1

Introduction to Reliability Engineering

1.1 What is Reliability Engineering?

No one disputes the need for engineered products to be reliable. The average consumer is acutely aware of the problem of less than perfect reliability in domestic products such as TV sets and automobiles. Organizations such as airlines, the military and public utilities are aware of the costs of unreliability. Manufacturers often suffer high costs of failure under warranty. Argument and misunderstanding begin when we try to quantify reliability values, or try to assign financial or other cost or benefit values to levels of reliability.

The simplest, purely producer-oriented or inspectors' view of quality is that in which a product is assessed against a specification or set of attributes, and when passed is delivered to the customer. The customer, having accepted the product, accepts that it might fail at some future time. This simple approach is often coupled with a warranty, or the customer may have some protection in law, so that he may claim redress for failures occurring within a stated or reasonable time. However, this approach provides no measure of quality over a period of time, particularly outside a warranty period. Even within a warranty period, the customer usually has no grounds for further action if the product fails once, twice or several times, provided that the manufacturer repairs the product as promised each time. If it fails often, the manufacturer will suffer high warranty costs, and the customers will suffer inconvenience. Outside the warranty period, only the customer suffers. In any case, the manufacturer will also probably incur a loss of reputation, possibly affecting future business.

We therefore come to the need for a time-based concept of quality. The inspectors' concept is not time-dependent. The product either passes a given test or it fails. On the other hand, reliability is usually concerned with failures in the time domain. This distinction marks the difference between traditional quality control and reliability engineering.

Whether failures occur or not, and their times to occurrence, can seldom be forecast accurately. Reliability is therefore an aspect of engineering uncertainty. Whether an item will work for a particular period is a question which can be answered as a probability. This results in the usual engineering definition of reliability as:

> *The probability that an item will perform a required function without failure under stated conditions for a stated period of time*.

Reliability can also be expressed as the number of failures over a period.

Durability is a particular aspect of reliability, related to the ability of an item to withstand the effects of time (or of distance travelled, operating cycles, etc.) dependent mechanisms such as fatigue, wear, corrosion,

Practical Reliability Engineering, Fifth Edition. Patrick D. T. O'Connor and Andre Kleyner.

electrical parameter change, and so on. Durability is usually expressed as a minimum time before the occurrence of *wearout* failures. In repairable systems it often characterizes the ability of the product to function after repairs.

The objectives of reliability engineering, in the order of priority, are:

1 To apply engineering knowledge and specialist techniques to prevent or to reduce the likelihood or frequency of failures.
2 To identify and correct the causes of failures that do occur, despite the efforts to prevent them.
3 To determine ways of coping with failures that do occur, if their causes have not been corrected.
4 To apply methods for estimating the likely reliability of new designs, and for analysing reliability data.

The reason for the priority emphasis is that it is by far the most effective way of working, in terms of minimizing costs and generating reliable products.

The primary skills that are required, therefore, are the ability to understand and anticipate the possible causes of failures, and knowledge of how to prevent them. It is also necessary to have knowledge of the methods that can be used for analysing designs and data. The primary skills are nothing more than good engineering knowledge and experience, so reliability engineering is first and foremost the application of good engineering, in the widest sense, during design, development, manufacture and service.

Mathematical and statistical methods can be used for quantifying reliability (prediction, measurement) and for analysing reliability data. The basic methods are described in Chapter 2, to provide an introduction for some of the applications described subsequently. However, because of the high levels of uncertainty involved these can seldom be applied with the kind of precision and credibility that engineers are accustomed to when dealing with most other problems. In practice the uncertainty is often in orders of magnitude. Therefore the role of mathematical and statistical methods in reliability engineering is limited, and appreciation of the uncertainty is important in order to minimize the chances of performing inappropriate analysis and of generating misleading results. Mathematical and statistical methods can make valuable contributions in appropriate circumstances, but practical engineering must take precedence in determining the causes of problems and their solutions. Unfortunately not all reliability training, literature and practice reflect this reality.

Over-riding all of these aspects, though, is the management of the reliability engineering effort. Since reliability (and very often also safety) is such a critical parameter of most modern engineering products, and since failures are generated primarily by the people involved (designers, test engineers, manufacturing, suppliers, maintainers, users), it can be maximized only by an integrated effort that encompasses training, teamwork, discipline, and application of the most appropriate methods. Reliability engineering "specialists" cannot make this happen. They can provide support, training and tools, but only managers can organize, motivate, lead and provide the resources. Reliability engineering is, ultimately, effective management of engineering.

1.2 Why Teach Reliability Engineering?

Engineering education is traditionally concerned with teaching how manufactured products work. The ways in which products fail, the effects of failure and aspects of design, manufacture, maintenance and use which affect the likelihood of failure are not usually taught[1], mainly because it is necessary to understand how a

[1] Mechanical engineering curricula normally include basic failure processes such as fracture mechanics, wear and corrosion.

product works before considering ways in which it might fail. For many products the tendency to approach the failed state is analogous to entropy. The engineer's tasks are to design and maintain the product so that the failed state is deferred. In these tasks he faces the problems inherent in the variability of engineering materials, processes and applications. Engineering education is basically deterministic, and does not usually pay sufficient attention to variability. Yet variability and chance play a vital role in determining the reliability of most products. Basic parameters like mass, dimensions, friction coefficients, strengths and stresses are never absolute, but are in practice subject to variability due to process and materials variations, human factors and applications. Some parameters also vary with time. Understanding the laws of chance and the causes and effects of variability is therefore necessary for the creation of reliable products and for the solution of problems of unreliability.

However, there are practical problems in applying statistical knowledge to engineering problems. These problems have probably deterred engineers in the past from using statistical methods, and texts on reliability engineering and mathematics have generally stressed the theoretical aspects without providing guidance on their practical application. To be helpful a theoretical basis must be credible, and statistical methods which work well for insurance actuaries, market researchers or agricultural experimenters may not work as well for engineers. This is not because the theory is wrong, but because engineers usually have to cope with much greater degrees of uncertainty, mainly due to human factors in production and use.

Some highly reliable products are produced by design and manufacturing teams who practise the traditional virtues of reliance on experience and maintenance of high quality. They do not see reliability engineering as a subject requiring specialist consideration, and a book such as this would teach them little that they did not already practise in creating their reliable products. Engineers and managers might therefore regard a specialist reliability discipline with scepticism. However, many pressures now challenge the effectiveness of the traditional approaches. Competition, the pressure of schedules and deadlines, the cost of failures, the rapid evolution of new materials, methods and complex systems, the need to reduce product costs, and safety considerations all increase the risks of product development. Figure 1.1 shows the pressures that lead to the overall perception of risk. Reliability engineering has developed in response to the need to control these risks.

