PART A

Overview

CHAPTER 1

Overview

1.1 INTRODUCTION

In the real world, problems arise in many different contexts. Problem solving is an activity that has a history as old as the human race. Models have played an important role in problem solving and can be traced back to well beyond the recorded history of the human race. Many different kinds of models have been used. These include physical (full or scaled) models, pictorial models, analog models, descriptive models, symbolic models, and mathematical models. The use of mathematical models is relatively recent (roughly the last 500 years). Initially, mathematical models were used for solving problems from the physical sciences (e.g., predicting motion of planets, timing of high and low tides), but, over the last few hundred years, mathematical models have been used extensively in solving problems from biological and social sciences. There is hardly any discipline where mathematical models have not been used for solving problems.

Two different approaches to building mathematical models are as follows:

1. *Theory-Based Modeling*. Here, the modeling is based on the established theories (from physical, biological, and social sciences) relevant to the problem. This kind of model is also called *physics-based model* or *white-box model* as the underlying mechanisms form the starting point for the model building.
2. *Empirical Modeling*. Here, the data available forms the basis for the model building, and it does not require an understanding of the underlying mechanisms involved. As such, these models are used when there is insufficient understanding to use the earlier approach. This kind of model is also called *data-dependent model* or *black-box model.*

In empirical modeling, the type of mathematical formulations needed for modeling is dictated by a preliminary analysis of data available. If the analysis indicates

Weibull Models, by D.N.P. Murthy, Min Xie, and Renyan Jiang.
ISBN 0-471-36092-9

that there is a high degree of variability, then one needs to use models that can capture this variability. This requires probabilistic and stochastic models to model a given data set.

Effective empirical modeling requires good understanding of (i) the methodology needed for model building, (ii) properties of different models, and (iii) tools and techniques to determine if a particular model is appropriate to model a given data set.

A variety of such models have been developed and studied extensively. One such class of models is the *Weibull models*. These are a collection of probabilistic and stochastic models derived from the Weibull distribution. These can be divided into univariate and multivariate models and each, in turn, can be further subdivided into continuous and discrete. Weibull models have been used in many different applications to model complex data sets.

1.1.1 Aims of the Book

This book deals with Weibull models and their applications in reliability. The aims of the book are as follows:

1. Develop a taxonomy to integrate the different Weibull models.
2. Review the literature for each model to summarize model properties and other issues.
3. Discuss the use of Weibull probability paper (WPP) plots in model selection. It allows the model builder to determine whether one or more of the Weibull models are suitable for modeling a given data set.
4. Highlight issues that need further study.
5. Illustrate the application of Weibull models in reliability theory.

The book provides a good foundation for empirical model building involving Weibull models. As such, it should be of interest to practitioners from many different disciplines. The book should also be of interest to researchers as some topics for future research are defined as part of the exercises at the end of several chapters.

1.1.2 Outline of Chapter

The outline of the chapter is as follows. We start with a collection of real-world problems in Section 1.2 and discuss the data aspects and empirical models to obtain solutions to the problems. Section 1.3 deals with the modeling methodology, and we discuss the different issues involved. We highlight the role of statistics, probability theory, and stochastic processes in the context of the link between data and model. Section 1.4 starts a brief historical perspective and then introduces the standard Weibull model (involving the two-parameter Weibull distribution). Following this, a taxonomy to classify the different Weibull models is briefly discussed. Given a univariate continuous data set, a question of great interest to model builders is whether one of the Weibull models is suitable for modeling the given data set or not. This topic is discussed in Section 1.5. Section 1.6 deals with the applications of Weibull models where we start with a short list of applications to highlight the diverse range of applications of the Weibull models in different disciplines. However, in this book we focus on the application of Weibull models in the context of product reliability from a product life perspective. We discuss this briefly so as to set the scene for the discussion on Weibull model applications later in the book. We finally conclude with an outline of the book in Section 1.7.

1.2 ILLUSTRATIVE PROBLEMS

In this section we give a few illustrative problems and the types of data available to build models to obtain solutions to the problems.

