cost argues that many forms of life cycle cost analysis are not actually life cycle cost analysis but can contribute to life cycle cost analysis. This is an important distinction because wide ranges of analyses are often termed as life cycle cost analysis, whereas they are partial analysis.

Rose concludes that many life cycle cost studies relate to part of a system rather than a complete system. He has shown the relationship between capital and revenue costs, the potential

trade-off between costs and engineering features, and the organizational functions of an enterprise. This is illustrated in Figure 1.

Hart (1985) in a paper titled, 'The interpretation of life cycle costs,' refers to a potential communication problem – first within the life cycle cost community and second with the receiver of the information for decision making. The suggested solution is the set of definitions at a high level in terms of boundaries, which consist of (1) system, (2) life and (3) cost. The system boundaries establish the extent of the life cycle cost analysis. Although this would appear straight forward, Hart observes that often analysis relates to the highly visible items of a system and disregards ancillary items needed to operate and maintain it. Hart describes the boundary of life in the following terms – 'The life of a project begins when there is recognition that a new asset is needed to meet the requirements of the organization. Resources are then expended by the owner, to test the suitability and feasibility of the proposed asset, define it, acquire it, integrate it into service and then operate it'. Hart's description of life doesn't include the disposal of the system, which is important phase of life cycle. Hart defines cost as those resources sacrificed towards an objective.

Hart also discusses the problem of communicating life cycle cost information to decision-makers. He concludes that there should be two studies – the first an economic study and the second a budgetary study. The economic study considers all the resources which were committed in the past to procure assets and from which benefits can still be derived in addition to future consumption of resources. The budgetary study concerns only future procurement and is not concerned with the past (sunk) costs.

At the concept level, definition of life cycle cost and total cost of ownership itself is a challenging task. Definitions of the term life cycle cost in the literature are normally of a generalized form. Several definitions of LCC/TCO exist. It is important for any organization to define what they mean by LCC or

TCO. This will set the boundary for the costs that should be included for the analysis. In the following sections, we look at the some of the most common definitions of LCC/TCO.

2.1 Definitions of Life Cycle Cost and Cost of Ownership

White and Ostwald (1976)

" The life cycle cost of an item is the sum of all funds expended in support of the item from its conception and fabrication through its operation to the end of its useful life "

Michaels and Woods (1989)

" The total cost to the customers of acquisition and ownership of that system over its full life "

Dhillon (1989)

" The sum of all costs incurred during the life time of an item, i.e., the total of procurement and ownership costs "

Fabrycky and Blanchard (1991)

" All costs associated with the system or product as applied to the defined life cycle "

" Life cycle costing is all costs associated with the system as applied to the defined life cycle. The total cost of a system could be broken into four categories, (1) design and development cost, (2) production/manufacturing cost, (3) utilization cost, and (4) retirement and disposal cost "

Degraeve and Roodhooft (1999)

"The total cost of ownership is the true cost of buying a particular good or service and consists of price and other elements that reflect additional costs caused by the suppliers in the purchasing companies value chain"

2.3 Life Cycle Cost Technique

Harvey (1976) in his article comprehensively reviewed the LCC technique and proposed the general procedure for LCC, which is summarized in the Figure 2.
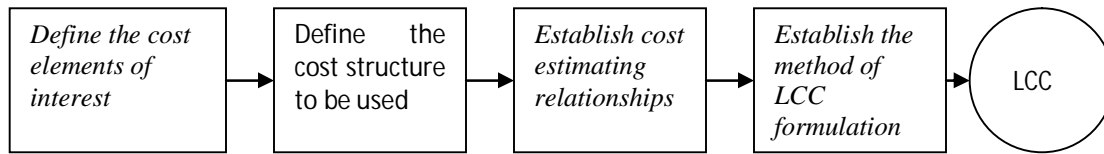
Figure 2 Harvey's life cycle costing procedure

Woodward (1997) has elaborated the different steps of Harvey's procedure as given below:

The cost elements of interest are all the cash flows that occur during the life of the asset. From the definition of LCC it is apparent that the LCC of an asset includes all expenditure incurred in respect of it, from acquisition until disposal at the end of its life.

Defining the cost structure involves grouping costs so as to identify potential trade-offs, thereby to achieve optimum LCC. The nature of the cost structure defined will depend on the required depth of the LCC study, and a number of alternative structures have been proposed in the literature (White and Ostwald 1976, Fabrycky and Blanchard 1991).

Cost estimating relationship is a mathematical expression that describes, for estimating purposes, the cost of an item or activity as a function of one or more independent variables.

Establishing the method of LCC formulation involves choosing an appropriate methodology to evaluate the asset's LCC.

Kaufman (1970) developed one of the earliest formulations of LCC where he has developed a model based on the eight-step approach indicated below and shown in Figure 3. The eight steps of Kaufman's LCC model are:

Ø   Establish the operating profile

Ø   Establish the utilization factors

249

Ø   Identify all cost elements

Ø   Determine all critical cost parameters

Ø   Calculate all costs at current prices.

Ø   Escalate current costs at assumed inflation rates;

Ø   Discount all costs to the base period:

Ø   Sum discounted costs to establish the net present value.

Step 1: The operating profile (OP) describes the periodic cycle, through which equipment will go, and indicates which equipment will, or alternatively will not be working. The operating profile should indicate the operating hours of the equipment throughout the life of that equipment.

Step 2: Utilization factors indicate in what way equipment will be functioning within each mode of the OP.

Step 3: Every cost element or area of cost must be identified.

Step 4: The critical cost parameters are those factors, which control the degree of costs incurred during the life of the equipment. Stevens (1976) has suggested the most significant of these are:

§   Mean Time between failures (MTBF)

§   Mean Time Between Overhauls (MTBO)

§   Mean Time To Repair (MTTR)

§   Time Between Scheduled Maintenance

§   Energy use rate

Step 5: All costs are first calculated at current rates.

Step 6: All costs need to be projected forward at appropriate rate (that is, differential) rates of inflation.

Step 7: Money has a time value and the cash flows occurring in different time periods should be discounted back to the base period to ensure comparability.

Step 8: Summing all the cash flows involved will enable the LCC of the asset to be established. Comparisons between competing assets can then be undertaken, and the fallacy of opting simply for the asset with lowest capital cost will then be exposed.
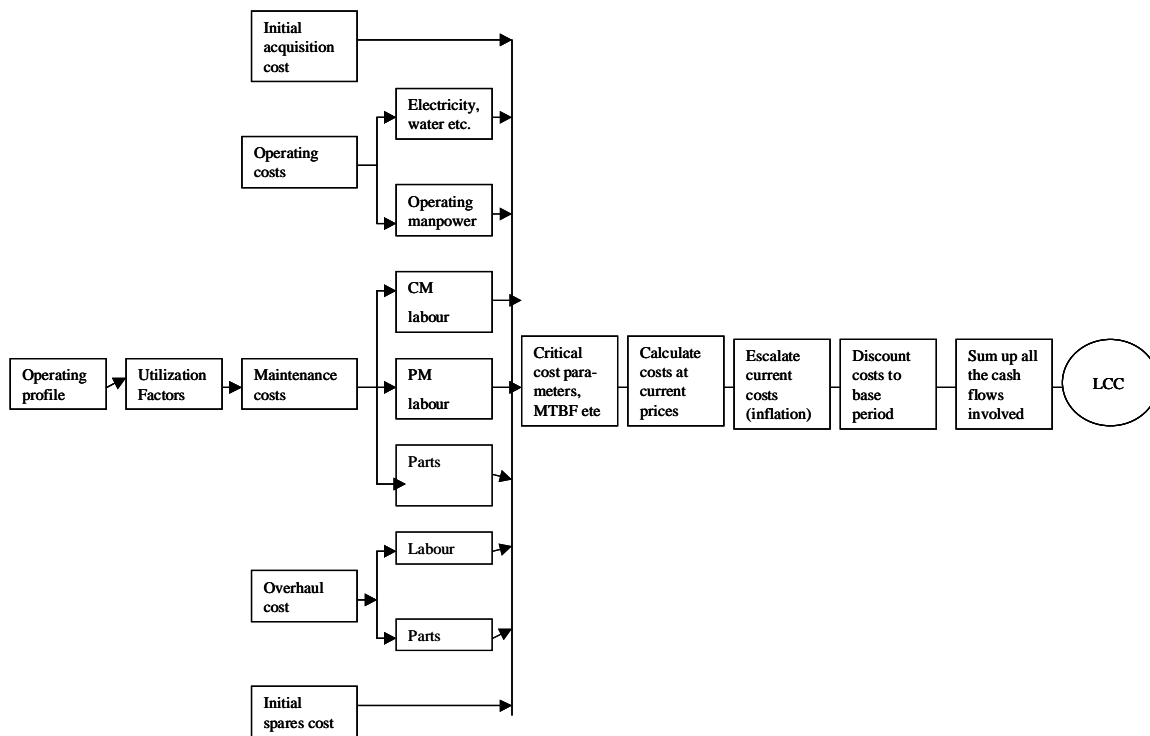
Figure 3. Kaufman's life cycle costing formulation

54. 2.4 REVIEW OF LIFE CYCLE COST MODELS

In general, The LCC models can be classified into the following categories:

1. Accounting models (models that sum LCC components).

2. Cost estimating relationship (CER) models (models used to analyse design alternatives).

3. Heuristic models.

251

4. Failure free warranty models (models used to analyse warranty periods)

5. Reliability models (used to apportion reliability and maintainability), and

6. Economic analysis models (models dealing with general cost effectiveness).

However, Sherif and Kolarik (1981) classify LCC models into three general forms: (1) conceptual, (2) analytical models, and (3) heuristic models. Conceptual models consist of a set of hypothetical relationships expressed in a qualitative framework. Conceptual models are generally constructed at macro level. Analytic models consist of a set of mathematical relationships, which are used to describe a certain aspect of the system. Such models range from models covering very specific aspects of a system to models, which address total system life cycle cost.

Gupta (1983) identifies three types of analytic models: (1) design trade-off models, (2) total cost models, and (3) logistic support models. Design trade-off models relate to the design phase of the life cycle cost and attempt to minimise cost to meet a given value of design parameters such as reliability and availability to maximize the value of design parameters for given cost constraints. Total cost models are termed true life cycle cost models and usually encompass the total life of the system. They attempt to minimize the total life cycle cost of the system while maximizing its performance and effectiveness by evaluating various parameters such as reliability, maintainability, availability etc, which affect life cycle cost. Logistic support models are concerned with the operations phase of the life cycle. Usually the objective of such models is to determine costs for alternative support plans and effect on the system's effectiveness. They reflect operations cost parameters as variable costs and research, development, test and evaluation and acquisitions costs as fixed costs. These models are inconsistent in that design parameters such as reliability and maintainability heavily influence operations costs and therefore fall short of determining optimal life cycle cost.

Dhillon (1989) simply divides life cycle cost models into two forms: (1) general life cycle cost models, and (2) specific life cycle cost models. General life cycle cost models are not related to any specific

equipment or system whereas specific life cycle cost models have been developed for particular types of equipment or system. Given the specific interrelationships and interactions of a particular system, the application of general models is clearly limited. Recently, Daniel (1991) has classified life cycle cost models into two broad categories: (1) accounting models which attempt to assemble and distribute costs, determined elsewhere so as to describe the total cost of a system, and (2) predictive models which are used to forecast the values of the various cost elements required as input to the accounting models.