Later chapters will show how reliability engineering methods can be applied to design, development, manufacturing and maintenance to control the level of risk. The extent to which the methods are applicable must be decided for each project and for each design area. They must not replace normal good practice, such as safe design for components subject to cyclic loading, or application guidelines for electronic components.

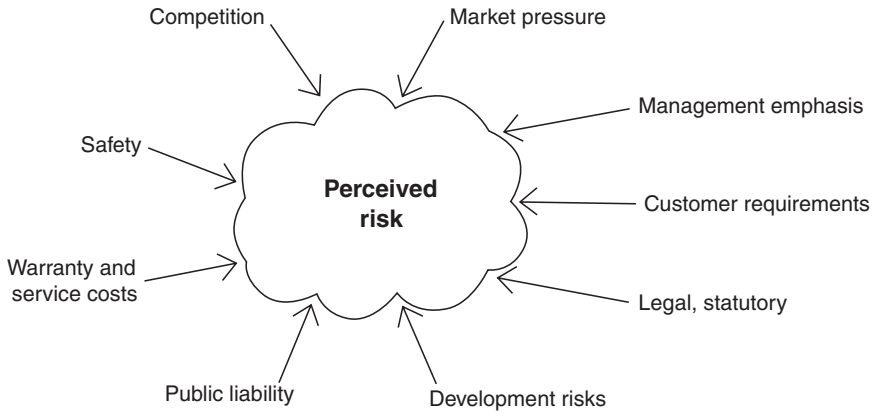

Figure 1.1 *Perception of risk.*

They should be used to supplement good practice. However, there are times when new risks are being taken, and the normal rules and guidelines are inadequate or do not apply. Sometimes we take risks unwittingly, when we assume that we can extrapolate safely from our present knowledge. Designers and managers are often overoptimistic or are reluctant to point out risks about which they are unsure.

It is for these reasons that an understanding of reliability engineering principles and methods is now an essential ingredient of modern engineering.

1.3 Why Do Engineering Products Fail?

There are many reasons why a product might fail. *Knowing, as far as is practicable, the potential causes of failures is fundamental to preventing them*. It is rarely practicable to anticipate all of the causes, so it is also necessary to take account of the uncertainty involved. The reliability engineering effort, during design, development and in manufacture and service should address all of the anticipated and possibly unanticipated causes of failure, to ensure that their occurrence is prevented or minimized.

The main reasons why failures occur are:

1 The design might be *inherently incapable*. It might be too weak, consume too much power, suffer resonance at the wrong frequency, and so on. The list of possible reasons is endless, and every design problem presents the potential for errors, omissions, and oversights. The more complex the design or difficult the problems to be overcome, the greater is this potential.
2 The item might be *overstressed* in some way. If the stress applied exceeds the strength then failure will occur. An electronic component will fail if the applied electrical stress (voltage, current) exceeds the ability to withstand it, and a mechanical strut will buckle if the compression stress applied exceeds the buckling strength. Overstress failures such as these do happen, but fortunately not very often, since designers provide margins of safety. Electronic component specifications state the maximum rated conditions of application, and circuit designers take care that these rated values are not exceeded in service. In most cases they will in fact do what they can to ensure that the in-service worst case stresses remain below the rated stress values: this is called 'de-rating'. Mechanical designers work in the same way: they know the properties of the materials being used (e.g. ultimate tensile strength) and they ensure that there is an adequate margin between the strength of the component and the maximum applied stress. However, it might not be possible to provide protection against every possible stress application.
3 Failures might be caused by *variation*. In the situations described above the values of strength and load are fixed and known. If the known strength always exceeds the known load, as shown in Figure 1.2, then failure will not occur. However, in most cases, there will be some uncertainty about both. The actual strength values of any population of components will vary: there will be some that are relatively strong, others that are relatively weak, but most will be of nearly average strength. Also, the loads applied will be variable. Figure 1.3 shows this type of situation. As before, failure will not occur so long as the applied load does not exceed the strength. However, if there is an overlap between the distributions of load and strength, and a load value in the high tail of the load distribution is applied to an item in the weak tail of the strength distribution so that there is overlap or *interference* between the distributions (Figure 1.4), then failure will occur. We will discuss load and strength interference in more detail in Chapter 5.
4 Failures can be caused by *wearout*. We will use this term to include any mechanism or process that causes an item that is sufficiently strong at the start of its life to become weaker with age. Well-known examples of such processes are material fatigue, wear between surfaces in moving contact, corrosion, insulation deterioration, and the wearout mechanisms of light bulbs and fluorescent tubes. Figure 1.5 illustrates this kind of situation. Initially the strength is adequate to withstand the applied loads, but as weakening occurs