Example 1: Tidal Heights
At a popular tourist beach the cyclone season precedes the tourist season. Very high tides during the cyclone season cause the erosion of sand on the beach. The erosion is related to the amplitude of the high tide, and it takes a long time for the beach to recover naturally from the effect of such erosion. Often, sand needs to be pumped to restore the loss and to ensure high tourist numbers. A problem of interest to the city council responsible for the beach is the probability that a high tide during the cyclone season exceeds some specified height resulting in the council incurring the sand pumping cost. The data available is the amplitude of high tides over several years.

Example 2: Efficacy of Treatment
In medical science, a problem of interest is in determining the efficacy of a new treatment to control the spread of a disease (e.g., cancer). In this case, clinical trials are carried out for a certain period. The data available are the number entering the program, the time instants, and the age at death for the patients who died during the trial period, ages of the patients who survived the test period, and so on. Similar data for a sample not given the new treatment might also be available. The problem is to determine if the new treatment increases the life expectancy of the patients.

Example 3: Strength of Components
Due to manufacturing variability, the strength of a component varies significantly. The component is used in an environment where it fails immediately when put into use if its strength is below some specified value. The problem is to determine the probability that a component manufactured will fail under a given environment. If this probability is high, changing the material, the process of manufacturing, or redesigning might be the alternatives that the manufacturer might need to explore. The data available is the laboratory test data. Here items are subjected to increasing levels of stress and the stress level at failure being recorded.

Example 4: Insurance Claims
Whenever there is a legitimate claim, a car insurance company has to pay out. The pay out indicates a high degree of variability (since it can vary from a small to a

very large amount). The insurance company has used the expected value as the basis for determining the annual premium it should charge its customers. It is planning to change the premium and is interested in assessing the probability of an individual claim exceeding five times the premium charged. The data available is the insurance claims over the last few years.

Example 5: Growth of Trees

Paper manufacturing requires wood chips. One way of producing wood chips is through plantations where trees are harvested when the trees reach a certain age. The height of the tree at the time of the harvesting is critical as the volume of wood chips obtained is related to this height. The heights of trees vary significantly. As a result, the output of a plantation can vary significantly, and this has an impact on the profitability of the operation. The operator of a plantation is faced with the problem of choosing between two different types of trees. The data available (from other plantations) are the heights of trees at the time of harvesting for both species.

Example 6: Maintenance of Street Lights

The life of electric bulbs used for street lighting is uncertain and is influenced by a variety of factors (variability in the material used and in the manufacturing process, fluctuations in the voltage, etc.). Replacement of an individual failed item is in general expensive. In this case the road authority might decide on some preventive maintenance action where the bulbs are replaced by new ones at set time instants $t = kT, k = 1, 2, \ldots$. The cost of replacing a bulb under such a replacement policy is much cheaper, but it involves discarding the remaining useful life of the bulb. Any failure in between results in the failed item being replaced by a new one at a much higher cost. The problem facing the authority is to determine the optimal T that minimizes the expected cost. The data available is the historical record of failures and preventive replacements in the past.

Example 7: Stress on Offshore Platform

An offshore platform must be designed to withstand the buffeting of waves. The impact of each wave on the structure is determined by the energy contained in the wave. The wave height is an indicator of the energy in a wave. The data available are the heights of successive waves over a certain time interval, and this exhibits a high degree of variability. The problem is to determine the risk of an offshore platform collapsing if designed to withstand waves up to a certain height.

Example 8: Wind Velocity

Windmills are structures that harness the energy in the wind and convert it into electrical or mechanical energy. The wind velocity fluctuates, and as a result the output of the windmill fluctuates. The economic viability of a windmill is dependent on it being capable of generating a certain minimum level of output for a specified fraction of the day. The problem is to determine the viability of windmills based on the data for wind speeds measured every 5 minutes over a week.