Noble and Tanchoco (1990) developed a conceptual framework for concurrent design and economic justification of the system. A prototype implementation was developed to explore the usefulness of the design justification concept. Actual data from the design of an electromagnetic/radio frequency shield, a component in electrical metering equipment, was used to demonstrate the model.

Woodward (1997) in his paper titled, 'Life Cycle Costing – Theory, Information Acquisition and Application,' presented a case on total cost of ownership on South Yorkshire Passenger Transport (SYPT). SYPT's main activity is the provision of passenger transport services by road. It's fixed assets were worth $ 43,327,500 out of which the passenger vehicles accounted for about 17,662,500. The company purchases vehicles that form a major part of the capital expenditure, on a regular basis and the decision to purchase them is based on the LCC technique. The estimated life cycle costs are discounted at an assumed monetary cost of capital of 15%, after including a standard inflation rate assumed over the life of the asset. If the two alternatives have similar discounted costs, then a choice will be made by the financial director taking into account non-financial factors such as the credibility, reliability etc. of the suppliers. Although, the case was on passenger transport by road, the concept is valid for any system, including airborne defence equipment.

Degraeve and Roodhooft (1999) developed a mathematical programming model that uses total cost of ownership information to select suppliers and determine order quantities over a multi-period time

253

horizon. The total cost of ownership quantifies all costs associated with purchasing process and is based on the activities and cost drivers determined by an Activity Based Costing (ABC) system. They have also discussed a case on the purchasing problem of heating electrodes at *Cockerill Sambre*, a Belgian multinational steel producer. In this case, quality issues accounted for more than 70% of the total cost of ownership making the quality of the supplier a critical success factor in the supplier selection process.

LCC model can be a simple series of cost estimation relationships (CERs). LCC analysis during the conceptual or preliminary design phases may require the use of basic accounting techniques (Fabrycky and Blanchard, 1991). The most important task in LCC modeling is the construction of Cost Breakdown Structure (CBS), which shows various cost categories that combine to provide the total cost. Cost breakdown structure should exhibit the following basic characteristics (Blanchard *et al* 1995):

1. All system cost elements must be considered.

2. Cost categories are generally identified with a significant level of activity or some major item of hardware.

3. The cost structure and categories should be coded in such a manner as to allow for the analysis of certain specific areas of interest (e.g., system operation, energy consumption, equipment design, spares, maintenance personnel and support, maintenance equipment and facilities). In some instances, the analyst may wish to pursue a designated area in depth while covering other areas with gross top-level estimates. This will certainly occur from time to time as a system evolves through the different phases of its life cycle.

4. When related to a specific program, the cost structure should be compatible (through cross-indexing, coding etc.) with the contract work breakdown structure (WBS) and with management accounting procedures used in collecting costs.

5. For program, where subcontracting is prevalent, it is often desirable and necessary to separate supplier costs (i.e., initial bid price and follow-on program costs) from other costs. The cost

structure should allow for the identification of specific work packages that require close monitoring and control.

An example of a cost breakdown structure adopted from Blanchard (1991) is shown in Figure 4. Referring to Figure 4, costs may be accumulated at different levels depending on the areas of interest and the depth of detail required. Most of the LCC models can be a simple series of cost estimation relationships. Estimating models used in industry can be broadly classified as parametric models, analogous models and detailed models (Asiedu and Gu, 1998).
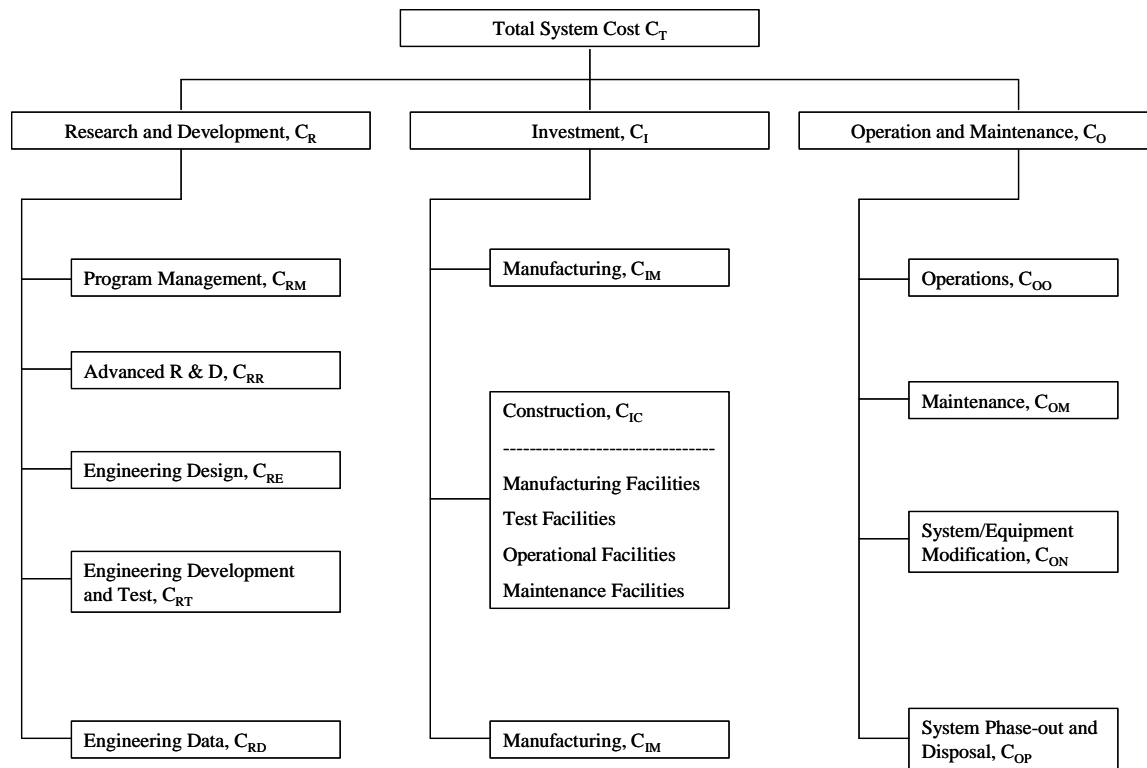
Figure 4. Cost Breakdown Structure (Blanchard, 1991)

Parametric models involve generation and application of equations that describe relationships between cost schedules and measurable attributes of a system that must be brought forth, sustained and retired (Dean, 1995). Cost estimation with a parametric model is based on predicting a product's cost either in total or for various activities, by the use of regression analysis based on historical cost and technical information. A simple parametric CER is the relation between the cost of buildings and the floor area. Most of the cost estimating relationships for airborne military systems relates the cost to the parameters such as weight, cruise speed, etc. of the system. Parametric estimating can involve considerable effort because of the systematic collection and revision process required to keep the CERs updated, but once this data is available estimates can be produced fairly rapidly (Greves and Schreiber, 1993). There are several commercial models available now. The most widely used is the Lockheed Martin's PRICE system. Establishments such as British Aerospace, The European Space Agency and

NASA use PRICE system. However, it is not recommended for estimating the cost of products that utilize new technologies.

55.   ANALOGOUS MODELS

Cost estimating made by analogy identifies a similar product or component and adjusts for differences between it and the target product (Shields and Young 1991). The effectiveness of this method depends heavily on an ability to identify correctly the differences between the case in hand and those deemed to be comparable. The main disadvantage of estimating by analogy is the high degree of judgment required.

56.   DETAILED MODELS

Detailed models use estimates of labour times and rates and also material quantities and prices to estimate the direct costs of a product or activity (Shields and Young, 1991). An allocation rate is then used to allow for indirect/overhead costs. This is known as bottom-up estimating and is widely used to allow indirect/overhead costs. It is the most time consuming and costly approach and requires a very detailed knowledge of the product and processes. However, the most accurate cost estimates can be made using this approach.   The method involves (Asiedu and Gu, 1998) estimation the time needed to perform an activity and the hourly rates for the man and machine, and then multiply times and rates to get costs. Time standards can be industry standards, in-house standards or based on expert guesses. In-house standards are the best but most difficult to develop. Industrial time standards for production operations exist for many common tasks.

57. IN THE NEXT FEW SECTIONS WE DESCRIBE FEW LCC MODELS THAT ARE POPULAR AMONG PRACTITIONERS THAT CAN BE USED FOR ESTIMATION OF LIFE CYCLE COST/COST OF OWNERSHIP OF AIRBORNE MILITARY EQUIPMENT.

58. 2.5 TAYLOR'S LCC MODEL

Taylor's model focuses on the capital and revenue costs. Taylor claims that in any discussion of trade-offs between initial and subsequent costs, a point that is frequently made is that there is a major distinction between initial capital costs and revenue costs. It is claimed that companies and public bodies faced with limited capital budget or cost limits do not have the facility to increase initial capital costs on the chance that there will be future revenue gains. However, Taylor claims that the distinction between revenue expenditure and capital is an accounting one which doesn't affect the life cycle cost concept based on the cash flows throughout the life of the asset.
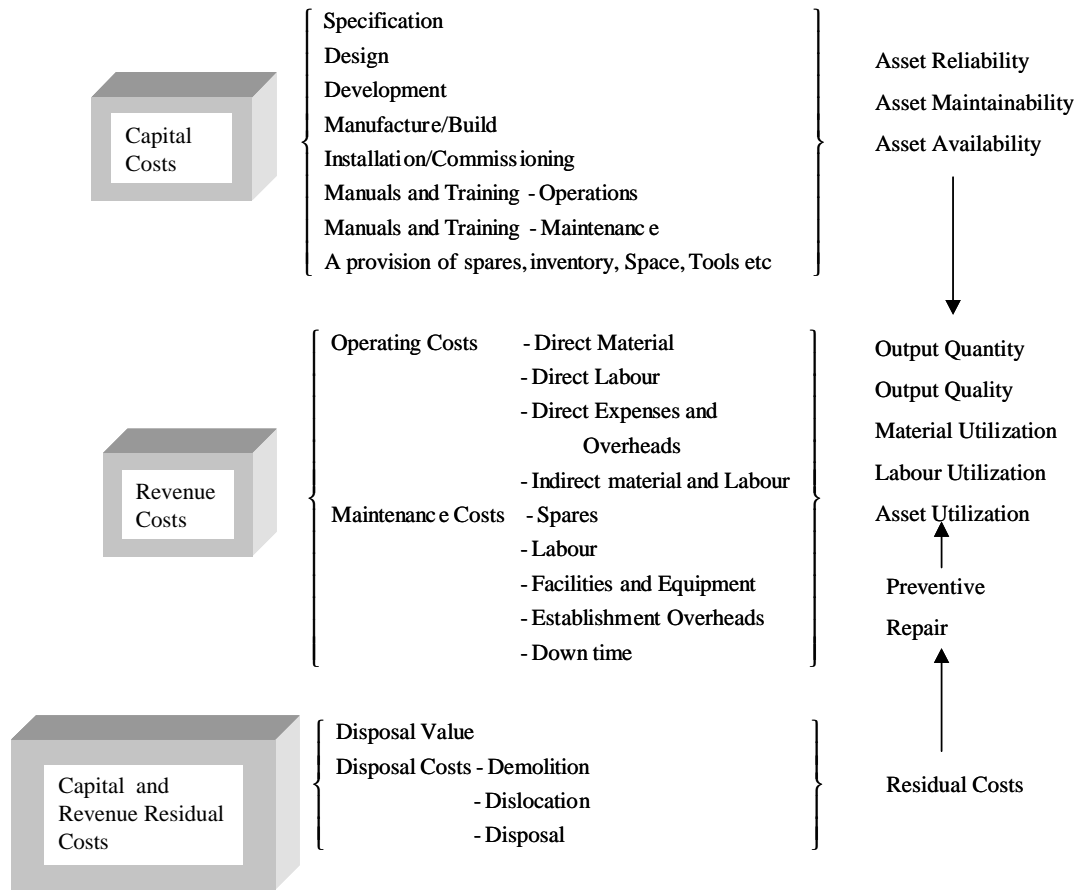
Figure 5. Taylor's LCC cost elements and interaction

Taylor's costs of owning physical asset are shown in Figure 5. The costs fall into three groups, first the initial capital costs secondly the revenue costs of operating and maintaining the asset during its operational life and thirdly the cost of asset disposal, which may be revenue of capital if it is substantial. The initial costs for an organization which designs and constructs physical assets for its own use or for resale would be:

Ø   Research and Development

Ø   Design and Specification.