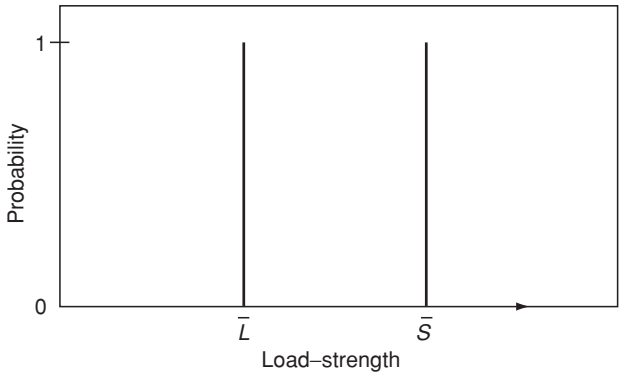

Figure 1.2 *Load–strength – discrete values.*

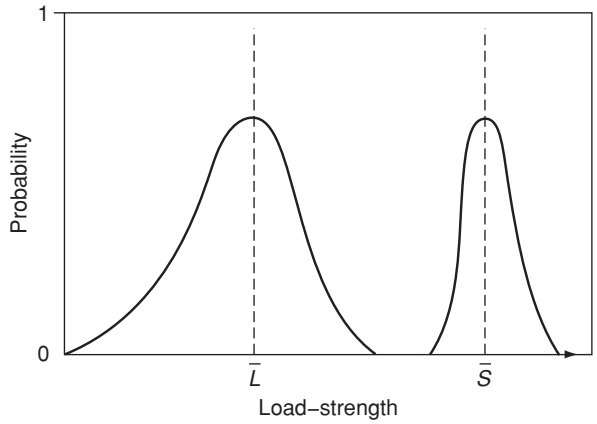

Figure 1.3 *Load–strength – distributed values.*

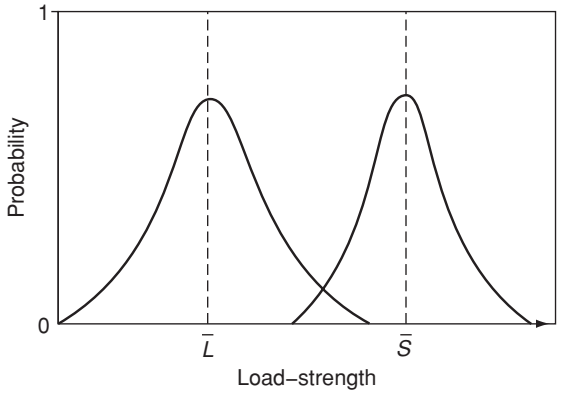

Figure 1.4 *Load–strength – interfering distributions.*

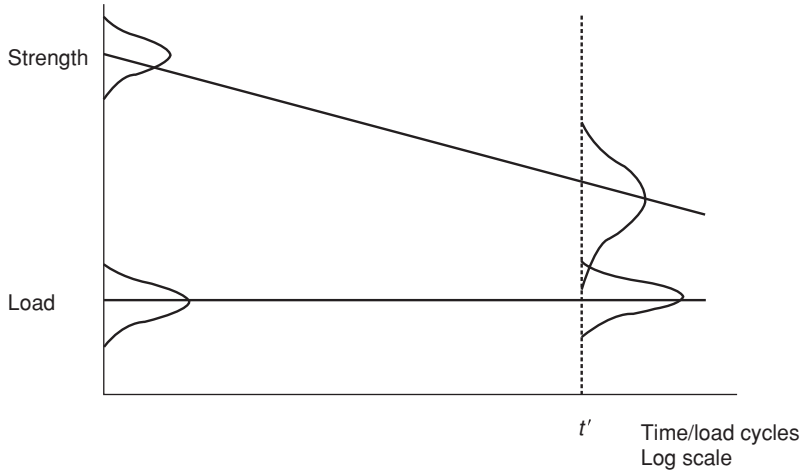

Figure 1.5 *Time-dependent load and strength variation.*

over time the strength decreases. In every case the average value falls and the spread of the strength distribution widens. This is a major reason why it is so difficult to provide accurate predictions of the lives of such items.

5 Failures can be caused by other time-dependent mechanisms. Battery run-down, creep caused by simultaneous high temperature and tensile stress, as in turbine discs and fine solder joints, and progressive drift of electronic component parameter values are examples of such mechanisms.

6 Failures can be caused by *sneaks*. A sneak is a condition in which the system does not work properly even though every part does. For example, an electronic system might be designed in such a way that under certain conditions incorrect operation occurs. The fatal fire in the Apollo spacecraft crew capsule was caused in this way: the circuit design ensured that an electrical short circuit would occur when a particular sequence was performed by the crew. Sneaks can also occur in software designs.

7 Failures can be caused by *errors*, such as incorrect specifications, designs or software coding, by faulty assembly or test, by inadequate or incorrect maintenance, or by incorrect use. The actual failure mechanisms that result might include most of the list above.

8 There are many other potential causes of failure. Gears might be noisy, oil seals might leak, display screens might flicker, operating instructions might be wrong or ambiguous, electronic systems might suffer from electromagnetic interference, and so on.

Failures have many different causes and effects, and there are also different perceptions of what kinds of events might be classified as failures. The burning O-ring seals on the Space Shuttle booster rockets were not classed as failures, until the ill-fated launch of Challenger. We also know that all failures, in principle and almost always in practice, can be prevented.

1.4 Probabilistic Reliability

The concept of reliability as a probability means that any attempt to quantify it must involve the use of statistical methods. An understanding of statistics as applicable to reliability engineering is therefore a necessary basis for progress, except for the special cases when reliability is perfect (we know the item will

never fail) or it is zero (the item will never work). In engineering we try to ensure 100 % reliability, but our experience tells us that we do not always succeed. Therefore reliability statistics are usually concerned with probability values which are very high (or very low: the probability that a failure does occur, which is 1 – reliability). Quantifying such numbers brings increased uncertainty, since we need correspondingly more information. Other sources of uncertainty are introduced because reliability is often about people who make and people who use the product, and because of the widely varying environments in which typical products might operate.

Further uncertainty, often of a subjective nature, is introduced when engineers begin to discuss failures. Should a failure be counted if it was due to an error that is hoped will not be repeated? If design action is taken to reduce the risk of one type of failure, how can we quantify our trust in the designer's success? Was the machine under test typical of the population of machines?

Reliability is quantified in other ways. We can specify a reliability as the mean number of failures in a given time (failure rate), or as the *mean time between failures* (MTBF) for items which are repaired and returned to use, or as the *mean time to failure* (MTTF) for items which are not repaired, or as the proportion of the total population of items failing during the mission life.

The application and interpretation of statistics to deal with the effects of variability on reliability are less straightforward than in, say, public opinion polls or measurement of human variations such as IQ or height. In these applications, most interest is centred around the behaviour of the larger part of the population or sample, variation is not very large and data are plentiful. In reliability we are concerned with the behaviour in the extreme tails of distributions and possibly unlikely combinations of load and strength, where variability is often hard to quantify and data are expensive.

Further difficulties arise in the application of statistical theory to reliability engineering, owing to the fact that variation is often a function of time or of time-related factors such as operating cycles, diurnal or seasonal cycles, maintenance periods, and so on. Engineering, unlike most fields of knowledge, is primarily concerned with change, hopefully, but not always, for the better. Therefore the reliability data from any past situation cannot be used to make credible forecasts of the future behaviour, without taking into account non-statistical factors such as design changes, maintainer training, and even imponderables such as unforeseeable production or service problems. The statistician working in reliability engineering needs to be aware of these realities.

Chapter 2 provides the statistical basis of reliability engineering, but it must always be remembered that quality and reliability data contain many sources of uncertainty and variability which cannot be rigorously quantified. It is also important to appreciate that failures and their causes are by no means always clear-cut and unambiguous. They are often open to interpretation and argument. They also differ in terms of importance (cost, safety, other effects). Therefore we must be careful not to apply only conventional scientific, deterministic thinking to the interpretation of failures. For example, a mere count of total reported failures of a product is seldom useful or revealing. It tells us nothing about causes or consequences, and therefore nothing about how to improve the situation. This contrasts with a statement of a physical attribute such as weight or power consumption, which is usually unambiguous and complete. Nevertheless, it is necessary to derive values for decision-making, so the mathematics are essential. The important point is that the reliability engineer or manager is not, like an insurance actuary, a powerless observer of his statistics. Statistical derivations of reliability are not a guarantee of results, and these results can be significantly affected by actions taken by quality and reliability engineers and managers.

1.5 Repairable and Non-Repairable Items

It is important to distinguish between repairable and non-repairable items when predicting or measuring reliability.

