

Chapter 2

Reliability of Engineered Systems

2.1 Introduction

Making decisions about the design and operation of infrastructure requires estimating the future performance of systems, which implies evaluating the system's ability to perform as expected during a predefined time window. This evaluation fits within what is known as *reliability analysis*. This chapter presents an introduction to the basic concepts and the theory of reliability in engineering, which provides the foundation for constructing degradation models (see Chaps. 4–7), performing life-cycle cost analyses (see Chaps. 8 and 9), and to designing maintenance strategies (Chap. 10). In the first part of this chapter, we present some conceptual issues about reliability and a description of basic reliability approaches. The second part of the chapter, Sect. 2.7 and onward, presents an overview of reliability models and sets the basis for theory that will be used and discussed in the rest of the book.

2.2 The Purpose of Reliability Analysis

Reliability analysis is the study of how things fail. Any engineered system, be it a facility (e.g., power plant) or infrastructure component (e.g., bridge), an electro-mechanical device, a consumer product, or even a manufacturing process, is designed and built to perform a specific *function* for a *specified duration* (the *mission* of the system). Once in use, the physical properties of the system will inevitably decline, and any engineered system will eventually *fail* (i.e., be unable to perform its designated function), possibly before completion of the mission. Moreover, engineered systems are typically operated in environments that are neither controllable nor predictable, and even well-designed and constructed systems may not fulfill their intended purpose due to unforeseen or unexpected events. As technology improves and new products enter the marketplace, consumers have become accustomed to expecting

dependable performance in the goods and services they buy and in the infrastructure developed to support their operation. Reliability analysis is the quantitative study of system failures and is an integral aspect of ensuring high-quality system performance.

As an engineering discipline, the field of reliability engages engineers of all disciplines, as well as physicists, statisticians, operations researchers, and applied probabilists. Furthermore, it encompasses a wide range of activities, which include, among others:

- collecting and analyzing data from physical and virtual experiments (*design of experiments, statistical, and simulated life testing*);
- characterizing the physical processes that lead to system failure (*physics of failure and degradation modeling*) and modeling the uncertainties that govern those failures (*probabilistic lifetime modeling*); and
- understanding the logical structure that determines the interactions and the dependencies between system components and their influence on overall system performance (*reliability systems analysis*).

The purpose of reliability analysis is not simply to describe how, when, and why systems fail, but rather to use information about failures to support decisions that improve the system's quality, safety and performance, and to reduce its cost. This aspect is especially important in areas where failures have serious consequences, for example, where public safety is involved or where significant financial investments are at stake (e.g., bridge failure). The acceptable performance of a system can be achieved in many ways; for example, through improvements in design and manufacture, and through better planning of operations (e.g., maintenance policies and warranty procedures); within this context, reliability analysis provides a quantitative foundation to support decisions that make these activities more efficient.

Reliability evaluation methods have been presented and discussed in a wide variety of applications, and many journals and books are available on the topic; see for instance [1–8]. This chapter presents some of the fundamental concepts of reliability analysis and introduces reliability methods which are of particular importance to support of decisions about future investments (e.g., design, manufacture, operation, and maintenance). Several references have been included for the reader to find more detailed information.

2.3 Background and a Brief History of Reliability Engineering

The field of reliability analysis began in earnest after World War II, when the U.S. and Soviet militaries both began systematic studies of newly developed weapons systems with the goal of improving their operation. In subsequent years, reliability engineering permeated the military, aerospace (particularly during the “space race”), and nuclear energy sectors. These sectors were still highly regulated by governmental entities, which led to the development of many standards, specifications, and procedures that govern product development in these sectors. Driven by increasing

competition and demands for high-quality and dependable consumer products, eventually, reliability analysis became widely adopted by many commercial enterprises, such as automotive manufacturing, consumer electronics, software, and appliances, to name just a few. In these industries, reliability analysis remains an important part of the product development and manufacturing process. Many reliability engineering techniques, such as fault tree analysis (FTA), failure mode, effects and criticality analysis (FMECA), and root cause analysis, are commonly used in the design and planning of engineered systems. Reliability analysis has also driven the development of fatigue and wear models, crack propagation models, corrosion models, and other methods of modeling physical wear out.