Example 9: Rock Blasting

Mining involves blasting ore formation using explosives. The effect of explosion is that it fragments the ore into different sizes. Ore smaller than the minimum acceptable size is of no value as it are unsuited for processing. Ore lumps bigger than the maximum acceptable size need to be broken down, which involves additional cost. The problem of interest to a mine operator is to determine the size distribution of ore under different blasting strategies so as to decide on the best blasting strategy. In this case, the data available are the size distribution of ore randomly sampled after a blast.

Example 10: Spare Part Planning

For commercial equipment (e.g., aircraft, locomotive) downtime implies a loss of revenue. Downtime occurs due to failure of one or more components of the equipment. Failure of a component is dependent on the reliability of the component. The downtime is dependent on whether a spare is available or not and the time to get a spare if one is not available. When the component is expensive, one must manage the inventory of spare parts properly. Carrying a large inventory implies too much capital being tied up. On the other hand, having a small inventory can lead to high downtimes. The problem is to determine the optimal spare part inventory for components. The data available are the failure times for the different components over a certain period of time.

1.3 EMPIRICAL MODELING METHODOLOGY

The empirical modeling process involves the following five steps:

Step 1: Collecting data
Step 2: Analysis of data
Step 3: Model selection
Step 4: Parameter estimation
Step 5: Model validation

In this section we briefly discuss each of these steps.

Step 1: Collecting Data

Data can be either laboratory data or field data. Laboratory data is often obtained under controlled environment and based on a properly planned experiment. In contrast, field data suffers from variability in the operation environment as well as other uncontrollable factors.

The form of data can vary. In the case of reliability data, it could be continuous valued (e.g., life of an individual item) or discrete valued (e.g., number of items failing in a specified interval). In the former case, it could represent failure times or censored times (the lives of nonfailed items when data collection was stopped) for items. We shall discuss this issue in greater detail in Chapter 4.

Finally, when the data needed for modeling is not available, one needs to collect data based on a proper experiment on expert judgment in some cases. The experiment, in general, is discipline specific. We will discuss this issue in the context of product reliability later in the book.

Step 2: Preliminary Analysis of Data

Given a data set, one starts with a preliminary analysis of the data. Suppose that the data set available is given by $(t_1, t_2, \ldots, t_n)$. In the first stage, one computes various sample statistics (such as max, min, mean, sample variance, median, and first and third quartiles) based on the data. If the range (= max – min) is small relative to the sample mean, one might ignore the variability in the data and model the data by the sample mean. However, when this is not the case, then the model needs to mimic this variability in the data. In the case of time-ordered data, preliminary analysis is used to determine properties such as trends (increasing or decreasing), correlation over time, and so forth.

The main aim of the analysis is to assist in determining whether a particular model is appropriate or not to model a given data set. Many different plots have been developed to assist in this. Some of these plots (e.g., histogram) are general and others (e.g., Weibull probability paper plot) were originally developed for a particular model but have since been used for a broader class of models.

Step 3: Model Selection

Suppose that the data set $(t_1, t_2, \ldots, t_n)$ exhibits significant variability. In this case the data set needs to be viewed as an observed value of a set of random variables $(T_1, T_2, \ldots, T_n)$. If the random variables are statistically independent, then each T can be modeled by a univariate probability distribution function:

$$F(t; \theta) = P(T \leq t) \qquad -\infty < t < \infty \tag{1.1}$$

where θ denotes the set of parameters for the distribution. In some cases the range of t is constrained. For example, if T represents the lifetime of an item, then it is constrained to be nonnegative so that $F(t; \theta)$ is zero for $t < 0$.

Model selection involves choosing an appropriate model formulation (e.g., a distribution function) to model a given data set. In order to execute this step, one needs to have a good understanding of the properties of different model formulations suitable for modeling. Some basic concepts are discussed in Chapter 3. Probability theory deals with such study for a variety of model formulations. An important feature of modeling is that often there is more than one model formulation that will adequately model a given data set. In other words, one can have multiple models for a given data set.