For a non-repairable item such as a light bulb, a transistor, a rocket motor or an unmanned spacecraft, reliability is the survival probability over the item's expected life, or for a period during its life, *when only one failure can occur*. During the item's life the instantaneous probability of the first and only failure is called the *hazard rate*. Life values such as the mean life or *mean time to failure* (MTTF), or the expected life by which a certain percentage might have failed (say 10 %.) (*percentile life*), are other reliability characteristics that can be used. Note that non-repairable items may be individual parts (light bulbs, transistors, fasteners) or systems comprised of many parts (spacecraft, microprocessors).

For items which are repaired when they fail, reliability is the probability that failure will not occur in the period of interest, *when more than one failure can occur*. It can also be expressed as the *rate of occurrence of failures* (ROCOF), which is sometimes referred as the *failure rate* (usually denoted as λ). However, the term failure rate has wider meaning and is often applied to both repairable and non-repairable systems expressing the number of failures per unit time, as applied to one unit in the population, when one or more failures can occur in a time continuum. It is also sometimes used as an averaged value or practical metric for the hazard rate.

Repairable system reliability can also be characterized by the *mean time between failures* (MTBF), but only under the particular condition of a constant failure rate. It is often assumed that failures do occur at a constant rate, in which case the failure rate $\lambda = (\text{MTBF})^{-1}$. However, this is only a special case, valuable because it is often true and because it is easy to understand.

We are also concerned with the *availability* of repairable items, since repair takes time. Availability is affected by the rate of occurrence of failures (failure rate) and by maintenance time. Maintenance can be corrective (i.e. repair) or preventive (to reduce the likelihood of failure, e.g. lubrication). We therefore need to understand the relationship between reliability and maintenance, and how both reliability and maintainability can affect availability.

Sometimes an item may be considered as both repairable and non-repairable. For example, a missile is a repairable system whilst it is in store and subjected to scheduled tests, but it becomes a non-repairable system when it is launched. Reliability analysis of such systems must take account of these separate states. Repairability might also be determined by other considerations. For example, whether an electronic circuit board is treated as a repairable item or not will depend upon the cost of repair. An engine or a vehicle might be treated as repairable only up to a certain age.

Repairable system reliability data analysis is covered in Chapter 13 and availability and maintainability in Chapter 16.

1.6 The Pattern of Failures with Time (Non-Repairable Items)

There are three basic ways in which the pattern of failures can change with time. The hazard rate may be decreasing, increasing or constant. We can tell much about the causes of failure and about the reliability of the item by appreciating the way the hazard rate behaves in time.

Decreasing hazard rates are observed in items which become less likely to fail as their survival time increases. This is often observed in electronic equipment and parts. 'Burn-in' of electronic parts is a good example of the way in which knowledge of a decreasing hazard rate is used to generate an improvement in reliability. The parts are operated under failure-provoking stress conditions for a time before delivery. As substandard parts fail and are rejected the hazard rate decreases and the surviving population is more reliable.

A constant hazard rate is characteristic of failures which are caused by the application of loads in excess of the design strength, at a constant average rate. For example, overstress failures due to accidental or transient circuit overload, or maintenance-induced failures of mechanical equipment, typically occur randomly and at a generally constant rate.

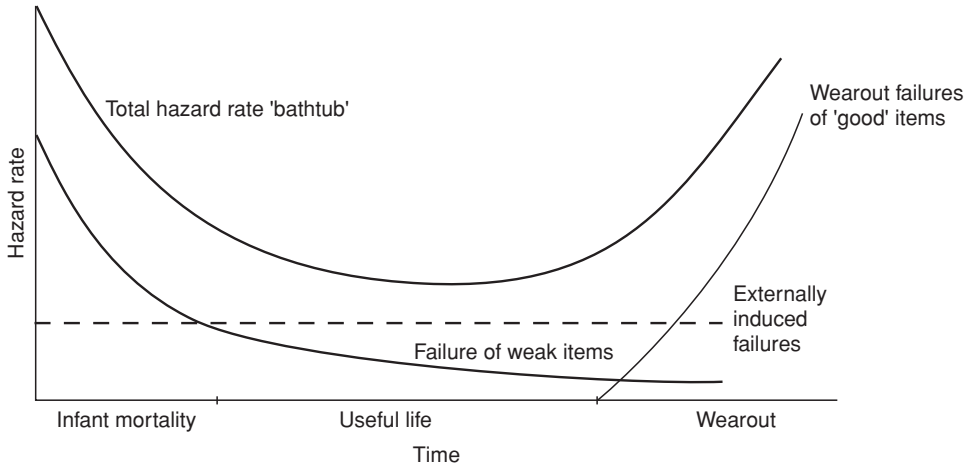

Figure 1.6 *The 'bathtub' curve.*

Wearout failure modes follow an increasing hazard rate. For example, material fatigue brought about by strength deterioration due to cyclic loading is a failure mode which does not occur for a finite time, and then exhibits an increasing probability of occurrence.

The combined effect generates the so-called *bathtub curve* (Figure 1.6). This shows an initial decreasing hazard rate or *infant mortality* period, an intermediate *useful life* period and a final *wearout* period. Death is a good analogy to failure of a non-repairable system, and the bathtub curve model is similar to actuarial statistical models.

1.7 The Pattern of Failures with Time (Repairable Items)

The failure rates (or ROCOF) of repairable items can also vary with time, and important implications can be derived from these trends.

A constant failure rate (CFR) is indicative of externally induced failures, as in the constant hazard rate situation for non-repairable items. A CFR is also typical of complex systems subject to repair and overhaul, where different parts exhibit different patterns of failure with time and parts have different ages since repair or replacement. Repairable systems can show a decreasing failure rate (DFR) when reliability is improved by progressive repair, as defective parts which fail relatively early are replaced by good parts. 'Burn in' is applied to electronic systems, as well as to parts, for this purpose.

An increasing failure rate (IFR) occurs in repairable systems when wearout failure modes of parts begin to predominate.

The pattern of failures with time of repairable systems can also be illustrated by use of the bathtub curve (Figure 1.6), but with the failure rate (ROCOF) plotted against age instead of the hazard rate.

The statistical treatment of failure data is covered in Chapters 2 and 3.

1.8 The Development of Reliability Engineering

Reliability engineering, as a separate engineering discipline, originated in the United States during the 1950s. The increasing complexity of military electronic systems was generating failure rates which resulted in

greatly reduced availability and increased costs. Solid state electronics technology offered long term hope, but conversely miniaturization was to lead to proportionately greater complexity, which offset the reliability improvements expected. The gathering pace of electronic device technology meant that the developers of new military systems were making increasing use of large numbers of new component types, involving new manufacturing processes, with the inevitable consequences of low reliability. The users of such equipment were also finding that the problems of diagnosing and repairing the new complex equipment were seriously affecting its availability for use, and the costs of spares, training and other logistics support were becoming excessive. Against this background the US Department of Defense and the electronics industry jointly set up the Advisory Group on Reliability of Electronic Equipment (AGREE) in 1952. The AGREE report concluded that, to break out of the spiral of increasing development and ownership costs due to low reliability, disciplines must be laid down as integral activities in the development cycle for electronic equipment. The report laid particular stress on the need for new equipments to be tested for several thousand hours in high stress cyclical environments including high and low temperatures, vibration and switching, in order to discover the majority of weak areas in a design at an early enough stage to enable them to be corrected before production commenced. Until that time, environmental tests of tens of hours' duration had been considered adequate to prove the suitability of a design. The report also recommended that formal demonstrations of reliability, in terms of statistical confidence that a specified MTBF had been exceeded, be instituted as a condition for acceptance of equipment by the procuring agency. A large part of the report was devoted to providing detailed test plans for various levels of statistical confidence and environmental conditions.