Reliability of infrastructure is, to a large extent, linked with the history of structural reliability. The first papers utilizing a probabilistic approach in design and analysis of structures were published in the late 1940s by Freudenthal [9], who discussed the basic reliability problem in structural components subjected to random loading, and in the early 1950s by Johnson [10], who proposed the first comprehensive formulation of structural reliability and economical design. These papers basically set the basis for a new field in structural engineering. In the 1960s, the basic concepts of safety (e.g., safety margin and safety index) were developed by Basler [11] and Cornell [12, 13], although there were also important contributions by other researchers such as Ferry-Borges [14] and Pugsley [15]. During the period from 1967 until 1974, the area of structural reliability attracted a great deal of interest in the academic community; however, its application and use in practice evolved only very slowly [3]. The work of Hasofer and Lind [16] and Veneziano [17] in the early 1970s, among others, led to the first standard in limit state format based on a probabilistic approach, the CSA [18], published in 1974. This publication was followed by development of other worldwide standards, and nowadays the probabilistic approach (mostly through partial safety factors) is used in almost every code of practice. More recently, the Joint Committee on Structural Safety (<http://www.jcss.byg.dtu.dk/>) has been working extensively to improve the general knowledge and understanding within the fields of safety, risk, reliability, and quality assurance in infrastructure design and development.

Interestingly, there are several important commercial sectors, where reliability engineering is still in a relatively nascent phase. These sectors include medical device manufacturing and food engineering. In medical device manufacturing, only relatively simple, qualitative techniques are commonly employed, and then primarily to respond to regulatory requirements. While it may appear somewhat unorthodox to consider food as an engineered system, many new methods of treating, processing, and packaging food are under development, and only very few studies on their reliability have been performed. Thus there is still a great need for engineers educated in the principles of reliability analysis among all sectors of the economy.

Despite the fact that the field of reliability now comprises a mature body of work, it is by no means a closed subject. In particular, there is still much work to be done in dealing with complex models such as those that describe the performance of large infrastructure systems. New developments in the theory and analysis of random processes have appeared that lend themselves particularly well to the performance analysis of infrastructure systems. At the same time, the increasing scrutiny of

the financial performance of massive infrastructure projects, and the socio-technical aspects of project execution and operation [19, 20], has demanded advanced reliability models to support both public and private investment decisions.

2.4 How do Systems Fail?

Before describing a mathematical framework for reliability analysis, it is valuable to develop a simple conceptual model for the “natural history” of an engineered object (e.g., a bridge, an electronic circuit, or an engine). A newly produced engineered object is imbued during manufacture with an initial physical capacity/resistance, commonly referred to as the *nominal life* of the object. Nominal life is measured in terms of a physical quantity (or indeed, a *vector* of physical quantities) whose units will be referred to as “life units.” Because of variations in materials, manufacturing or construction processes, etc., the nominal life of an engineered object is taken to be a random quantity. In the object’s operating environment, the physical capacity/resistance of the object declines through the process of *degradation*; see Chap. 4. Conceptually, degradation is the process of stripping out life units from the object over time; it can be described mathematically as a random process that measures the life units removed from the object over time. At any point in time, the *remaining life* (or *remaining capacity/resistance*)¹ is defined to be the difference between the nominal life and the accumulated degradation up to that time. When the remaining life declines to zero (or reaches a minimum performance threshold), the object fails. The time at which failure occurs is referred to as the *lifetime* of the system.

Two examples, shown in Fig. 2.1, illustrate the concepts of nominal life and deterioration. The first example considers an incandescent lightbulb. The lightbulb contains a filament that converts electrical energy into light energy (photons) and heat. Over time, the heat causes a reduction in the material of the filament (tungsten atoms detach from the filament), and eventually, when the filament becomes thin enough, detachment results in the loss of the electrical conductivity, and the bulb fails. In this example, we use the initial width of the filament as the nominal life of the bulb. Deterioration is the process by which the width of the filament is reduced during use. The second example is the case of a bridge located in a seismic region. In this example, nominal life is measured in terms of the bridge’s initial structural capacity (e.g., stiffness). Degradation can then be described by two mechanisms, one related to the weakening of the structure due to steady wear out over time, the second related to sudden decreases in capacity as a result of earthquakes of various magnitudes. As a result of these two phenomena, the structural capacity is reduced over time until it reaches a minimum performance threshold that defines system failure.

This characterization of failure is useful for several reasons. It suggests that in most systems there are two distinct and independent factors that determine system

¹Throughout the book the terms “remaining life” and “remaining capacity/resistance” will be used interchangeably.