Ø   Manufacturing.

Ø   Quality control and testing.

Ø   Monitoring Performance.

The second group of costs are incurred during the operational life of the asset and this would include the costs of: operating the assets including the labour, materials, tools, fixtures and overheads, maintenance including spares and labour. Finally there are disposal costs which include costs of demolition and removal, dislocation of existing production capacity. Against this may be any disposal value of the physical asset.

2.6   Raymer's LCC Model

Raymer's life cycle costing model is based on the Development and Procurement Costs of Aircraft (DAPCA IV) model developed by the Rand Corporation. The Rand corporation developed several cost estimation relationships for estimating various costs for all departments including engineering, tooling, manufacturing and quality control groups. DAPCA assumes a ten-year product life, which also is an industry standard. Rand Corporation claims that DAPCA, coupled with appropriate factors is accurate to within +/- 5% of actual costs.

2.7   Roskam LCC Model

Roskam model divides the LCC into four major categories: (1) Research and Development, test and evaluation, (2) Program acquisition cost that includes manufacturing cost and manufacturer's profit, (3) Operating cost, and (4) Disposal cost. The expression for LCC is given by the equation:

$$LCC = C_{RDTE} + C_{AC} + C_{OM} + C_{D}$$
(1)

Where, C_{RDTC} is the R&D, test and evaluation cost, C_{AC} is the acquisition cost, C_{OM} is the operation and maintenance cost and C_{D} is the disposal cost.

R&D cost is further broken into cost elements such as: (1) airframe engineering and design, (2) test flight aircraft and flight test operations, (3) test & simulation facilities, (4) cost to finance. Each of these cost elements is estimated using parametric methods using aircraft weight, maximum design speed, and number of aircraft built. Similarly acquisition cost is calculated using parameters such as the number of aircrafts manufactured, manufacturing cost, take-off weight, design cruise speed etc.

Operation costs are broken into the material costs, direct and in-direct personnel cost and logistic support costs. The disposal cost is taken as 1% of the LCC cost. Roskam developed several cost estimation relationships for estimation of the various costs given in above equation. Most of these models were developed using weight of the aircraft as a dependent variable.

2.8 Fabrycky and Blanchard's LCC Model

Fabrycky and Blanchard (1991) developed the detailed LCC model. The most important task in their model is to develop the cost breakdown structure (CBS, shown in Figure 4). There is no method set for breaking down the costs as long as the method used can be tailored to the specific application. Primarily the cost is divided into the following four categories:

- Ø Research and development
- Ø Production and construction costs
- Ø Operation and maintenance costs
- Ø Retirement and disposal costs

Thus the total cost is (C) is calculated using the expression:

$$C = C_R + C_P + C_O + C_D$$

(2)

Where,

C_R = R & D cost, C_P = Production and Construction Cost,
C_O = Operation and Maintenance Cost, C_R = Retirement and Disposal Cost

The total cost, C, includes all future costs associated with the acquisition, utilization, and subsequent disposal of system equipment.

Research and development cost includes all costs associated with conceptual feasibility studies, basic and advanced research and development, engineering design, fabrication and test of engineering prototype models (hardware), and associated documentation. Also covers all related program management functions. The R&D cost is given by:

$$C_R = C_{RM} + C_{RR} + C_{RE} + C_{RT} + C_{RD}$$

(3)

Where,

C_{RM} = Program management cost, C_{RR} = Advanced R&D Cost,
C_{RE} = Engineering design cost, C_{RT} = Engineering development/test cost
C_{RD} = Engineering data cost

Operations and maintenance cost includes all costs associated with the operation and maintenance support of the system throughout the life cycle subsequent to the equipment delivery in the field. Specific categories cover the cost of system operation, maintenance, sustaining logistic support, equipment modifications. Thus, the operation and maintenance cost is given by:

$$C_O = C_{OO} + C_{OM} + C_{ON} + C_{OP}$$

(4)

Where,

C_{OO} = Cost of system life cycle operations, C_{OM} = Cost of system life cycle maintenance,
C_{ON} = Cost of system life cycle modifications, C_{OP} = Cost of system disposal

The costs in equations (2) – (4) can be further divided into various cost elements.

2.9   Burns Life Cycle Cost Model

Burns developed a cost estimation relationship for predicting life cycle cost of aircraft based on its weight. Burns model is a simple extension of Roskam's life cycle cost model. The model also includes judgement factor for computing airframe-engineering hours for development and production. A complete analysis of Burns model is presented in Jayakrishnan (2002).

2.10 PRICE Life Cycle Costing System

The PRICE system consists of parametric cost estimation models for predicting the life cycle cost of weapon systems developed by the Lockheed Martin. The PRICE system's tool includes a set of four parametric cost estimation models, each with a different specialty area. The models consists of:

PRICE M: This model specifically addresses electronic module level hardware development and production costs.

PRICE H: This model specifically addresses the costs associated with development and production of hardware. This tool can use outputs of the PRICE M tool.

PRICE HL: This model uses data generated by PRICE H and calculates the hardware life-cycle costs, including sparing for a deployment environment.

PRICE Software: This model estimates both development costs and life cycle support costs for software.

2.11 Equipment Designer's Cost Analysis System (EDCAS) Model

EDCAS is one of the popular commercial systems available for life cycle cost prediction. EDCAS is a sequential model and is applicable for design to LCC in front-end design analysis. Over 500 government

and industry use the system worldwide. For example, U.S. Air Force uses EDCAS for aircraft and airborne weapons and electronic systems.

2.12   LCC Models Using Markov Chain

Stump (1988) developed a LCC model based on Markov chains and illustrated the model for a hypothetical remotely piloted vehicle (RPV).   The Markov chain is used to estimate the operation, maintenance and support costs. The model assumes that the system goes through a number of states. For any state, the number of visits per cycle multiplied by the cost per visit and the expected life of the RPV in cycles will yield a life cost for that state. Summing over all states will yield a total life cost. To fully implement this life cycle cost methodology, the following information is needed:

1.   A list of system states.

2.   A list of transition probabilities from any state to any other state (zero if the states do not communicate).

3.   A list of the costs of entering the states.

4.   The average number of visits per cycle for each state.

5.   The expected life of the system.

6.   Cost estimating relationships for computing costs.

The life cycle cost for state i is:

$$LC_i = L \times a_i \times C_i$$

(5)

Where, a_i, is the state probability for state i and C_i is the average cost of entry into state i and L is the expected life of the system. The total life cost of the RPV is simply the sum of the LC_i for all states. One of the major problems with Markov chain model is that the system is likely to have large number of states.

2.13 LCC model for labour factor

Dahlen and Bolmsjo (1996) developed a life cycle cost model for the labour factor that covers the costs for an employee over the whole employment cycle – from the requirement until retirement. The costs are divided into three basic categories:

1. Employment costs: consisting of costs for recruiting, introduction and training of new employees – to compare with acquisition costs such as projecting, installation and start-up of the new equipment.

2. Operations costs: consisting of wages, and labour related overhead – to compare with depreciation, maintenance and repairs.

3. Work environmental costs: consisting of a additional costs for absenteeism, rehabilitation and pensions – to compare with costs for increased maintenance and repairs and finally to scrap the equipment.

The basic categories of labour life cycle cost can be divided into: employment costs, operation costs and work environmental costs. The employment costs can be divided into three major sub categories: (1) recruitment costs, (2) additional productions costs and (3) education costs.

Operation costs are incurred when the employee is introduced and masters the work tasks, the costs consists of wages and overheads. The third category, work environmental costs, includes costs for absence, sickness benefits, rehabilitation costs and disability pension costs.

2.14 LCC Models in Designing for Logistic Support

Hatch and Bedinelli (1999) developed a model that carries out a concurrent optimization of a product design and its associated manufacturing and logistic support system. The model is constructed

which links together the decisions associated with three major phases of the life cycle: product design, manufacturing and logistic system design, production and field operation control. The model included an optimization scheme that concurrently optimizes the decision variable of the linked model. The final solution prescribed by the model is based on a multi-criteria value function formed from the individual objectives of minimizing life cycle cost and maximizing availability. The model evaluates alternative design solutions by calculating the associated operational availability as well as manufacturing and logistic support costs. The two main performance measures can be combined into the following bi-criteria model formulation:

Min   Life Cycle Cost

Max   System Availability

Subject to   Product Design Requirements

2.15   Applications of LCC/Cost of Ownership Models

The literature survey carried out by Nicholas (1999) indicates that the term life-cycle cost/cost of ownership is applied to varying forms of analysis, which are undertaken for a range of different purposes. These applications can be broadly described as evaluation and decision-making, planning and budgeting, cost management and control, project/program management and control, life-cycle management, contracting, and marketing. These applications are shown in Table 1 in relation to recently published papers.

Although life-cycle cost/cost of ownership is applied to analysis, which is used for a wide range of purposes, it can be argued that in essence the analysis is being used to provide information for two fundamental purposes - planning and control. Planning includes i) decision-making - the allocation of

266

resources for future periods through the identification, evaluation and selection of alternative courses of action and ii) budgeting - the identification of means required to implement the selected courses of action. Principal areas of control include cost control and contractual arrangements. Cost control includes procedures to influence cost through design (Michaels and Wood, 1989) and throughout the process of acquisition (US Department of Defense, 1996). Contractual arrangements are designed to control cost through legal agreements. Contractual arrangements include guarantees for part of life-cycle cost such as support cost (Baathe, 1995) and guarantees for total life cycle cost (Akselsson and Burstrom, 1994). Planning and control are integrated through management which includes project and program management during acquisition and much wider life-cycle management which extends to the complete life-cycle of the system.