The AGREE report was accepted by the Department of Defense, and AGREE testing quickly became a standard procedure. Companies which invested in the expensive environmental test equipment necessary soon found that they could attain levels of reliability far higher than by traditional methods. It was evident that designers, particularly those working at the fringes of advanced technology, could not be expected to produce highly reliable equipment without it being subjected to a test regime which would show up weaknesses. Complex systems and the components used in them included too many variables and interactions for the human designer to cope with infallibly, and even the most careful design reviews and disciplines could not provide sufficient protection. Consequently it was necessary to make the product speak for itself, by causing it to fail, and then to eliminate the weaknesses that caused the failures. The Department of Defense (DOD) reissued the AGREE report on testing as US Military Standard (MIL-STD) 781, *Reliability Qualification and Production Approval Tests*.

Meanwhile the revolution in electronic device technology continued, led by integrated micro circuitry. Increased emphasis was now placed on improving the quality of devices fitted to production equipments. Screening techniques, in which all production devices are subjected to elevated thermal, electrical and other stresses, were introduced in place of the traditional sampling techniques. With component populations on even single printed circuit boards becoming so large, sampling no longer provided sufficient protection against the production of defective equipment. These techniques were formalized in military standards covering the full range of electronic components. Components produced to these standards were called 'Hi-rel' components. Specifications and test systems for electronic components, based upon the US Military Standards, were developed in the United Kingdom and in continental Europe, and internationally through the International Electrotechnical Commission (IEC).

However, improved quality standards in the electronic components industry resulted in dramatic improvements in the reliability of commercial components. As a result, during the 1980s the US Military began switching from military grade electronic components to "commercial off the shelf" (COTS) parts in order to reduce costs, and this approach has spread to other applications.

Engineering reliability effort in the United States developed quickly, and the AGREE and reliability programme concepts were adopted by NASA and many other major suppliers and purchasers of high technology equipment. In 1965 the DOD issued MIL-STD-785–*Reliability Programs for Systems and Equipment*. This document made mandatory the integration of a programme of reliability engineering activities with the

traditional engineering activities of design, development and production, as it was by then realized that such an integrated programme was the only way to ensure that potential reliability problems would be detected and eliminated at the earliest, and therefore the cheapest, stage in the development cycle. Much written work appeared on the cost-benefit of higher reliability, to show that effort and resources expended during early development and during production testing, plus the imposition of demonstrations of specified levels of reliability to MIL-STD-781, led to reductions in in-service costs which more than repaid the reliability programme expenditure. The concept of life cycle costs (LCC), or whole life costs, was introduced.

In the United Kingdom, Defence Standard 00–40, *The Management of Reliability and Maintainability* was issued in 1981. The British Standards Institution issued BS 5760 – *Guide on Reliability of Systems, Equipments and Components*. In the 1990s the series of European Reliability/Dependability[2] standards began to be developed, and became integrated into the International Standards Organization (ISO). For example ISO/IEC 60 300 describes the concepts and principles of dependability management systems. It identifies the generic processes for planning, resource allocation, control, and tailoring necessary to meet dependability objectives. At present, there is a large number of ISO standards regulating testing, validation, reliability analysis, and various other aspects of product development.

Starting in the early 1980s, the reliability of new Japanese industrial and commercial products took Western competitors by surprise. Products such as automobiles, electronic components and systems, and machine tools achieved levels of reliability far in excess of previous experience. These products were also less expensive and often boasted superior features and performance. The 'Japanese quality revolution' had been driven by the lessons taught by American teachers brought in to help Japan's post-war recovery. The two that stand out were J.R. Juran and W. Edwards Deming, who taught the principles of 'total quality management' (TQM) and continuous improvement. Japanese pioneers, particularly K. Ishikawa, also contributed. These ideas were all firmly rooted in the teaching of the American writer on management, Peter Drucker (Drucker, 1995), that people work most effectively when they are given the knowledge and authority to identify and implement improvements, rather than being expected to work to methods dictated by 'management'.

These ideas led to great increases in productivity and quality, and thus in reliability and market penetration, as Drucker had predicted. Many Western companies followed the new path that had been blazed and also made great improvements. The message is now almost universally applied, particularly with the trend to international outsourcing of manufacturing.

The Western approach had been based primarily on formal procedures for design analysis and reliability demonstration testing, whereas the Japanese concentrated on manufacturing quality. Nowadays most customers for systems such as military, telecommunications, transport, power generation and distribution, and so on, rely upon contractual motivation, such as warranties and service support, rather than on imposition of standards that dictate exactly how reliability activities should be performed.

Another aspect of reliability thinking that has developed is the application of statistical methods, primarily to the analysis of failure data and to predictions of reliability and safety of systems. Since reliability can be expressed as a probability, and is affected by variation, in principle these methods are applicable. They form the basis of most teaching and literature on the subject. However, variation in engineering is usually of such an uncertain nature that refined mathematical and quantitative techniques can be inappropriate and misleading. This aspect will be discussed in later chapters.

1.9 Courses, Conferences and Literature

Reliability engineering and management are now taught in engineering courses at a large number of universities, colleges and polytechnics, and on specialist short courses.

[2] In this context dependability is defined as including reliability, maintainability, availability and safety.

Conferences on general and specific reliability engineering and management topics have been held regularly in the United States since the 1960s and in Europe and elsewhere since the 1970s. The best known is the annual US Reliability and Maintainability Symposium (RAMS), sponsored by most of the important engineering associations and institutions in the United States. It is held every year and its conference proceedings contain much useful information and are often cited. The European Safety and Reliability Conference (ESREL) is also held annually and publishes proceedings on a variety of reliability topics, and conferences take place in other countries.

Journals on reliability have also appeared; some are referenced at the end of this chapter. Several books have been published on the subjects of reliability engineering and management; some of these are referenced at the end of other chapters.

Much of the reliability literature has tended to emphasize the mathematical and analytical aspects of the subject, with the result that reliability engineering is often considered by designers and others to be a rather esoteric subject. This is unfortunate, since it creates barriers to communication. It is important to emphasize the more practical aspects and to integrate reliability work into the overall management and engineering process. These aspects are covered in later chapters.