The data source often provides a clue to the selection of an appropriate model. In the case of failure data, for example, lognormal or Weibull distributions have been used for modeling failures due to fatigue and exponential distributions for failure of electronic components. In order to use this knowledge, the model builder must be familiar with earlier models for failures of different items.

If the data are not independent, one needs to use models involving multivariate distribution functions. If time is a factor that needs to be included in the model explicitly, then the model becomes more complex. The building of such models requires concepts from stochastic processes.

Step 4: Parameter Estimation
Once a model is selected, one needs to estimate the model parameters. The estimates are obtained using the data available. A variety of techniques have been developed, and these can be broadly divided into two categories—graphical and analytical. The accuracy of the estimate is dependent on the size of the data and the method used. Graphical methods yield crude estimates while analytical methods yield better estimates and confidence limits for the estimates. The basic concepts are discussed in in Chapter 4 and in later chapters in the context of specific models.

Step 5: Model Validation
One can always fit a model to a given data set. However, the model might not be appropriate or adequate. An inappropriate model, in general, will not yield the desired solution to the problem. Hence, it is necessary to check the validity of the model selected. There are several methods for doing this. The basic concepts are discussed in Chapter 5 and in later chapters in the context of specific models.

Comments

1. Steps 2, 4, and 5 deal with *statistical inference*. In probability theory, one models the uncertainty (randomness) through a distribution function, and then makes statements, based on the model, about the nature (e.g., variability) of the data that may result if the model is correct. The principal objective of statistical inference is to use data to make statements about the model, either in terms of probability distribution itself or in terms of its parameters or some other characteristics. Thus, probability theory and statistical inference may be thought of as inverse of one another as indicated:

$$\begin{aligned} \text{Probability theory:} &\quad \text{Model} \rightarrow \text{Data} \\ \text{Statistics:} &\quad \text{Data} \rightarrow \text{Model} \end{aligned}$$

2. Statistical inference requires concepts, tools, and techniques from the *theory of statistics*. Understanding a model requires studying the properties of the model. This requires concepts, tools, and techniques from the *theory of probability* and the *theory of stochastic processes.*
3. In this book we discuss both model properties and statistical inference for Weibull models.

1.4 WEIBULL MODELS

1.4.1 Historical Perspective

The three-parameter Weibull distribution is given by the distribution function

$$F(t;\theta) = 1 - \exp\left[-\left(\frac{t-\tau}{\alpha}\right)^{\beta}\right] \qquad t \geq \tau \tag{1.2}$$

The parameters of the distribution are given by the set $\theta = \{\alpha, \beta, \tau\}$ with $\alpha > 0, \beta > 0$, and $\tau \geq 0$. The parameters α, β, and τ are the *scale, shape*, and *location parameters* of the distribution, respectively. The distribution is named after Waloddi Weibull who was the first to promote the usefulness of this to model data sets of widely differing character. The initial study by Weibull (Weibull, 1939) appeared in a Scandinavian journal and dealt with the strength of materials. A subsequent study in English (Weibull, 1951) was a landmark work in which he modeled data sets from many different disciplines and promoted the versatility of the model in terms of its applications in different disciplines.

A similar model was proposed earlier by Rosen and Rammler (1933) in the context of modeling the variability in the diameter of powder particles being greater than a specific size. The earliest known publication dealing with the Weibull distribution is a work by Fisher and Tippet (1928) where this distribution is obtained as the limiting distribution of the smallest extremes in a sample. Gumbel (1958) refers to the Weibull distribution as the third asymptotic distribution of the smallest extremes.

Although Weibull was not the first person to propose the distribution, he was instrumental in its promotion as a useful and versatile model with a wide range of applicability. A report by Weibull (Weibull, 1977) lists over 1000 references to the applications of the basic Weibull model, and a recent search of various databases indicate that this has increased by a factor of 3 to 4 over the last 30 years.