There is an important distinction to be made between the nature of the use of analysis to provide information for planning and control. In planning, in both decision-making and budgeting, life-cycle cost is used as an ex ante concept to predict future cost. In control, it is used as both an ex ante and an ex post concept. It is used as an ex ante concept to establish targets or performance criteria but as an ex post concept when monitoring and comparing cost performance in terms of planned cost against actual cost. As indicated in the discussion of concepts of cost below, this will involve the application of different concepts of cost.

Analysis using the concept of life-cycle cost is not only used for different purposes but also involves different functional users. Examples of different users include policy-makers (Kirkpatrick, 1996), managers (Riggs and Jones, 1990, Greene, 1991) and engineers (Various - see references and bibliography). Policy-makers are involved in strategic decisions that involve the long-term commitment of funds. Managers may be project managers or budget managers whose interest is principally in control. Engineers may include design engineers, production engineers, systems

engineers, logistics engineers and others who have an interest in life-cycle cost for engineering decision-making.

Table 1. Application of LCC/TCO in the literature

Application of Life Cycle Cost and Cost of Ownership Models	
Design evaluation	Dacko and Darlington (1988) Takagishi (1989) Gibbs and King (1989) Johnson (1990) Keene and Keene (1993) Stahl and Wallace (1995) Plebani, Rosi and Zanetta (1996) Asiedu and Gu (1998)
Materials selection	Winkel (1996)
Choice of design life	Howard (1991) Asiedu and Gu (1998)
Environmental evaluation	Fiksel and Wapman (1994) Vivona (1994)
Evaluation of technology developments	Curry (1993) Vacek, Hopkins and MacPherson (1995)
Production/manufacturing	Wilkinson (1990) Malkki, Enwald and Toivonen (1991)
Reliability analysis	Zhou and Cai (1994)
Failure analysis	Rooney and Jackson (1996)
Availability analysis	Fairclough (1989)
Maintenance	Lansdowne (1994) Dinesh Kumar (2000)
Maintainability	Govil (1992), Dinesh Kumar (2000)
Condition monitoring	Hutton (1994)
Logistics support analysis	McArthur and Snyder (1989)
Operation and support	Curry (1989) Snyder (1990) Stone, Drubka and Braun (1994)
Transportation	Wonsiewicz (1988) Tzemos (1990)
Value engineering	Harding (1996)
Life-cycle cost benefit analysis	Adler, Herkamp, Wiesler and Williams (1995)
Planning and budgeting	
Procurement strategy	Profitt (1994)
Business planning	Jones (1994)
Budgetary provision	Kirkpatrick (1995)
Manpower, personnel and training planning	Cole (1991)
Cost management and control	
Cost management	Fabrycky and Blanchard (1991)
Design for/to Cost	Michaels and Wood (1989) Dean and Unal (1991)
Project/program management and control	
Program management	Greene (1991) Zhi (1993)
Project control	Goble and Paul (1995)

58.1   Table 1 Continued

Management	
Life-cycle management	Hell (1995)
Physical asset management	Hodges (1996) Sherwin (1996)
Activity-based management	Brimson and Antos (1994)
Contracting	

Contract provision	Akselsson and Burstrom (1994) Baathe (1995)
Marketing	
Marketing of commercial products	Carruba (1992)

59.

List of References

Adler, D.J., Herkamp, J.A., Wiesler, J.R. and Williams, S.B. (1995) 'Life-Cycle Cost and Benefits of Process Automation in Bulk Pharmaceuticals', *ISA Transactions*, Vol. 34, No. 2, pp. 133-139.

Ahmed, N. U. (1995) 'A design and Implementation Model for Life Cycle Cost Management System', *Information and Management,* Vol. 28, pp. 261-269.

Akselsson, H. and Burstrom, B. (1994) 'Life-Cycle Cost Procurement of Swedish State Railways High-Speed Train X2000', *Proceedings of the Institution of Mechanical Engineers, Part F, Journal of Rail and Rapid Transit*, Vol. 208, No. 1, pp.51-59.

Asiedu, Y. and Gu, P. (1998) 'Product Life Cycle Cost Analysis: State of the Art Review', *International Journal of Production Economics,* Vol. 36, No. 4, 883-908.

Baathe, O. (1995) 'Life Cycle Management for the A/C JAS 39 Grippen', Proceedings of 11th International Congress, Society of Logistics Engineers, Stockholm, pp. 159-167.

Blanchard, B.S. (1972) '*Design and Manage to Life Cycle Cost*', M/A Press, Oregon, USA.

Blanchard, B.S. (1986) *Logistics Engineering and Management*, 3rd Edition, Prentice-Hall, Englewood Cliffs, New Jersey.

Blanchard, B.S. (1998) *System Engineering Management*, 2nd Edition, John Wiley, New York.

Blanchard, B.S. and Fabrycky, W.J. (1990) *Systems Engineering and Analysis*, 2nd Edition, Prentice-Hall, Englewood Cliffs, New Jersey.

Brimson, J.A. and Antos, J. (1994) *Activity-Based Management*, John Wiley, New York.

Burns, J. W. (1994) 'Aircraft Cost Estimation Methodology and Value of a Pound Derivation for Preliminary Design Development Applications, SAWE Paper No. 2228, Long Beach, USA.

Carrubba, E.R. (1992) 'Integrating Life-Cycle Cost and Cost of Ownership in the Commercial Sector', *Proceedings of the Annual Reliability and Maintainability Symposium*, pp. 101-108.

Checkland, P. (1981) *Systems Thinking, Systems Practice*, John Wiley, New York.

Curry, E.E. (1989) 'STEP - A Tool for Estimating Avionics Life Cycle Costs', *IEEE Aerospace and Electronic Systems Magazine*, Vol. 4, No. 1, pp. 30-32.

Curry, E.E. (1993) 'FALCCM-H: Functional Avionics Life Cycle Cost Mode for Hardware', *1993 IEEE National Aerospace and Electronics Conference*, Dayton, Ohio, May 24-28, pp.950-953.

Dacko, L.M. and Darlington, R.F. (1988) 'Life Cycle Cost Procedure for Commercial Aircraft Subsystem', Proceedings of the *Annual Reliability and Maintainability Symposium, 1988*, Los Angeles, January 26-28, pp.389-394.

Dahlen, P., and Bolmsjo, B.S. (1996) 'Life Cycle Cost Analysis of Labour Factor,' *International Journal of Production Economics,* Vol. 46-47, pp. 459-467.

Dean, E. B. (1995) 'Parametric Cost Deployment', Proceedings of the Seventh Symposium on Quality Function Deployment, MI, USA, pp. 27-34.

Degraeve, Z., and Roodhooft F. (1999) 'Improving the Efficiency of the Purchasing Process Using the Total Cost of Ownership Information: The Case of Heating Electrodes at Cockerill Sambre S.A.,' *European Journal of Operational Research*, Vol. 112, pp. 42-53.

De La Mare, R.F. (1982) *'Manufacturing Systems Economics: The Life Cycle Costs and Benefits of Industrial Assets'*, Holt, Rinehart and Winston, London.

Dewhurst. P., and Boothroyd, G. (1988) 'Early Cost Estimating in Product Design ,' *Journal of Manufacturing Systems,'* Vol. 7, pp. 183-191.

Dhillon, B.S. (1989) *Life Cycle Costing : Techniques, Models and Applications*, Gordon and Breach, New York.

Dienemann, P.F. (1966) 'Estimating Cost Uncertainty Using Monte Carlo Techniques', *Memorandum RM-4854-PR, The RAND Corporation*, Santa Monica, California, January.

Dowlatshahi, S. (1992) 'Product Design in Concurrent Engineering Environment: An Optimization Approach', *Journal of Production Research,* Vol. 30, 1803-1818.

Fabrycky, W.J. (1994) 'Modeling and Indirect Experimentation in System Design Evaluation', *The Journal of NCOSE*, Vol. 1, No. 1, pp. 133-144.

Fabrycky, W.J. and Blanchard, B.S. (1991) *Life-Cycle Cost and Economic Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey.

Fabrycky, W.J. and Thuesen, G.J. (1980) *Economic Decision Analysis*, 2nd edition, Prentice-Hall, Englewood Cliffs, New Jersey.

Fairclough, M.R. (1989) 'Arena - A Software Aid for Measuring System Availability and Life Cycle Cost', *Reliability '89, Part 2*, London, June 14-16, pp.p4B/1-4B/1/6.

Fiksel, J. and Wapman, K. (1994) 'How to Design for Environment and Minimize Life Cycle Cost', *Proceedings of the 1994 IEEE International Symposium on Electronics and the Environment*, San Francisco, pp. 75-80.

Finkelstein, W. and Guertin, J.A.R. (1988) *Integrated Logistics Support : The Design Engineering Link*, IFS Publications, Kempston, Bedford.

Fisher, G.H. (1971) *Cost Considerations in Systems Analysis*, American Elsevier, New York.

Flanagan, R. and Norman, G. (1983), *Life Cycle Costing for Construction, Quantity* Surveyors Division of the Royal Institution of Charted Surveyors, London.

Gibbs, D.J.L. and King, R.J. (1989) 'Life Cycle Costing in the Design of Naval Equipment', *Reliability '89*, Part 1, No. 1989, June 14-16, pp. 2A/5/1 - 2A/ 5/8.

Govil, K.K. (1992) 'A Simple Model for Life-Cycle vs Maintainability Function', *Microelectronics and Reliability*, Vol. 32, No. 1-2, pp. 269-270.

Greene, L. E., and Shaw, B. L. 1990 'The Steps for Successful Life Cycle Cost Analysis,' *Proceedings of the IEEE 1990 National Aerospace and Electronics Conference,* Dayton, Ohio, pp. 1209-1216.

Greves, Y. P., and Schreiber, B. (1993) 'Engineering Costing Techniques in ESA, Internet site: http://www.esapub.esriu.esa.it.pointtobullet/greves.html.

Gupta, Y.P. (1983) 'Life Cycle Cost Models and Associated Uncertainties' in Skwirzynski, J. K. (1983) (Ed.), *Electronic Systems Effectiveness and Life Cycle Costing*, Springer-Verlag, Berlin, pp. 535 - 549.

Gupta, Y., and Chow, W. S. (1985) 'Twenty-five Years of Life Cycle Costing – Theory and Application: A Survey', *The International Journal of Quality and Reliability Management,* Vol. 2, pp. 551-576.

Habayeb, A.R. (1987) *System Effectiveness*, Pergamon Press, Oxford.

Hall, A.D. (1962) *A Methodology for Systems Engineering*, Van Nostrand, Princeton, New Jersey.

Hanssmann, F. (1968) *Operations Research Techniques for Capital Investment*, John Wiley, New York.

Hart, G.W. (1985) 'Interpretation of Life Cycle Costs', *Reliability '85 - Proceedings* , Vol. 1, July 10-12, pp. 3A. 4. 1-3A. 4. 5.

Hart, G.W. (1985) 'Interpretation of Life Cycle Costs', *Reliability '85 - Proceedings*, National Centre of Systems Reliability, UKAEA and Institute of Quality Assurance, Vol. 1, 10-12 July, pp. 3A.4.1 - 3A.4.5.

Harvey, G. (1976) 'Life Cycle Costing: A Review of the Technique', *Management Accounting*, pp. 343-347.