1.10 Organizations Involved in Reliability Work

Several organizations have been created to develop policies and methods in reliability engineering and to undertake research and training. Amid those organizations it is important to mention ASQ (American Society for Quality), which became a truly international organization with members in almost every country in the world. ASQ has many internal organizations including the Reliability Division which is the worldwide professional group with the focus on reliability specific training, education, networking and best practices.

1.11 Reliability as an Effectiveness Parameter

With the increasing cost and complexity of many modern systems, the importance of reliability as an effectiveness parameter, which should be specified and paid for, has become apparent. For example, a radar station, a process plant or an airliner must be available when required, and the cost of non-availability, particularly if it is unscheduled, can be very high. In the weapons field, if an anti-aircraft missile has a less than 100 % probability of functioning correctly throughout its engagement sequence, operational planners must consider deploying the appropriate extra quantity to provide the required level of defence. The Apollo project second stage rocket was powered by six rocket motors; any five would have provided sufficient impulse, but an additional motor was specified to cater for a possible failure of one. As it happened there were no failures, and every launch utilized an 'unnecessary' motor. These considerations apply equally to less complex systems, such as vending and copying machines, even if the failure costs are less dramatic in absolute terms.

As an effectiveness parameter, reliability can be 'traded off' against other parameters. Reliability generally affects availability, and in this context maintainability is also relevant. Reliability and maintainability are often related to availability by the formula:

$$\text{Availability} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}$$

where MTTR is the mean time to repair. This is the simplest steady-state situation. It is clear that availability improvements can be achieved by improving either MTBF or MTTR. For example, automatic built-in test

equipment can greatly reduce diagnostic times for electronic equipment, at a cost of a slight reduction in overall reliability and an increase in unit costs. Many other parameters can be considered in trade-offs, such as weight, redundancy, cost of materials, parts and processes, or reduction in performance.

The greatest difficulty in estimating relationships for reliability trade-offs derives from the fact that, whereas it is possible to estimate quite accurately such factors as the cost and weight penalties of built-in test equipment, the cost of materials and components, or the worth of a measurable performance parameter, the effect on reliability cannot generally be forecast accurately, and reliability measurements can at best be made only within statistical limits imposed by the amount of data available. Selection of trade-offs must therefore be very much a matter of experience of similar projects in the knowledge that wide margins of error can exist.

1.12 Reliability Programme Activities

What, then, are the actions that managers and engineers can take to influence reliability? One obvious activity already mentioned is quality assurance (QA), the whole range of functions designed to ensure that delivered products are compliant with the design. For many products, QA is sufficient to ensure high reliability, and we would not expect a company mass-producing simple diecastings for non-critical applications to employ reliability staff. In such cases the designs are simple and well proven, the environments in which the products will operate are well understood and the very occasional failure has no significant financial or operational effect. QA, together with craftsmanship, can provide adequate assurance for simple products or when the risks are known to be very low. Risks are low when safety margins can be made very large, as in most structural engineering. Reliability engineering disciplines may justifiably be absent in many types of product development and manufacture. QA disciplines are, however, essential elements of any integrated reliability programme.

A formal reliability programme is necessary whenever the risks or costs of failure are not low. We have already seen how reliability engineering developed as a result of the high costs of unreliability of military equipment, and later in commercial applications. Risks of failure usually increase in proportion to the number of components in a system, so reliability programmes are required for any product whose complexity leads to an appreciable risk.

An effective reliability programme should be based on the conventional wisdom of responsibility and authority being vested in one person. Let us call him or her the reliability programme manager. The responsibility must relate to a defined objective, which may be a maximum warranty cost figure, an MTBF to be demonstrated or a requirement that failure will not occur. Having an objective and the authority, how does the reliability programme manager set about his or her task, faced as he or she is with a responsibility based on uncertainties? This question will be answered in detail in subsequent chapters, but a brief outline is given below.

The reliability programme must begin at the earliest, conceptual phase of the project. It is at this stage that fundamental decisions are made, which can significantly affect reliability. These are decisions related to the risks involved in the specification (performance, complexity, cost, producibility, etc.), development time-scale, resources applied to evaluation and test, skills available, and other factors.

The shorter the project time-scale, the more important is this need, particularly if there will be few opportunities for an iterative approach. The activities appropriate to this phase are an involvement in the assessment of these trade-offs and the generation of reliability objectives. The reliability staff can perform these functions effectively only if they are competent to contribute to the give-and-take inherent in the trade-off negotiations, which may be conducted between designers and staff from manufacturing, marketing, finance, support and customer representatives.

As the project proceeds from initial study to detail design, the reliability risks are controlled by a formal, documented approach to the review of design and to the imposition of design rules relating to selection of components, materials and processes, stress protection, tolerancing, and so on. The objectives at this stage are to ensure that known good practices are applied, that deviations are detected and corrected, and that areas of uncertainty are highlighted for further action. The programme continues through the initial hardware manufacturing and test stages, by planning and executing tests to show up design weaknesses and to demonstrate achievement of specified requirements and by collecting, analysing and acting upon test and failure data. During production, QA activities ensure that the proven design is repeated, and further testing may be applied to eliminate weak items and to maintain confidence. The data collection, analysis and action process continues through the production and in-use phases. Throughout the product life cycle, therefore, the reliability is assessed, first by initial predictions based upon past experience in order to determine feasibility and to set objectives, then by refining the predictions as detail design proceeds and subsequently by recording performance during the test, production and in-use phases. This performance is fed back to generate corrective action, and to provide data and guidelines for future products.

The elements of a reliability programme are outlined in documents such as US MIL-STD-785, UK Defence Standard 00–40 and British Standard 5760 (see Bibliography). The activities are described fully in subsequent chapters.

1.13 Reliability Economics and Management

Obviously the reliability programme activities described can be expensive. Figure 1.7 is a commonly-described representation of the theoretical cost–benefit relationship of effort expended on reliability (and production quality) activities. It shows a U-shaped total cost curve with the minimum cost occurring at a reliability level somewhat lower than 100 %. This would be the optimum reliability, from the total cost point of view.