1.4.2 Taxonomy

The two-parameter Weibull distribution is a special case of (1.2) with $\tau = 0$ so that

$$F(t;\theta) = 1 - \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right] \qquad t \geq 0 \tag{1.3}$$

We shall refer to this as the *standard* Weibull model with $\alpha(>0)$ and $\beta(>0)$ being the scale and shape parameters respectively. The model can be written in alternate parametric forms as indicated below:

$$F(t;\theta) = 1 - \exp\left[-(\lambda t)^{\beta}\right] \tag{1.4}$$

with $\lambda = 1/\alpha$;

$$F(t;\theta) = 1 - \exp\left(-\frac{t^{\beta}}{\alpha'}\right) \tag{1.5}$$

with $\alpha' = \alpha^{\beta}$; and

$$F(t;\theta) = 1 - \exp(-\lambda' t^{\beta}) \tag{1.6}$$

with $\lambda' = (1/\alpha)^{\beta}$. Although they are all equivalent, depending on the context a particular parametric representation might be more appropriate. In the remainder of the book, the form for the standard Weibull model is (1.3) unless indicated otherwise.

A variety of models have evolved from this standard model. We propose a taxonomy for classifying these models, and it involves seven major categories labeled Types I to VII. In this section, we briefly discuss the basis for the taxonomy, and the different models in each category are discussed in Chapter 2.

Let T denote the random variable from the standard Weibull model. Let the distribution function for the derived model be $G(t;\theta)$, and let Z denote the random variable from this distribution. The links between the standard Weibull model and the seven different categories of Weibull models are as follows:

Type I Models Here Z and T are related by a transformation. The transformation can be either (i) linear or (ii) nonlinear.

Type II Models Here $G(t,\theta)$ is related to $F(t,\theta)$ through some functional relationship.

Type III Models These are univariate models derived from two or more distributions with one or more being a standard Weibull distribution. As a result, $G(t,\theta)$ is a univariate distribution function involving one or more standard Weibull distributions.

Type IV Models The parameters of the standard Weibull model are constant. For models belonging to this group, this is not the case. As a result, they are either a function of the variable t or some other variables (such as stress level) or are random variables.

Type V Models In the standard Weibull model, the variable t is continuous valued and can assume any value in the interval $[0,\infty)$. As a result, T is a continuous random variable. In contrast, for Type V models Z can only assume nonnegative integer values, and this defines the support for $G(t,\theta)$.

Type VI Models The standard Weibull model is a univariate model. Type VI models are multivariate extensions of the standard Weibull model. As a result, $G(\cdot)$ is a multivariate function of the form $G(t_1,t_2,\ldots,t_n)$ and related to the standard Weibull in some manner.

Type VII Models These are stochastic point process models with links to the standard Weibull model.

1.5 WEIBULL MODEL SELECTION

Model selection tends to be a trial-and-error process. For Types I to III models the Weibull probability paper plot provides a systematic procedure to determine whether one of these models is suitable for modeling a given data set or not. It is based on the Weibull transformations

$$y = \ln\{-\ln[1 - F(t)]\} \quad \text{and} \quad x = \ln(t) \tag{1.7}$$

A plot of y versus x is called the Weibull probability plot. In the early 1970s a special paper was developed for plotting the data under this transformation and was referred to as the Weibull probability paper (WPP) and the plot called the WPP plot. These days, most reliability software packages contain programs to produce these plots automatically given a data set. We use the term WPP plot to denote the plot using computer packages.

1.6 APPLICATIONS OF WEIBULL MODELS

Weibull models have been used in many different applications and for solving a variety of problems from many different disciplines. Table 1.1 gives a small sample of the application of Weibull models along with references where interested readers can find more details.

1.6.1 Reliability Applications

All man-made systems (ranging from simple products to complex systems) are unreliable in the sense that they degrade with time and/or usage and ultimately fail. The following material is from Blischke and Murthy (2000).

The *reliability* of a product (system) is the probability that the product (system) will perform its intended function for a specified time period when operating under normal (or stated) environmental conditions.

Product Life Cycle and Reliability

A product life cycle (for a consumer durable or an industrial product), from the point of view of the manufacturer, is the time from initial concept of the product to its withdrawal from the marketplace. It involves several stages as indicated in Figure 1.1.