Hodges, N.W. (1996) *'The Economic Management of Physical Assets'*, Mechanical Engineering Publications, London.

Howard, R.J. (1991) 'Road Life Cycle Costing', *National Conference Publication - Institution of Engineers, Australia*, No. 91, Part 14, pp. 55-59.

Jayakrishnan, V. E. (2002) '*Model for Fighter Aircraft Life Cycle Cost Estimation,*' M.Tech. Thesis, Indian Institute of Technology Bombay, India

Johnson, V.S. (1990) 'Minimizing Life-Cycle Cost for Subsonic Commercial Aircraft', *Journal of Aircraft*, Vol. 27, No. 2, pp. 139-145.

Jones, C. (1994) 'Life Cycle Cost Models', *Colloquium on Life Cycle Costing and the Business Plan*, London, February 1, pp.2/1 - 2/6.

Jordan, N. (1960) 'Some Thinking About 'System'', in Optner (1973), pp. 53-72.

Kaufman, R. J. (1970) 'Life Cycle Costing: A Decision Making Tool of Capital Equipment Acquisition,' *Cost and Management*, pp. 21-28.

Keene, S.J. and Keene, K.C. (1993) 'Reducing the Life Cycle Cost of Software through Concurrent Engineering', *Proceedings of the 1993 Annual Reliability and Maintainability Symposium*, Atlanta, Georgia, January 26-28, pp.305-309.

Kirkpartick, P. S., and Jones, A. L. (1974) 'Life Cycle Cost Comparisons of Avionic System Design Alternatives,' *Proceedings of the IEEE 1974 National Aerospace and Electronics Conference,* Dayton, Ohio, pp. 514-520.

Knezevic, J. (1993) *Reliability, Maintainability and Supportability : A Probabilistic Approach*, McGraw-Hill , London.

Knezevic, J. (1997) *Systems Maintainability : Analysis, Engineering and Management*, Chapman and Hall, London.

Knotts, R. M. H. 1996 '*Analysis of the Impact of Reliability, Maintainability and Supportability on the Business and Economics of Civil Air Transport Aircraft Maintenance and Support*, M. Phil. Thesis, University of Exeter, UK

Lamar, W.E. (1981) 'A Review and Assessment of System Cost Reduction Activities', *AGARD Conference Proceedings*, No. 289, Design to Cost and Life Cycle Cost' in Dhillon, B.S. (1989) *Life Cycle Costing : Techniques, Models and Applications*, Gordon and Breach, New York.

Lamb, R.G. (1995) *Availability Engineering and Management for Manufacturing Plant Performance*, Prentice-Hall, Englewood Cliffs, New Jersey.

Lansdowne, Z.F., (1994) 'Built-In-Test Factors in a Life-Cycle Cost Model', *Reliability Engineering and System Safety*, Vol. 43, No. 3, pp. 325-330.

Macedo, M.C., Dobrow, P. V., and O'Rourke, J J. (1978) '*Value Management for Construction*,' John Wiley & Sons, Chiester, UK.

Malkki, H., Enwald, P. and Toivonen, J. (1991) 'Experience of Transferring Life Cycle Costing to Manufacturing Industry', *American Society of Mechanical Engineers*, pp. 1-8.

McNichols, G. R. (1983a) A Reported Presentation on Cost Growth in Skwirzynski, J. K. (1983) (Ed.), *Electronic Systems Effectiveness and Life Cycle Costing*, Springer-Verlag, Berlin, pp. 567 - 582.

McNichols, G. R. (1983b) 'Uncertainties of LCC Predictions' in Skwirzynski, J. K. (1983) (Ed.), *Electronic Systems Effectiveness and Life Cycle Costing*, Springer-Verlag, Berlin, pp. 583 - 594.

Michaels, J.V. and Wood, W.P. (1989) *Design to Cost*, John Wiley, New York.

Nicholas, C.S. (1999) '*Life Cycle Costing',* M.Phil. Thesis, University of Exeter, UK

Nicholas, C.S. (1994) 'Determining and Achieving Operational Availability - Economic Impact', *Proceedings of 10th International Congress*, Society of Logistics Engineers, Exeter, March, pp. 275-283.

Nicholas, C.S. (1996a) 'The Impact of Discounting Life-Cycle Cost on Supportability Engineering', *Communications in Reliability, Maintainability, Supportability and Logistics'*, Vol. 3, No. 1, pp. 42-46.

Nicholas, C.S. (1996b) 'Supportability Engineering - The Economic Dimension', *Proceedings of the Society of Logistics Engineers Annual International Symposium*, Washington, pp. 151-157.

Noble, J. S., and Tanchoco, J. M. A. (1990) 'Concurrent Design and Economic Justification in Developing a Product', *International Journal of Production Research, Vol. 28, pp. 1225-1238.*

Ostwald, P.F. (1992) *Cost Estimating*, 3rd edition, Prentice-Hall, Englewood Cliffs, New Jersey.

Plebani, S., Rosi, R. and Zanetta, L. (1996) 'Life Cycle Costs Comparison and Sensitivity Analysis for Multimedia Networks', *Proceedings of the 1996 IEE Colloquium on Optical and Hybrid Access Networks*, No. 052, London, March 4-5, pp. 3/1-3/6.

Proffitt, J.T. (1994) 'Life Cycle Cost of Aircraft Systems', Colloquium on Life Cycle Costing and the Business Plan, *IEE Colloquium (Digest)*, No. 013, pp. 1/1-1/3.

Perera, H.S.C., Nagarur, N., and Tabucanon, M.T. (1999) 'Component Part Standardization: A Way to Reduce the Life Cycle Costs of Products,' *International Journal of Production Economics,'* Vol. 60-61, pp. 109-116.

Quade, E.S. (1963b) 'The Limitations of A Cost-Effectiveness Approach to Military Decision-Making', *Report P-2798*, The RAND Corporation, Santa Monica, California.

Quade, E.S. (1965) 'Cost-Effectiveness Analysis: An Appreciation, *Report P-3248*, The RAND Corporation, Santa Monica, California.

Raymer, D.P. (1992) *Aircraft Design: A Conceptual Approach*, American Institute of Aeronautics and Astronautics, Washington.

Riggs, J.L. and Jones, D. 'Flowgraph representation of Life Cycle Cost Methodology - A New Perspective for Project Managers', IEEE Transactions on Engineering Management, Vol. 37, No.2, pp.147-152.

Riggs, J.L. (1968) *Economic Decision Models for Engineers and Managers*, McGraw-Haill, New York.

Rooney, C. and Jackson, E. (1996) 'Unit Failure: A Life Cycle Cost Analysis', *Glass Digest*, Vol. 75, No.2, pp.7.

Rose, A. (1984) 'Life Cycle Costing Today', *28th Annual Meeting of the American Association of Cost Engineers and the Eighth International Cost Engineering Congress*, Montreal, June 24-27, pp. J.3.1 - J.3. 4.

Roskam, J. (1990) '*Airplane Design Part VIII: Airplane Cost Estimation Design, Development, Manufacturing and Operating,* Roskam Aviation and Engineering Corporation Report, Kansas, USA.

Sabbagh, K. (1995) *21st Century Jet : The Making of the Boeing 777*, Macmillan, London.

SAE (1992) 'Life Cycle Cost - Techniques and Applications', *Aerospace Recommended Practice*, SAE ARP4293, SAE, Warrendale.

Scherer, F.M. (1964) *The Weapons Acquisition Process : Economic Incentives*, Harvard University, Harvard.

Sherif, Y.S. and Kolarik, W. J. (1981) 'Life Cycle Costing: Concept and Practice', *OMEGA, The International Journal of Management Science*, Vol. 9, No. 3, pp. 287 - 296.

Sherwin, D. (1996) 'Economic and Accounting Aspects Of Terotechnology and Life-Cycle Costing', *Proceedings of the International Conference of Maintenance Societies*, Melborne, May 22-23, Paper 46.

Shields, M. D., and Young, S. M. 1991 'Managing Product Life Cycle Cost: An Organizational Model, *Cost Management,* pp. 39-52.

Simon, H. A. (1959) 'Theories of Decision Making in Economics and Behavioural Science', American Economic Review, June.

Singleton, W.T. (1974) *Man-Machine Systems*, Penguin, Harmondsworth.

Snyder, H.M. (1990) 'Life Cycle Cost Model for Dormant Systems', *Proceedings of the IEEE 1990 National Aerospace and Electronics Conference - NAECON 1990*, Vol.3, Dayton, May 21-25, pp.1217-1219.

Srinivasan, N. K., Nageswara A. M., and Prabhakar B K. (1997) '*CASSA Model for Life Cycle Cost Estimation for Aircraft,*' Report No. S-97-1, CASSA, DRDO, Bangalore.

Stahl, L. and Wallace, M. (1995) 'Life Cycle Cost Comparison - Traditional Cooling Systems vs. Natural', *Proceedings of the 1995 17th International Telecommunications Energy Conference*, Hague, October 29 - November 1, pp.259-265.

Stone, K.W., Drubka, R.E.and Braun, H. (1994) 'Impact of Stirling Engine Operational Requirements on Dish Stirling System Life Cycle Costs', *Proceedings of the 1994 ASME/JSME/JSES International Solar Energy Conference*, San Francisco, March 27-30, pp.529-534.

Stump, E. J. (1988) 'An Application of Markov Chains to Life Cycle Cost Analysis,' *Engineering Costs and Production Economics,* Vol. 14, pp. 151-156.

Takagishi, S.K. (1989) 'Electric Power versus Petrol, Methanol or Gas. Life Cycle Cost Comparison', *9th International Electric Vehicle Symposium, Electric Vehicle Developments,*Vol. 8, No. 3, Toronto, November 13-15, pp.77-81.

Taylor, W.B. (1981) 'The Use of Life Cycle Costing in Acquiring Physical Assets', *Long Range Planning,* Vol. 14, pp. 32-43.

Thuesen, G.J. and Fabrycky, W.J. (1989) *Engineering Economy*, 7th edition, Prentice-Hall, Englewood Cliffs, New Jersey.

UK Department of Industry (1977) 'Life Cycle Costing in the Management of Assets', *Department of Industry, Committee for Terotechnology*, H.M.S.O., London.

US Department of Defense, (1983a) 'Design to Cost', *Directive 4245.3*, US Department of Defense.

US Department of Defense, (1983b) 'Life Cycle Cost in Navy Acquisitions', MIL-HDBK-259 (Navy), US Department of Defense.

US Department of Defense, (1996) 'Major Systems Acquisitions', *Directive 5000.1 (Revised)*, US Department of Defense.

Vacek, R.M., Hopkins, M. and MacPherson, W.H. (1996) 'Development, Demonstration and Integration of Advanced Technologies to Improve the Life Cycle Costs of Space Systems', *Proceedings of the 1996 IEEE Aerospace Applications Conference*, Vol. 1, IEEE, Los Alamos, pp. 217-225.

Verma, D. (1995) 'Conceptual System Design Synthesis, Analysis and Evaluation', *Technical Monograph 1*, Centre for Management of Industrial Reliability, Cost and Effectiveness, Faculty of Engineering, University of Exeter, Exeter.

Vivona, M.A. (1994) 'Audit Environmental Processes Using Life-Cycle Costs', *Hydrocarbon Processing*, Vol. 73, No. 8, p. 115.