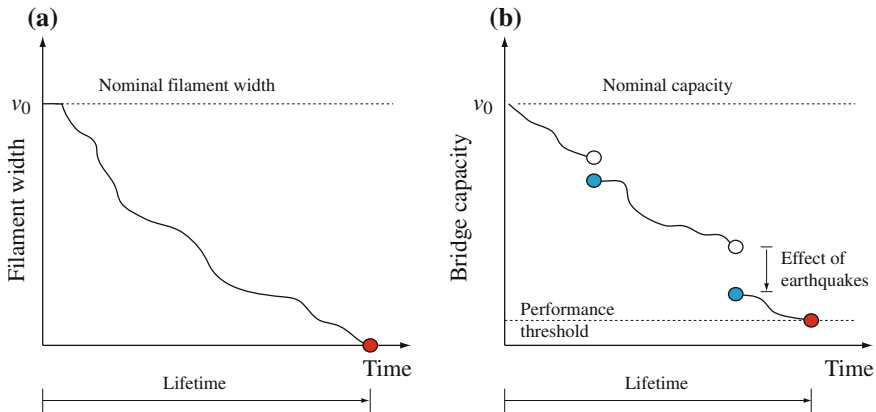


Fig. 2.1 Sample path of degradation of two systems over time: **a** the filament thickness of a light bulb; and **b** a bridge structural capacity

lifetime, namely the manufacturing or construction process that establishes nominal life (resistance/capacity), and the operating conditions and environmental processes that govern deterioration. Thus it is natural to study these factors separately. There are a variety of tools for studying the distribution of nominal life (resistance/capacity) via manufacturing process variability (e.g., quality control) and static design reliability estimation (see for instance Sect. 2.8). On the other hand, studying system degradation involves generally more elaborated stochastic process models such as Markov chains, Brownian motion, compound Poisson processes, and Lévy processes; some of them will be discussed in Chaps. 3–5 and 7. The characterization of failure in terms of these two factors (manufacturing and environment) is central to evaluating the system lifetime as a random variable.

2.5 The Concept of Reliability

The widely used and general accepted definition of reliability, and one which will be adopted in this book, is the following:

The reliability of a system² is the likelihood that it will perform its required functions under stated conditions for a specified period of time.

Note that for any given situation, it is necessary to define exactly what is understood by the terms used above. Thus, unavoidably, engineering judgement is required in defining essential concepts such as “required functions,” “stated conditions,” and “specified period of time”; these make up the *mission* of the system. Furthermore,

²In this book, we will use also the terms *system*, *device* or *component* as the object of a reliability study. Most of the concepts and theory presented here are applicable to a wide range of objects, therefore, the term *system* is used as a general description of the object of study.

the notion that the system “performs its required functions” suggests the need to distinguish clearly between two possible system operating states, namely “satisfactory” and “not satisfactory” (i.e., failed).

The definition of reliability presented above also introduces the need to measure a “likelihood,” and hence, it rests on the mathematical foundations of probability theory as the means by which reliability is characterized. Taking the system’s *lifetime* to be its operating time, the definition of reliability above can be rephrased as follows:

The reliability of a system is the probability that the system’s lifetime exceeds a specific period of time (e.g., its mission time).

Finally, in terms of a system’s performance indicator, an alternate (equivalent) definition of reliability is:

The reliability of a system is the probability that the system’s performance indicator remains above a predefined threshold within a specific period of time.

In the definition of reliability based on a system performance indicator (e.g., resistance measure), the threshold is the minimum value above which the system is deemed to operate successfully. This threshold is a very important concept in engineering design and is frequently referred to as the *limit state*: the value of a performance measure below which a system fails to perform its function satisfactorily.

The limit state concept has been used extensively as a design and operating criterion in mechanical problems and especially in various civil engineering fields such as soil mechanics, pavements, and structures. Although different limit states can be defined, there are two of particular importance, which will be used throughout this book: *ultimate* and *serviceability* limit states. Ultimate limit states describe the system’s condition beyond which its operation is unacceptable, for instance partial or total structural instability, structural collapse, attainment of the maximum resistance (for some components or the entire system) or unacceptable deterioration. On the other hand, serviceability limit states allow for the system to perform below the expectations but without failure, for instance, excessive deformations, vibration or noise, or esthetic degradation.

2.6 Risk and Reliability

Reliability is often associated with the terms “risk” and “risk analysis”; nevertheless, risk and reliability are different concepts. The field of risk analysis differs from reliability engineering in that it takes a broader approach to threats and their consequences. Risk analysis is a process of collecting evidence of possible unwanted future scenarios (consequences of detrimental outcomes) throughout the system’s life cycle; therefore, both qualitative and quantitative analysis are important. The results from reliability analysis can be used as evidence in risk analysis. In risk analysis, aspects such as the socioeconomic evaluation of consequences, communication, management, and policy are very important. Frequently, probabilistic risk

analysis (PRA), which is commonly taken as a systematic evaluation of the likelihoods of some consequences, is seen as subsumed in reliability analysis. However, although they might sometimes look similar, there are some important differences in the fundamentals of both approaches.