The process begins with an idea to build a product to meet some customer requirements regarding performance (including reliability) targets. This is usually based on a study of the market and the potential demand for the product being planned. The next step is to carry out a feasibility study. This involves evaluating whether it is possible to achieve the targets within the specified cost limits. If this analysis indicates that the project is feasible, an initial product design is undertaken. A prototype is then developed and tested. It is not unusual at this stage to find that achieved performance level of the prototype product is below the target value. In this case, further product development is undertaken to overcome the problem. Once this is achieved, the next step is to carry out trials to determine performance of the product in the field and to start a preproduction run. This is required because

Table 1.1 Sample of Weibull Model Applications

Application	Reference
Yield strength of steel	Weibull (1951)
Size distribution of fly ash	Weibull (1951)
Fiber strength of Indian cotton	Weibull (1951)
Fatigue life of ST-37 steel	Weibull (1951)
Tensile strength of optical fibers	Phani (1987)
Fineness of coal	Rosen and Rammler (1933)
Size of droplets in sprays	Fraser and Eisenklam (1956)
Particle size	Fang et al. (1993)
Pitting corrosion in pipe	Sheikh et al. (1990)
Time to death following exposure to carcinogens	Pike (1966), Peto and Lee (1973)
Strength of fibers from coconut husks	Kulkarni et al. (1973)
Forecasting technology change	Sharif and Islam (1980)
Dielectric breakdown voltage	Nossier et al. (1980)
	Mu et al. (2000)
	Wang et al. (1997)
Fracture strength of glass	Keshvan et al. (1980)
Size of Antarctic icebergs	Neshyba (1980)
Inventory lead time	Tadikamalla (1978)
Wave heights in English Channel	Henderson and Weber (1978)
Size of rock fragments	Rad and Olson (1974)
Ball bearing failures	Lieblein and Zelen (1956)
Failure of carbon fiber composites	Durham and Padgett (1997),
	Padgett et al. (1995)
Machining center failures	Yazhou et al. (1995)
Traffic conflict in expressway merging	Chin et al. (1991)
Partial discharge phenomena	Cacciari et al. (1995)
Flight load variation in helicopter	Boorla and Rotenberger (1997)
Precipitation in Pacific Northwest	Duan et al. (1998)
Adhesive wear in metals	Quereshi and Sheikh (1997)
Fracture in concrete	Xu and Barr (1995)
Latent failures of electronic products	Yang et al. (1995)
Damage in laminated composites	Kwon and Berner (1994)
Software reliability growth	Yamada et al. (1993)
Brittle material	Fok et al. (2001)
Thermoluminescence glow	Pagonis et al. (2001)
Flood frequency	Heo et al. (2001)
Temperature fluctuations	Talkner et al. (2000)
Wind speed distribution	Seguro and Lambert (2000)
	Lun and Lam (2000)
Earthquakes	Huillet and Raynaud (1999)
Failure of coatings	Almeida (1999)
Rain drop size	Jiang et al. (1997)
Discharge inference	Contin et al. (1994)

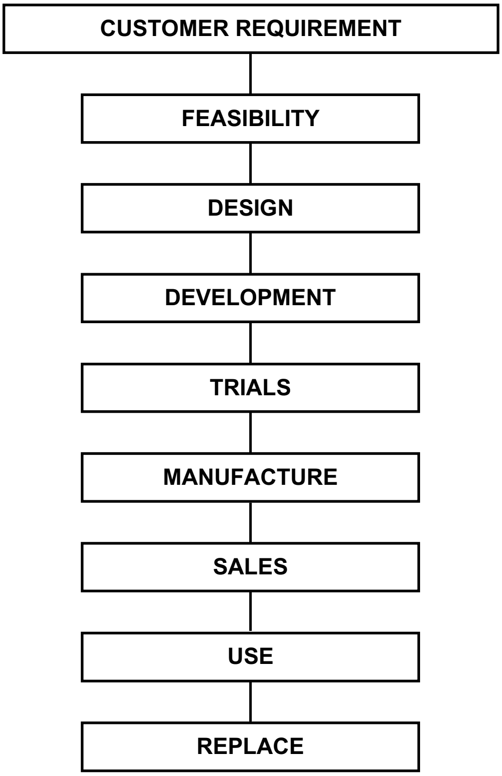

Figure 1.1 Different stages of product life cycle. (From Blischke and Murthy, 2000.)