White, G. E., and Ostwald, P. H. (1976) '*Management Accounting',* pp. 39-42.

White, J.A., Agee, M.H. and Case, K.E. (1977) *Principles of Engineering Economic Analysis*, John Wiley, New York.

Wilkinson, V.K. (1990) 'Life Cycle Cost Analyses of Government Production and Services', *Proceedings of the 1990 International Industrial Engineering Conference*, No. 1990, San Francisco, May 20-23, pp. 111-116.

Winkel, J.D. (1996) 'Use of Life Cycle Costing in New and Mature Applications', *Proceedings of the 1996 NPF/SPE European Production Operations Conference*, Society of Petroleum Engineers (SPE), Stavanger, Norway, April 16-17, pp. 239-248.

Wonsiewicz, T.J. (1988) 'Life Cycle Cost Analysis. Key Assumptions and Sensitivity of Results', *Transportation Research Record*, Lane Enterprises Inc, Camp Hill, pp.113-117.

Woodward D. G. (1997) 'Life Cycle Costing – Theory, Information, Acquisition and Application,' *International Journal of Project Management,* Vol. 15, No. 6, pp. 335-344.

Woud, J.K., Smit, K. and Vucinic, B. (1997) 'Maintenance Programme Design for Minimal Life Cycle Costs and Acceptable Safety Risks', *International Shipbuilding Progress*, Vol. 44, No. 437, Delft University Press, Delft, Netherlands, pp.77-100.

Zhi, H. (1993) 'Simulation Analysis in Project Life Cycle Cost', *Cost Engineering*, Vol. 35, No. 12, Morgantown, West Virginia, pp. 13-17.

Zhou, X. and Cai, N. (1994) 'Dynamic Study of the Relationship between Life Cycle Cost and Reliability in Electronic Product', *Tien Tzu Hsueh Pao/Acta Electronica Sinica*, Vol. 22, No. 11, Hangzhou Inst. of Electronics Engineering, Hangzhou, China, pp.22-27.

Total Cost of Ownership for Railway Assets: A Case Study on BOXN Wagons of Indian Railways

U Dinesh Kumar*, Gopinath Chattopadhyay** and H S Pannu***

*Indian Institute of Management Calcutta, Joka, D H Road, Kolkata 700104, INDIA

**Dept. of Mechanical, Manufacturing and Medical Engineering, Queensland University of Technology, Brisbane, QLD 4001, Australia

***South Eastern Railway, Garden Reach Road, Kolkata 700043, INDIA

dinesh@iimcal.ac.in, g.chattopadhyay@qut.edu.au, h_s_pannu@yahoo.co.in

ABSTRACT

Decision to purchase any capital equipment must be based on its total cost of ownership (TCO) rather than the usual practice of procurement based on the initial purchase price. In the recent years, TCO has become a part of the strategic cost management and the concept can be applied for effective procurement of railway assets. TCO provides an insight into the total cost of acquisition and sustenance and thus effectively support decision-making in evaluation of various alternatives.   The primary objective of this paper is to develop models for prediction of cost of ownership of capital assets. The models are developed using the Markov and renewal processes depending on the time to failure distribution of individual items within the capital equipment. The models developed in this paper are validated using the data from BOXN wagon used by the Indian Railways.

Keywords: Asset management, Maintenance, Markov and renewal processes, Total cost of ownership.

1.   INTRODUCTION

"Value for Money" has become one of the important criteria in an increasingly competitive business environment. Life Cycle Cost (LCC) and the Total Cost of Ownership (TCO) are two important financial measures that are used for decision making in acquisitions to evaluate the value of any capital equipment (Hampton, 2004; Humphries, 2004). Life cycle cost refers to all costs associated with the

product or system as applied to defined life cycle. That is, starting from requirement analysis, design, production, operation and maintenance till disposal. Total cost of ownership (TCO) is a philosophy, which is aimed at understanding the true cost of buying a particular product or service from a particular supplier. From its origins in defence equipment procurement in the US in early 1960s, the use of life cycle cost and cost of ownership has extended to other areas of the public and private sectors. LCC and TCO are being used to assist in decision-making, budget planning, cost control, and range of other activities that occur over the life of complex technological equipment.

LCC analysis is applied routinely to military projects (Blanchard, 1986). In the military sector the consumer, by funding the project and operating the related product, essentially bears the total life cycle cost covering the major cost elements in all stages of a product's life cycle. The term LCC analysis is rarely used in the commercial sector. Instead, the main focus is on TCO where related costs, covering acquisition (purchase or lease), operation, maintenance and support are borne by the customer. In addition, the customer can also incur costs when the product is not available for use, that is, '*down time costs*'.

The objectives of LCC/TCO are (Flanagan and Norman, 1983):

- To enable investment options to be more effectively evaluated.
- To consider the impact of all costs rather than only the initial capital costs.
- To assist in the effective management of completed projects.
- To facilitate choice between competing alternatives.

In the Defence industry the system's life cycle is divided into various phases, which allow proper planning and control of a project. The number of phases depends on the nature of the project, purpose and whether they are applied to commercial, military or space projects (Knotts, 1998). Commonly used phases are:

8. Requirements (Functional Specification).
9. Concept/Feasibility Studies.
10. Design and Development.
11. Production.
12. Testing and Certification.
13. Operation, Maintenance and Support.
14. Disposal

It is reported by the US Department of Defence that 70% of weapon system life cycle cost is committed by the end of concept studies, 85% by the end of system definition and 95% by the end of full scale development (Knotts, 1998). The US Department of Defence has formally used the concept of life cycle cost in weapon system acquisition since the early 1960s through life cycle costing and life cycle cost analysis.

The cost of ownership approach identifies all future costs and reduces them to their present value by use of the discounting techniques through which the economic worth of a product or product

options can be assessed. In order to achieve these objectives the following elements of cost of ownership have been identified (Woorward, 1997):

- Initial capital costs
- Life of the asset
- The discount rate
- Operating and maintenance costs
- Disposal cost

1.2   Initial capital costs

The initial capital costs can be divided into three sub-categories of cost namely: (1) purchase costs, (2) acquisition/finance costs, and (3) installation/commissioning/training costs. Purchase costs will include assessment of items such as land, buildings, fees, and equipment. Finance costs include alternative sources of funds. Basically, the initial capital cost category includes all the costs of buying the physical asset and bringing it into operation.

1.   Life of the Asset

The estimated life of an asset has a major influence on life cycle cost analysis. Ferry *et al* (1991) has defined the following five possible determinants of an asset's life expectancy:

- *Functional life* – the period over which the need for the asset is anticipated.

- *Physical life* – the period over which the asset may be expected to last physically, to when replacement or major rehabilitation is physically required.

- *Technological life* – the period until technical obsolescence dictates replacement due to the development of a technologically superior alternative.

- *Economic life* – the period until economic obsolescence dictates replacement with a lower cost alternative.

- *Social and legal life* – the period until human desire or legal requirement dictates replacement.

1.4   The discount rate

As the cost of ownership is discounted to their present value, selection of a suitable discount rate is crucial for TCO analysis. A high discount rate will tend to favour options with low capital cost, short life and high recurring cost, whilst a low discount rate will have the opposite effect.

1.5 Operations and Maintenance Costs

Cost of ownership, in many cases, is about operation and maintenance cost. Estimation of operation and maintenance costs is the essential to minimise the total cost of ownership of the asset. In the whole of TCO analysis, estimation of operation and maintenance is the most challenging task.

1.6 Disposal cost

This is the cost incurred at the end of as asset's working life in disposing of the asset. The disposal cost would include the cost of demolition, scrapping or selling the asset.

1.7 Uncertainties and Sensitivity Analysis

LCC/TCO is highly dependent on the assumptions and estimates made while collecting data. Even though it is possible to improve the quality of these estimates, there is always an element of uncertainty associated with these estimates and assumptions. Macedo *et al* (1978) identifies the following five major sources of uncertainty:

6. Differences between the actual and expected performance of the system could affect future operation and maintenance cost.
7. Changes in operational assumptions arising from modifications in user activity.
8. Future technological advances that could provide lower cost alternatives and hence shorten the economic life of any system/subsystem.
9. Changes in the price levels of major resources such as energy or manpower, relative to other resources can affect future alteration costs.
10. Error in estimating relationships, price rates for specific resources and the rate of inflation in overall costs from the time of estimation to the availability of the asset.

While undertaking a LCC/TCO analysis, there may be some key parameters about which uncertainty exists, usually because of the inadequacy of the input data. Blanchard (1972) suggested the following should be the subject of sensitivity analysis:

1. Frequency of the maintenance factor.
2. Variation of the asset's utilization or operating time.
3. Extent of the system's self-diagnostic capability.
4. Variation of corrective maintenance hours per operating hour.
5. Product demand rate.
6. The discount rate

In this paper, we develop mathematical models for prediction of total cost of ownership. The rest of the paper is organized as follows. In Section 2, we have developed a framework for estimating the total cost of ownership. Section 3 deals with the mathematical models for total cost of ownership. A case study on BOXN wagon used by the Indian Railways is used to illustrate the mathematical model developed in the paper.

2. Framework for Total Cost of Ownership Model

Total cost of ownership is driven by reliability, maintainability and supportability. The objective of total cost of ownership is to minimise TCO by optimizing reliability, maintainability and supportability. Figure 1 illustrates the relationship between the system operational effectiveness and other design parameters (Dinesh Kumar 2000). Total cost of ownership will decrease as the reliability increases. Similarly better maintainability and supportability would decrease the maintenance and support cost and hence will decrease the total cost of ownership. However, increasing reliability, maintainability and supportability may require additional resources during the design and product development stage and hence is likely to increase the initial procurement cost.
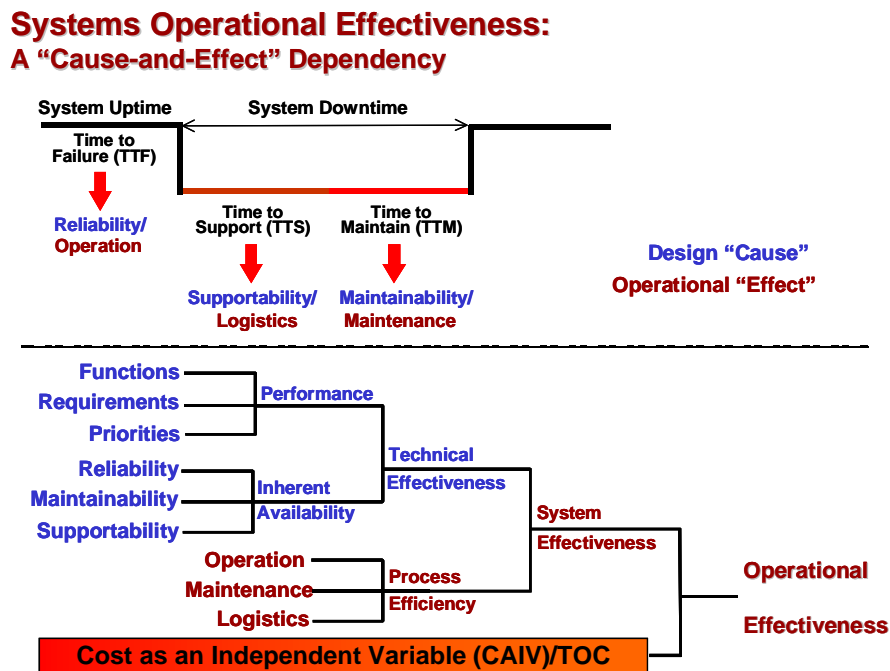
Figure 1. Cause and effect dependency between operational effectiveness, total cost of ownership and other design parameters