W.E. Deming presented a different model in his teaching on manufacturing quality (Deming, 1986). This is shown in Figure 1.8. He argued that, since less than perfect quality is the result of failures, all of which

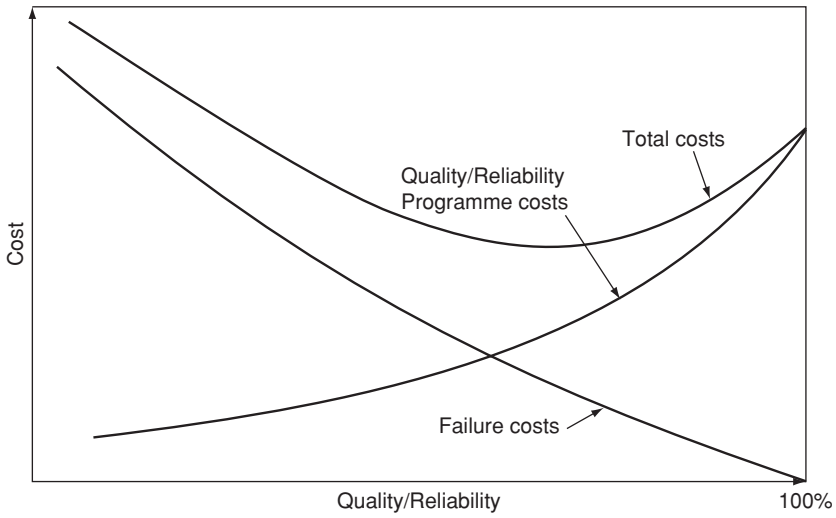

Figure 1.7 *Reliability and life cycle costs (traditional view).*

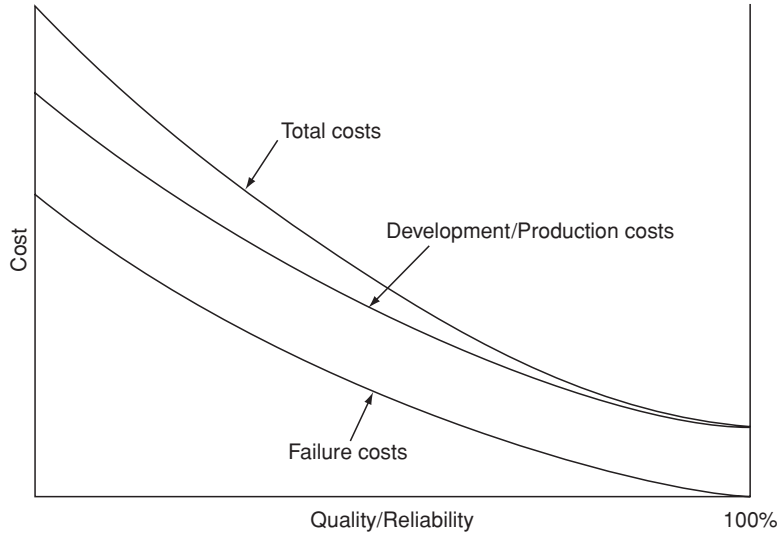

Figure 1.8 *Reliability/Quality and life cycle costs (Based on Deming's quality vs. cost model).*

have causes, we should not be tempted to assume that any level of quality is "optimum", but should ask 'what is the cost of preventing or correcting the causes, on a case by case basis, compared with the cost of doing nothing?' When each potential or actual cause is analysed in this way, it is usually found that it costs less to correct the causes than to do nothing. Thus total costs continue to reduce as quality is improved. This simple picture was the prime determinant of the post-war quality revolution in Japan, and formed the basis for the philosophy of *kaizen* (continuous improvement). 100 % quality was rarely achieved, but the levels that were achieved exceeded those of most Western competitors, and production costs were reduced.

In principle, the same argument applies to reliability: all efforts to improve reliability by identifying and removing potential causes of failures in service should result in cost savings later in the product life cycle, giving a net benefit in the longer term. In other words, an effective reliability programme represents an investment, usually with a large payback over a period of time. Unfortunately it is not easy to quantify the effects of given reliability programme activities, such as additional design analysis or testing, on achieved reliability. The costs (including those related to the effects on project schedules) of the activities are known, and they arise in the short term, but the benefits arise later and are often much less certain. However, achieving levels of reliability close to 100 % is often not realistic for complex products. Recent research on reliability cost modeling (Kleyner, 2010) showed that in practical applications the total cost curve is highly skewed to the right due to the increasing cost and diminishing return on further reliability improvements, as shown in Figure 1.9. The tight timescales and budgets of modern product development can also impact the amount of effort that can be applied. On the other hand there is often strong market pressure to achieve near perfect reliability. See more on cost of reliability in Chapters 14 and 17.

It is important to remember though that while achieving 100 % quality in manufacturing operations, or 100 % reliability in service, is extremely rare in real life applications, especially in high volume production, it should nevertheless be considered as an ultimate goal for any product development and production programme.

Achieving reliable designs and products requires a totally integrated approach, including design, test, production, as well as the reliability programme activities. The integrated engineering approach places high requirements for judgment and engineering knowledge on project managers and team members. Reliability specialists must play their parts as members of the team.

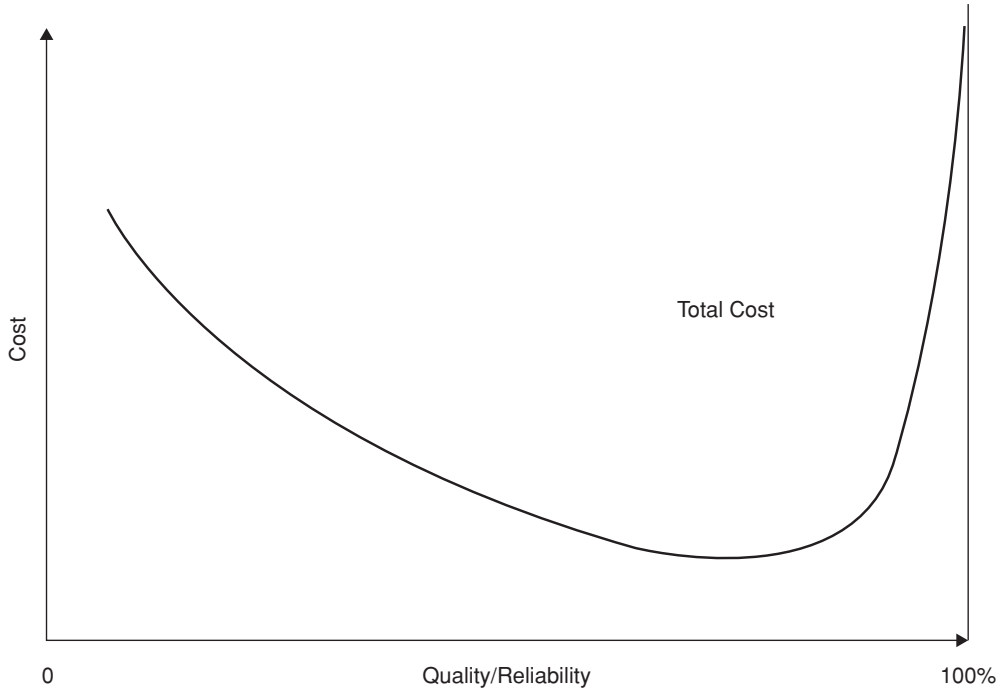

Figure 1.9 *Reliability and life cycle costs (practical applications).*

There are three kinds of engineering product, from the perspective of failure prevention:

1 *Intrinsically reliable components*, which are those that have high margins between their strength and the stresses that could cause failure, and which do not wear out within their practicable life times. Such items include nearly all electronic components (if properly applied), nearly all mechanical non-moving components, and all correct software.
2 *Intrinsically unreliable components*, which are those with low design margins or which wear out, such as badly applied components, light bulbs, turbine blades, parts that move in contact with others, like gears, bearings and power drive belts, and so on.
3 *Systems* which include many components and interfaces, like cars, dishwashers, aircraft, and so on, so that there are many possibilities for failures to occur, particularly across interfaces (e.g. inadequate electrical overstress protection, vibration nodes at weak points, electromagnetic interference, software that contains errors, and so on).