the manufacturing process must be fine tuned and quality control procedures established to ensure that the items produced have the same performance characteristics as those of the final prototype. After this, the production and marketing efforts begin. The items are produced and sold. Production continues until the product is removed from the market because of obsolescence and/or the launch of a new product.

We focus our attention on the reliability of the product over its life cycle. Although this may vary considerably, a typical scenario is as shown in Figure 1.2. A feasibility study is carried out using the specified target value for product reliability. During the design stage, product reliability is assessed in terms of part and component reliabilities. Product reliability increases as the design is improved. However, this improvement has an upper limit. If the target value is below this limit, then the design using available parts and components achieves the desired target value. If not, then a development program to improve the reliability through test–fix–test cycles is necessary. Here the prototype is tested until a failure occurs and the causes of the failure are analyzed. Based on this, design and/or manufacturing changes are introduced to overcome the identified failure causes. This process is continued until the reliability target is achieved.

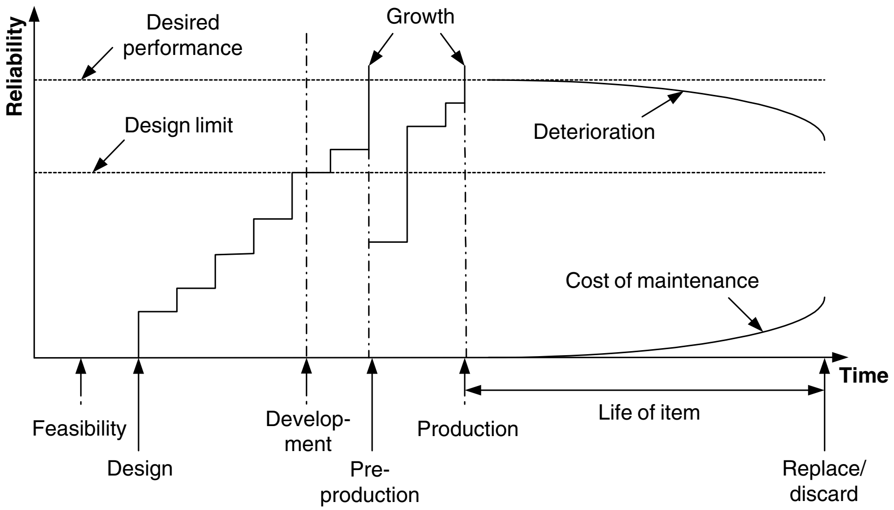

Figure 1.2 Reliability issues over product life cycle. (From Blischke and Murthy, 2000.)

The reliability of the items produced during the preproduction run is usually below that for the final prototype. This is caused by variations resulting from the manufacturing process. Through proper process and quality control, these variations are identified and reduced or eliminated, and the reliability of items produced is increased until it reaches the target value. Once this is achieved, full-scale production commences and the items are released for sale.

The reliability of an item in use deteriorates with age. This deterioration is affected by several factors, including environment, operating conditions, and maintenance. The rate of deterioration can be controlled through preventive maintenance as shown in Figure 1.2.

It is worth noting that if the reliability target values are too high, they might not be achievable with development. In this case, the manufacturer needs to revise the target value and start with a new feasibility study before proceeding further.

Weibull models have been used not only to model the reliability of the product but also to study other issues in the different stages of the product life cycle. We can group the different stages into two groups: (i) premanufacturing and (ii) postmanufacturing. The former deals with issues such as reliability assessment, accelerated testing, and reliability growth, and the latter with issues such as maintenance and warranties.