The framework for calculating total cost of ownership can be very complex depending on the procurement and asset management strategies used by the user. In this paper, we mainly focus on

procurement, operation, maintenance and disposal cost, which are more relevant for assets like wagons. The framework shown in Figure 2 is used for evaluation of total cost of ownership.

```
                    ┌─────────────────────────┐
                    │  Total Cost of Ownership │
                    └─────────────────────────┘
              ┌──────────────┴──────────────┐
    ┌──────────────────┐        ┌──────────────────┐
    │ Acquisition Cost │        │  In-Service Cost │
    └──────────────────┘        └──────────────────┘
                            ┌──────────┴──────────┐
                   ┌──────────────────┐  ┌──────────────────┐
                   │  Operating Cost  │  │   Support Cost   │
                   └──────────────────┘  └──────────────────┘
                                     ┌──────────┴──────────┐
                            ┌──────────────────┐  ┌──────────────────┐
                            │ Maintenance Cost │  │  Logistics Cost  │
                            └──────────────────┘  └──────────────────┘
```

Figure 2: Framework for calculation of total cost of ownership

3. Mathematical Models for Estimation of Total Cost of Ownership

In this section we develop mathematical models for estimation of various cost elements in the total cost of ownership. The main focus is on estimation of in-service cost. Since all the cost elements in the total cost of ownership need to be discoun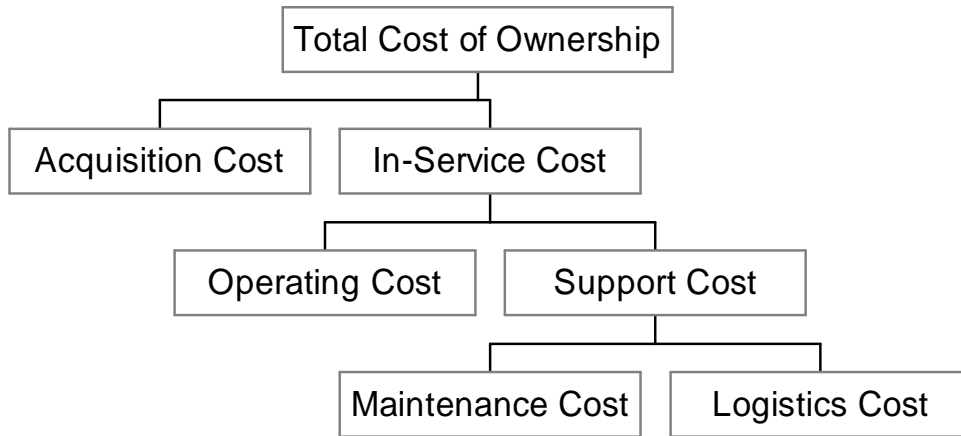ted to their present value, all the costs models explained in the subsequent sections are calculated on annual basis and finally discounted using appropriate discount rate.

3.1 Estimation of operating cost

The operating cost can be divided into two categories, direct operating cost and overhead costs. The direct operating cost is determined by the resources, which are required for operating the asset. The main resources for most of the system are energy consumed by the asset and the manpower required to operate the asset. The energy consumed by the asset will depend on the operational availability of the system calculated on annual basis. The operational availability, A_o, of any asset is given by (D Kumar et al 2000):

$$A_O = \frac{MTBM}{MTBM + DT} \tag{1}$$

where:

MTBM = Mean time between maintenance

DT = Down Time

Mean time between maintenance for duration, T, is given by:

$$MTBM = \frac{T}{M(T) + \dfrac{T}{T_{sm}}}$$ (2)

Where, M(T), is the number of failures resulting in unscheduled maintenance and T_{sm} is the time between scheduled maintenance. The down time, DT, can be estimated using the following equation:

$$DT = \frac{M(T) \times MCMT + \dfrac{T}{T_{sm}} \times MPMT}{M(T) + \dfrac{T}{T_{sm}}}$$ (3)

Where,

MCMT = Mean corrective maintenance time.

MPMT = Mean preventive maintenance time.

The number of failures resulting in unscheduled maintenance can be evaluated using renewal function and is given by:

$$M(T) = F(T) + \int_{0}^{T} M(T - x) f(x) dx$$ (4)

Where, F(T) is the cumulative distribution of the time-to-failure random variable and f(x) is the corresponding probability density function. Equation (4) is valid only when the failed units are replaced or when the repair is as good as new. However, in case of minimal repair and imperfect repair, one may have to use models based on non-homogeneous Poisson process or modified renewal process. For more details, the readers may refer to Ross (2000). If we assume that the energy cost and manpower

cost per unit time is C_{ou} and the annual usage of the asset is T life units. Then the annual operating cost is given by:

$$C_O = A_O \times T \times C_{ou} \qquad (5)$$

Assume that 'r' denotes the discount rate. Then the present value of the operating cost for n^{th} period (n^{th} year), $C_{O,n}$, is given by:

$$C_{O,n} = \frac{A_O \times T \times C_{ou}}{(1+r)^n} \qquad (6)$$

3.2 Estimation of Maintenance Cost

The main components of maintenance cost are corrective maintenance costs, preventive maintenance costs and overhaul costs. Maintenance resources that used in performing that particular maintenance drive these costs. The maintenance cost, C_M, can be estimated using the following equation:

$$C_M = M(T) \times C_{cm} + \frac{T}{T_{sm}} \times C_{pm} + \sum_{i=1}^{k} \delta_{i,n} \times C_{OH,i} \qquad (7)$$

Where:

C_{cm} = Average cost of corrective maintenance.

C_{pm} = Average cost of preventive maintenance

$$\delta_{i,n} = \begin{cases} 1, \text{if overhaul of type i is carried out during period n} \\ 0, \text{otherwise} \end{cases}$$

$C_{OH,i}$ represents the average cost of overhaul of type i. This cost will be added to the maintenance cost, if the type i overhaul is carried out during period n.

The maintenance cost for period 'n' is given by:

$$C_{M,n} = \frac{1}{(1+r)^n}\left(M(T) \times C_{cm} + \frac{T}{T_{sm}} \times C_{pm} + \sum_{i=1}^{k} \delta_{i,n} C_{OH,i} \right) \qquad (8)$$

3.3 Estimation of Logistic Support Cost

Logistic support cost covers the costs associated with maintaining spare parts, maintenance facilities, test equipment and other logistics costs such as transportation costs. The spare parts contribute significant portion of the total support cost. The number of spares stocked also plays a crucial role in the operational availability of the system. Practitioners decide on the number of spare parts to be purchased based on the target fill rate, α (probability that a demand for a particular spare part can be achieved from the available stock). Usually the target fill rate is 85%. Assume that N_s represent the minimum number of spares that should be stocked to achieve a target fill rate α. Then the value of N_s can be calculated using the following equation:

$$\sum_{k=0}^{N_s} \frac{\exp(-\lambda T) \times (\lambda T)^k}{k!} \qquad (9)$$

The above equation is valid only when the time-to-failure distribution follows exponential distribution, where λ is the failure rate. When the time-to-failure distribution is other than exponential, then we need to use renewal function to find the value of N_s to achieve the target availability.

The annual logistics cost for the period n, C_L,n is given by:

$$C_{L,n} = \frac{1}{(1+r)^n} N_s \times C_s \qquad (10)$$

3.4 Total Cost of Ownership

The total cost of ownership is obtained by adding the components given by equations (6), (8) and (10) over the designed life of the asset. If the designed life of the asset is D, then the total cost of ownership, TCO_D, is given by:

$$TCO_D = C_P + \sum_{n=1}^{D} [C_{O,n} + C_{M,n} + C_{L,n}] + C_{MF} \qquad (11)$$

Where C_P is the procurement price of wagon and C_{MF} is the one-time expenses of maintenance and support equipment.

4. Case Study on BOXN wagons used by the Indian Railways

Indian railways are the principal mode of transport for raw material for steel plants, finished steel from steel plants, coal, iron, oil, cement, petroleum products, fertilizer and food grains in India. Indian railways owns more than 5,00,000 wagons. The wagons have a design life of 35 years. The requirement of wagons for future is assessed on the basis of freight traffic projected and the anticipated level of productivity of wagons measured in terms of net ton kilometers (NTKM) per wagon per day likely to be achieved. The wagons are procured from the wagons India limited, which was incorporated in 1974. Wagons India limited supplies about 90% of the wagon requirement and the rest are purchased from the open market. The total cost of ownership is an important issue during the procurement of wagons. The procurement cost of wagon is $40,000.

The BOXN wagons are mainly used for carrying coal and are fitted with CASNUB bogies. CASNUB bogies are the critical subsystem of the wagon. The CASNUB bogie consists of two cast side frames and a floating bolster. The bolster is supported on the side frames through two groups of spring, which also incorporate the load proportional friction damping. The side frames of the CASNUB bogie are connected by a fabricated mild steel spring plank to maintain the bogie square. The salient features CASNUB bogies are shown in Table 1.

Table 1. Salient features of CASNUB bogie

59.1.1.1 Gauge	1676mm
Axle load	20.3t. However all bogies except Casnub 22HS can be upgraded upto 22.9t
Wheel Diameter	1000 mm (New). 956 mm ((New) for retrofitted Casnub 22W
Wheel Base	2000mm
Type of Axle	Casnub- 22W (M)
Bearing	(a) Cylindrical roller bearing axle box in limited no. on Casnub 22W Bogies only. (b) Standard AAR tapered cartridge bearing class 'E' suitable for 152.4x276.4mm (6"x11") narrow jaw.

Distance Between Journal centres	2260mm
Distance between side bearers	1474 mm
Type of side bearers	**59.2 Casnub 22W** **59.3** Roller Type (clearance type) **59.4 Retrofitted Casnub 22W, Casnub 22W(M), 22NLB** Constant contact type (Metal bonded rubber pad, housed inside side bearer housing). Casnub 22HS Spring loaded constant contact type side bearer.
Type of pivot	Casnub 22W IRS Type Top Pivot –RDSO Drg. No. W/BE-601. Bottom Pivot – RDSO Drg. No. W/BE-602 or similar mating profile integrally cast with bolster. **59.5 Casnub –22W(M), 22NL, 22NLB, 22HS** Spherical type RDSO Drg. No. WD-85079-S/2.
59.6 Anti rotation features	Anti rotation lugs have been provide between bogie bolster and sideframe.
59.6.1.1 Type of brake beam	**59.7 Casnub –22W, /22NL, 22NLB and 22HS** Unit type fabricated brake beam supported and guided in the brake beam pockets. Casnub—22W(M) Unit type cast steel brake beam suspended by hangers from side frame brackets.
Suspension details	Long travel helical springs
Elastomeric pads	On all type of bogies except Casnub

22Wsubsequently provided in retrofitment.