It is the task of design engineers to ensure that all components are correctly applied, that margins are adequate (particularly in relation to the possible extreme values of strength and stress, which are often variable), that wearout failure modes are prevented during the expected life (by safe life design, maintenance, etc.), and that system interfaces cannot lead to failure (due to interactions, tolerance mismatches, etc.). Because achieving all this on any modern engineering product is a task that challenges the capabilities of the very best engineering teams, it is almost certain that aspects of the initial design will fall short of the 'intrinsically reliable' criterion. Therefore we must submit the design to analyses and tests in order to show

not only that it works, but also to show up the features that might lead to failures. When we find out what these are we must redesign and re-test, until the final design is considered to meet the criterion.

Then the product has to be manufactured. In principle, every one should be identical and correctly made. Of course this is not achievable, because of the inherent variability of all manufacturing processes, whether performed by humans or by machines. It is the task of the manufacturing people to understand and control variation, and to implement inspections and tests that will identify non-conforming product.

For many engineering products the quality of operation and maintenance also influence reliability.

The essential points that arise from this brief and obvious discussion of failures are that:

1 Failures are caused primarily by people (designers, suppliers, assemblers, users, maintainers). Therefore the achievement of reliability is essentially a management task, to ensure that the right people, skills, teams and other resources are applied to prevent the creation of failures.
2 Reliability (and quality) specialists cannot by themselves effectively ensure the prevention of failures. High reliability and quality can be achieved only by effective team working by all involved.
3 There is no fundamental limit to the extent to which failures can be prevented. We can design and build for ever-increasing reliability.

Deming explained how, in the context of manufacturing quality, there is no point at which further improvement leads to higher costs. This is, of course, even more powerfully true when considered over the whole product life cycle, so that efforts to ensure that designs are intrinsically reliable, by good design, thorough analysis and effective development testing, can generate even higher pay-offs than improvements in production quality. The '*kaizen*' (continuous improvement) principle is even more effective when applied to up-front engineering.

The creation of reliable products is, therefore, primarily a management task. Guidance on reliability programme management and costs is covered in Chapter 17.

Questions

1. Define (a) failure rate, and (b) hazard rate. Explain their application to the reliability of components and repairable systems. Discuss the plausibility of the 'bathtub curve' in both contexts.
2. a Explain the theory of component failures derived from the interaction of stress (or load) and strength distributions. Explain how this theory relates to the behaviour of the component hazard function.
   b Discuss the validity of the 'bathtub curve' when used to describe the failure characteristics of non-repairable components.
3. What are the main objectives of a reliability engineering team working on an engineering development project? Describe the important skills and experience that should be available within the team.
4. Briefly list the most common basic causes of failures of engineering products.
5. It is sometimes claimed that increasing quality and reliability beyond levels that have been achieved in the past is likely to be uneconomic, due to the costs of the actions that would be necessary. Present the argument against this belief. Illustrate it with an example from your own experience.
6. Describe the difference between repairable and non-repairable items. What kind of effect might this difference have on reliability? List examples of repairable and non-repairable items in your everyday life.
7. Explain the difference between reliability and durability and how they can be specified in a product development programme.

8. a List the potential economic outcomes of poor reliability, and identify which costs are directly quantifiable and which are intangible. Explain how they can be minimised, and discuss the extent to which very high reliability (approaching zero failures) is achievable in practice.
   b What are the major factors that might limit the achievement of very high reliability?
9. After processing the existing programme cost data and running a regression model on the previous projects, the cost of product development and manufacturing (CDM) has been estimated to follow the equation: CDM = \$ 0.8 million +\$ 3.83 million $\times$ R^2 (R is the achieved product reliability at service life and is expected to be above 90 %). The cost of failure (CF) has been estimated as the sum of fixed cost of \$ 40 000 plus variable cost of \$ 150 per failure. The total number of the expected failures is $n \times (1 - R)$, where n is the total number of produced units. Considering that the production volume is expected to be 50 000 units, estimate the optimal target reliability and the total cost of the programme.
10. Select an everyday item (coffee maker, lawnmower, bicycle, mobile phone, CD player, computer, refrigerator, microwave oven, cooking stove, etc.).
    a Discuss the ways this item can potentially fail. What can be done to prevent those failures?
    b Based on the Figures 1.3 and 1.4, what would be an example of the load and strength for a critical component within this item? Do you expect load and strength for this component to be time-dependent?

Bibliography

British Standard, BS 4778 (1991) *Glossary of Terms Used in Quality Assurance* (including reliability and maintainability). British Standards Institution, London.

British Standard, BS 5760 (1996) *Reliability of Systems, Equipments and Components*. British Standards Institution, London.

Deming, W. (1986) *Out of the Crisis*, MIT University Press (originally published under the title *Quality, Productivity and Competitive Position*).

Drucker, P. (1995) *The Practice of Management*. Heinemann.

Kleyner, A. (2010) *Determining Optimal Reliability Targets*, Lambert Academic Publishing.

Misra, K. (ed.) (2008) *The Handbook of Performability Engineering*, Springer-Verlag, London.

UK Defence Standard 00–40. *The Management of Reliability and Maintainability*. HMSO.

US MIL-STD-721. *Definitions of Effectiveness Terms for Reliability, Maintainability, Human Factors and Safety*. National Technical Information Service, Springfield, Virginia.

US MIL-STD-785. *Reliability Programs for Systems and Equipment*. National Technical Information Service, Springfield, Virginia (suspended in 1976).

Periodic Publications

International Journal of Performability Engineering. Available at: http://www.ijpe-online.com/

Microelectronics Reliability, Elsevier (published monthly).

Proceedings of the US Reliability and Maintainability Symposium (RAMS). American Society for Quality and IEEE (published annually).

Quality and Reliability Engineering International, Wiley (published quarterly).

Reliability Engineering and Systems Safety, Elsevier (published monthly).

Transactions on Reliability, Institute of Electrical and Electronics Engineers (IEEE) (published quarterly).