1.7 OUTLINE OF THE BOOK

The book is structured in seven parts (Parts A to G) with each part containing one or more chapters.

Part A consists of two chapters. Chapter 1 gives an overview of the book, and Chapter 2 deals with a detailed discussion of the taxonomy for Weibull models and the mathematical structure of these models.

Part B consists of three chapters. Chapter 3 deals with the analysis and presents various results dealing with model properties. Chapter 4 deals with parameter estimation and examines different data structures and estimation methods and their properties. Chapter 5 deals with model selection and validation, where the focus is on deciding whether a specific model is appropriate to model a given data set or not. In these chapters various concepts and techniques are introduced, and these are referred to in later chapters. In these three chapters, the analysis, estimation, and validation issues for the standard Weibull model as well as the three-parameter Weibull model are discussed as the two models are very similar.

Part C consists of two chapters. Chapter 6 deals with Type I models and Chapter 7 deals with Type II models. Part D consists of four chapters and deals with Type III models, and each chapter looks at a different family of models. Chapter 8 deals with mixture models, Chapter 9 with competing risk models, Chapter 10 with the multiplicative models, and Chapter 11 with sectional models. Part E consists of four chapters. Chapter 12 deals with Type IV models and looks at four different models. Chapter 13 deals with Type V models, Chapter 14 with Type VI models, and Chapter 15 with Type VII models. In Chapters 6 to 15, for each model, the model properties and the statistical inference issues are discussed.

Part F consists of a single chapter (Chapter 16) and deals with model selection. Here the focus is on deciding whether one or more Type I to III models are suitable for modeling a given univariate continuous valued data set.

Part G deals with the application of Weibull models in reliability theory. It consists of three chapters. Chapter 17 deals with modeling failures at system and component levels. Chapter 18 deals with a variety of reliability-related decision problems at the premanufacturing, manufacturing, and postsale service stages of the product life cycle. We briefly review the literature dealing with such problems with failures modeled by Weibull models.

1.8 NOTES

The notion of random variable and some basic concepts from probability theory can be found in most textbooks on the probability theory or models. See, for example, Ross (1983). The four volumes by Johnson and Kotz (1969a, 1969b, 1970a, 1970b) deal with discrete, univariate, and multivariate distributions. Basic concepts from the theory of stochastic processes can be found in most texts on the subject. See, for example, Ross (1980).

EXERCISES

1.1. Define a set of problems from each of the following areas that requires the use of probability distribution function to model one or more of the variables of importance to the problem:

a. Biological sciences

b. Social sciences

c. Health sciences

d. Physical sciences

e. Environmental science

1.2. Carry out a search of computer databases (such as Inspec, Compendex, etc.) to compile a list of studies dealing with the problems defined in Exercise 1.1.

1.3. A city council operates a fleet of buses. At the workshop level, for each bus the date of failure and the component that caused the failure are recorded. Discuss what other data should have been recorded for proper understanding of bus failures? The recorded data is converted into monthly failures for reporting to top management. Discuss if this is appropriate or not.

1.4. Discuss why the modeling of a data set is both an art as well as a science.

1.5. The lifetime of a human is a random variable and affected by several factors that can be grouped into genetic, ethnic, environmental, and lifestyle. One can model the lifetime by either ignoring these factors or by incorporating them. Discuss problems when it is appropriate to ignore them and when one needs to include them in the model building.

1.6. An automobile is a complex system comprised of several components. The failure of an automobile is due to the failure of one or more components. As such, one can model the failures either at the system or component level. Discuss when it is more appropriate to use the modeling at the system level as opposed to the component level.

1.7. Discuss the different kinds of data that is generated at each stage of the product life cycle that is relevant for determining the reliability of the product.

1.8. Weibull distribution is one of many probability distributions that have been proposed and studied. Make a list of distributions that you are familiar with and can be used to model nonnegative continuous random variable.

1.9. Repeat Exercise 1.8 for nonnegative discrete random variable.