The CASNUB bogie assembly consists of the following components:

1. Wheel set with cylindrical roller bearing or wheel set with cartridge bearing.
2. Axle box/adapter, retainer bolt & side frame key assembly.
3. Sideframe with friction wear plates.
4. Bolster with wear liners.
5. Spring plank, fitbolts & rivets.
6. Load bearing springs and snubber springs.
7. Friction shoe wedge.
8. Centre Pivot arrangement comprising of centre pivot, centre pivot bottom, centre pivot pin, centre pivot retainer & locking arrangement.
9. Side bearers.
10. Elastomeric pads.
11. Bogie brake gear.
12. Brake beam.

Reliability and Maintenance of BOXN Wagon

Indian railways classifies the failures into the following three categories:

1. Vital – causing line failure.
2. Essential – causing delay to traffic.
3. Non-essential – causing no disturbance to traffic.

The above classification enables the Indian Railways to focus on vital and essential components and to study their reliability and maintainability in service and take adequate steps to improve their performance by modification or re-design. The following three types of maintenance are practiced for wagons:

1. Preventive maintenance (PM): Preventive maintenance is carried out after every 6000 Km for BOXN wagons (approximately 15 days).
2. Routine overhaul (ROH): Routine overhaul is carried out after every 24 months. During ROH, the bogie is dismantled and the wheels are de-wheeled.
3. Periodic Overhaul (POH): The periodic overhaul is carried out after every 48 months and involves complete overhaul of the wagon. However, the first POH is carried out after 6 years.

For this research, we looked at the most critical components (the components that contribute towards majority of the failures). Table 2 shows the vital components and their time-to-failure distribution along

with the estimated parameters. To maintain the confidentiality of the failure and maintenance data, we have used hypothetical data in the rest of the paper. The objective here is to illustrate the models developed in the paper.

Table 2. Vital components of BOXN wagons and their time-to-failure distribution (λ is the failure rate, η is the scale parameter and β is the shape parameter)

S. No.	Component	Time-to-Failure Distribution	Parameters
1.	Wheel	Weibull	η= 52, 0000 Km, β = 4
2.	Roller Bearing	Weibull	η= 250, 0000 Km, β = 3
3.	Brake Beam	Weibull	η= 160, 0000 Km, β = 4
4.	Brake Shoe	Weibull	η= 140, 0000 Km, β = 3
5.	CBC	Weibull	η= 70, 0000 Km, β = 3.5
6.	Panel Hatch	Weibull	η= 38, 0000 Km, β = 4.2
7.	Air Brake	Weibull	η= 48, 0000 Km, β = 3.5
8.	Wagon Door	Exponential	$\lambda = 6.6 \times 10^{-6}$
9.	Centre Pivot	Weibull	η= 55, 0000 Km, β = 3.8

All critical components except wagon door follow Weibull distribution. The time-to-failure distribution of the wagon door is exponential, since most of the wagon door failures are caused due to mishandling. The time-to-failure of the wagon itself follows exponential distribution with mean time between failures of 16000 Km.

Calculation of the operational availability of Wagon

All the life units are measured in terms of Kilometer and thus the PM, ROH, POH are converted in terms of Km. The preventive maintenance interval is approximately 15 days, that is after every 6000 Km, and during PM, the wagon is out of service for 2 days (that is 800 Km). Whenever, the wagon requires corrective maintenance, it is likely to be out of service for 4 days (that is 1600 Km). The usage of wagon for every month is 12,000 Km. Using these data, the mean time between maintenance in one year, (144,000 Km) is given by:

$$MTBM_{wagon} = \frac{144000}{M(144000) + \dfrac{144000}{6000}} = 4363 \, \text{Km} \tag{12}$$

The Down time is given by:

$$DT_{wagon} = \frac{M(T) \times MCMT + (\frac{T}{T_{sm}}) \times MPMT}{M(T) + \frac{T}{T_{sm}}} = \frac{9 \times 1600 + 24 \times 800}{9 + 24} = 1018 \text{ Km} \quad (13)$$

Using, (12) and (13), we get the operational availability of the wagon as:

$$A_{wagon} = \frac{MTBM_{wagon}}{MTBM_{wagon} + DT_{wagon}} = \frac{4363}{4363 + 1018} = 0.8108 \quad (14)$$

Thus, the operational availability of the wagon is 81.08%.

4.3 Operating cost for Wagon

The operational availability value can be now used to calculate the operating cost of the wagon. For the sake of mathematical simplicity, we calculate the cost of ownership for 6 years from commissioning of the wagon. Assume:

C_{ou} = $1 per Km

Then, the operating cost for first six years, at an interest rate of 6% is given in the following table (table 3):

Table 3: Present value of the operating cost

Year	PV of the operating cost (in $)
1	9178.868
2	8659.309
3	8169.16
4	7706.755
5	7270.523
6	6858.984
Total	**54314.34**

4.4 Maintenance cost for Wagon

The cost of maintenance for six years can be calculated using the equation (8). We make the following assumptions:

C_{cm} = Cost of corrective maintenance = $800

C_{pm} = Cost of preventive maintenance = $1500

$C_{OH,1}$ = Cost of regular overhaul (ROH) = $8000

$C_{OH,2}$ = Cost of periodic overhaul (POH) = $15000

Table 4, shows the present value of the maintenance cost for the first six years.

Table 4: Present value of the maintenance cost

Year	PV of the logistics cost (in $)
1	30849.06
2	36222.86
3	27455.55
4	37782.87
5	24435.34
6	28691.89
Total	**30849.06**

In table 4, one can notice, cost fluctuation during year 2, 4 and 6. This is due to ROH and POH carried out during that period.

4.5 Logistics Support Cost

The logistics support cost can be estimated using equation (10). This involves the use of renewal process to estimate the spares requirement for each of the components shown in table 2. Assuming, N_s x C_s = $ 3000, the present value of the logistics cost for six years is shown in Table 5.

Table 5: Present value of the logistics cost

Year	PV of the logistics cost (in $)
1	2830.189
2	2669.989
3	2518.858
4	2376.281
5	2241.775
6	2114.882
Total	**14751.97**

4.6 Total cost of ownership of wagon

The total cost of ownership of wagon for the first 6 years is obtained by adding the components given by equations (6), (8) and (10). Assume that C_{MF} = $20000. The total cost of ownership, TCO_D, for six years, is given by:

$$TCO_D = C_P + \sum_{n=1}^{D} [C_{O,n} + C_{M,n} + C_{L,n}] + C_{MF}$$
$$= 40000 + 54314.34 + 185437.6 + 14751.97 + 20000$$
$$= \$314502.6$$

The cost of ownership is calculated for six years, which is the first major overhaul period. The above cost can be divided by the duration, to calculate TCO per year, which then can be used for comparing different configurations.

5. Decision Making on the Basis of Cost of Ownership

In this section, we discuss how the cost of ownership derived in the previous section can be used for purchasing decisions. The following two approaches can be used for decision making.

5.1   Decision Making Based on TCO as only Criteria

If the cost of ownership is the only criteria on which the purchasing decision is based on then, the alternative with minimum total cost of ownership per period (per annum). That is, if there are n alternatives such that:

$TCO_{a,i}$ = Total cost of ownership per annum for the ith alternative.

Then, the alternative, *m*, such that:

$$TCO_{a,m} = Min\{TCO_{a,1}, TCO_{a,2}, ..., TCO_{a,n}\}$$ (15)

5.2   Decision Making Based on TCO as one of the criteria

In many situations, the purchasing decision is made using multiple criteria, including TCO as one of the criteria. In such cases, one can use multi criteria decision making techniques to choose the best alternative. In this paper, we suggest analytic hierarchy process (AHP) for choosing the best alternative (Saaty, 1980). AHP is a multi-criteria decision making technique which can be used to choose best alternative among number of alternatives. Let us assume that there are M alternatives and N decision criteria. Let a_{ij} denote the weight of the ith alternative on j^{th} criteria. Let W_j be the weight for criteria j. Then, the decision problem can be defined using the following matrix (Triantaphyllou *et al 1995*)

Criteria					
	1	2	3	...	N
Alternative	W_1	W_2	W_3	...	W_N
1	a_{11}	a_{12}	a_{13}	...	a_{1N}
2	a_{21}	a_{22}	a_{23}	...	a_{2N}
.				...	
.				...	

.				...	
M	a_{M1}	a_{M2}	a_{M3}		a_{MN}

Using the above data AHP finds the overall importance of the alternative and chooses the one with maximum weight (interested reader may refer to Saaty, 1980).

6. Conclusions

The main objectives of this paper are to develop framework for estimation of cost of ownership of capital assets and to develop mathematical models for estimation of various cost elements within the cost of ownership. The innovative approach used in this paper is the use of operational availability to estimate the operating cost. Most of the model use calendar time to estimate the operating time. The use of calendar time may be appropriate for certain elements of operating cost such as labor costs, the variable costs such as energy consumed would depend on the operational availability of the system. The models developed in the paper are used illustrated using the wagons used in the Indian railways. Although, the data used in the paper are modified, it captures the impact of cost of ownership. In the example, we have shown that the cost of ownership of wagon for six years is almost 8 times its initial procurement price. The main aim of this paper is to show the significance and total cost of ownership compared to the procurement price and thus to prove that all procurement decisions must be based on total cost of ownership and not on the basis of procurement cost.

References

Anon. (2002), Procurement of wagons by Indian Railways, Report Number 9A, Indian Railways.

Blanchard, B. S. (1986), *Logistics Engineering and Management, 3rd Edition,* Prentice Hall, New Jersey.

Blanchard, B. S. (1988), *Systems Engineering and Management, 2nd Edition,* John Wiley, New York.

Blanchard, B. S. and Fabrycky, W. J. (1990), *Systems Engineering and Analysis, 2nd Edition,* Prentice Hall, New Jersey.

Dinesh Kumar, U., Crocker, J, Knezevic, J and El-Haram. (2000), *Reliability, Maintenance and Logistics Support – A Life Cycle Approach,* Kluwer Academic Publishers, Boston.

Hampton, T. (2004), The Cost of Ownership, *ENR: Engineering News Record,* Vol. 253, 7, 24-29.

Harvey, G. (1976), Life Cycle Costing: A Review of the Technique, *Management Accounting, 343-347.*

Humphries, J. (2004), Optimizing Total Cost of Ownership, *Plant Engineers,* Vol. 58, 7, 23-27.

Ross, S. (2000), Introduction to Probability Models, *Academic Press,* New York.

Saaty, T. L. (1980), *The Analytic Hierarchy Process: Planning, Priority Setting and Resource Allocation*, McGraw Hill.

Sherif, Y. S., and Kolarik, W. J. (1981), Life Cycle Costing: Concept and Practice, *OMEGA,* 9,.3, 287-296.

Taylor, W.B. (1981), The Use of Life Cycle Costing in Acquiring Physical Asset, *Long Range Planning,* 14, 32-43.

Triantaphyllou, E., and Mann, S. (1995), Using Analytical Hierarchy Process in Decision Making in Engineering Applications: Some Challenges, *International Journal of Industrial Engineering: Applications and Practice,* Vol. 2, No. 1, 35-44