In this book, we will mention the term risk marginally; our focus is only on the theoretical aspects of reliability, as described in the following sections. Further reading on the conceptual aspects of risk analysis and its relationship with reliability can be found in [7, 21, 22]

2.7 Overview of Reliability Methods

Although there are many ways of approaching reliability, the selection of any strategy cannot be detached from the decision problem. This means that the analysis should balance relevance and precision so that the results become meaningful evidence for the decision. The selection of the approach that best suits the decision problem depends on the knowledge and understanding of the performance of the system, as well as on aspects such as the availability and quality of information, and the resources available.

The traditional way to classify reliability methods groups them in four levels based on the extent of information that is used [3, 5]. Thus, *level I* methods use one characteristic value of each uncertain parameter. It is basically a non-probabilistic approach and a generalized version of the *safety factor* commonly used in engineering design. In *level II* methods, random variables are described by two parameters (e.g., mean and variance), and they are usually assumed to be normally distributed. Furthermore, in these models the reliability problem is described by a simple limit state function. The reliability index presented in Sect. 2.8.1 is a case in point. The third category, *level III* methods, focuses on estimating the probability of failure, which requires information about the joint distribution of all uncertain parameters. This level also includes system reliability problems and transient (time-dependent models) analysis. Finally, *level IV* methods combine reliability models with information about the context, for example, cost-benefit analysis, life-cycle cost analysis, failure consequences, operation policies (maintenance and intervention strategies) and so on. Within this context, most of this book is about *level IV* reliability methods.

2.8 Traditional Structural Reliability Assessment

2.8.1 Basic Formulation

It is quite common in the civil engineering literature (cf. [3, 5, 8, 23]) to assess structural reliability in a static sense by comparing the (random) capacity/resistance (strength) of the system to the load/demand (stress) placed on the system. In the

literature, this approach, also termed *interference theory* [24] or the *basic reliability problem* [5], is most useful during the design phase, when physical models for determining the system capacity may be available.

In this case, the system is deemed to fail when the demand (e.g., load) exceeds the capacity (e.g., resistance) of the system. Thus, if we define a random variable C to be the capacity (with density f_C) and D to be the demand of the system (with density f_D), the limit state in this formulation is $C - D = 0$, where $C - D$ is the so-called *safety margin*. By definition, the reliability R of the system is given by

$$R = P(C > D) = P(C - D > 0) \quad (2.1)$$

If we further assume that C and D are independent and nonnegative random variables; then,

$$R = \int_{-\infty}^{\infty} f_D(x) \left(\int_x^{\infty} f_C(y) dy \right) dx, \quad (2.2)$$

which can also be written as

$$R = \int_{-\infty}^{\infty} f_D(x) [1 - F_C(x)] dx = \int_{-\infty}^{\infty} F_D(y) f_C(y) dy \quad (2.3)$$

Example 2.1 Consider a system subjected to a demand, which is assumed to be log-normally distributed. Three demand cases are considered. The density of all three possible demand functions (with the same mean but different COV) and the distribution of the resistance are shown in Fig. 2.2. The system's capacity (i.e., ability to accommodate the demand) is also assumed to follow a log-normal distribution with mean $\mu_C = 15$ and coefficient of variation $COV_C = 0.2$. Compute the reliability of the system.

The reliability of the system can be computed using Eq. 2.3:

$$R = \int_{-\infty}^{\infty} F_D(y) f_C(y) dy \quad (2.4)$$

For the particular case of lognormal demand and resistance, there is a close form solution; i.e.,

$$R = 1 - \Phi \left[- \frac{\ln \left(\frac{\mu_C}{\mu_D} \sqrt{\frac{1+COV_D^2}{1+COV_C^2}} \right)}{\sqrt{\ln[(1+COV_D^2)(1+COV_C^2)]}} \right] \quad (2.5)$$

where Φ is the normal standard distribution and $COV_{X_i} = \sigma_{X_i} / \mu_{X_i}$. Then, for the data used in this example, the reliability values for the three cases considered are: $R_{(COV=0.1)} = 0.961$, $R_{(COV=0.2)} = 0.926$, and $R_{(COV=0.3)} = 0.89$. These results

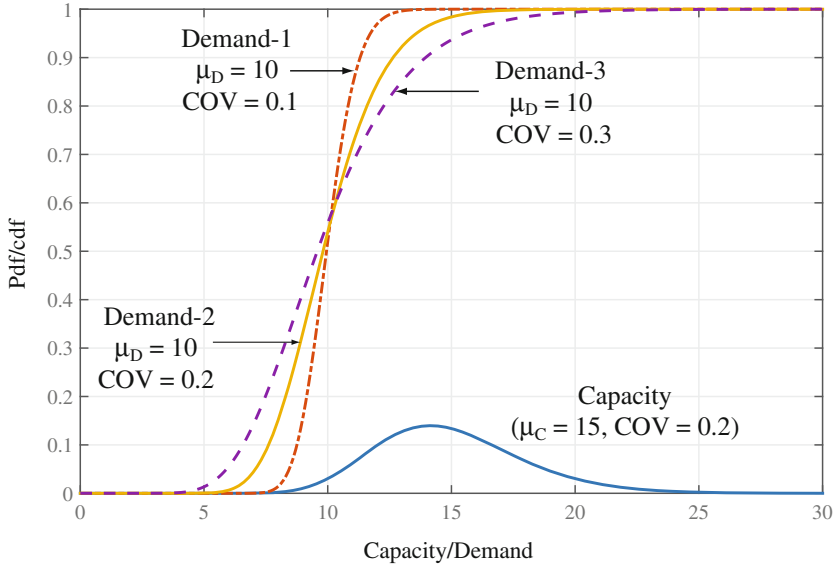


Fig. 2.2 Density function of the capacity and distribution function of the demand

show that larger variability implies larger failure probabilities and, therefore, smaller reliability values.

Let us now consider the special case where C and D in Eq. 2.3 are independent and normally distributed random variables. Let us further define $Z = C - D$, which is also normally distributed with parameters $\mu_Z = \mu_C - \mu_D$ and $\sigma_Z^2 = \sigma_C^2 + \sigma_D^2$; the density of Z is shown in Fig. 2.3. Then, the limit state can be defined as $Z = 0$. For this particular case, the reliability can be computed as:

$$R = \int_0^{\infty} f_Z(z) dz = 1 - \Phi\left(\frac{0 - \mu_Z}{\sigma_Z}\right) = 1 - \Phi(-\beta) \quad (2.6)$$

where $\beta = \mu_Z/\sigma_Z$ is called *safety* or *reliability index* [5]. The index β is a central concept in structural reliability. It is frequently used as a surrogate of failure probability and is widely used as a criteria for engineering design. For example, typical safety requirements for standard civil infrastructure (e.g., bridge design [25]) use $\beta \approx 3.5$ – 4.0 as an acceptable performance criteria [25].

2.8.2 Generalized Reliability Problem

Often, the formulation of the reliability problem (limit state) in terms of capacity, C , and demand, D , alone (Eq. 2.3) is not feasible, or it is incomplete because additional

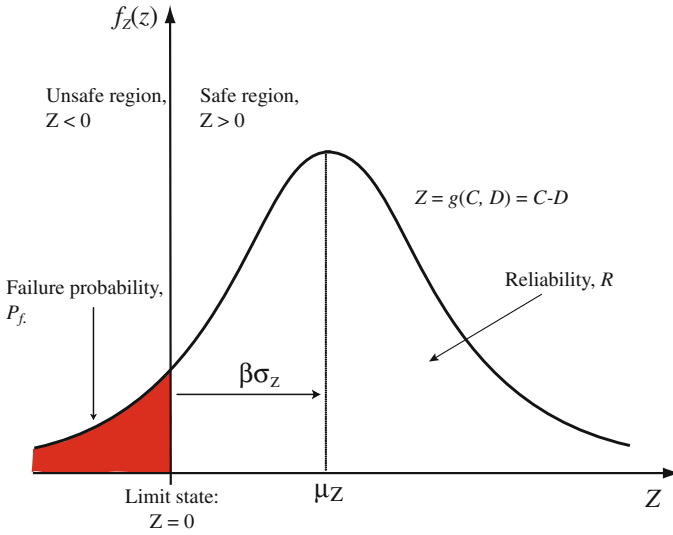


Fig. 2.3 Definition of the *reliability index* for the case of two normal random variables

information needs to be considered. In these cases, it may be of interest to describe the reliability problem in terms of a set of basic variables: $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$. In this n -dimensional variable space, the limit state $g(\mathbf{X}) = 0$ separates the safe ($g(\mathbf{X}) > 0$) and failure ($g(\mathbf{X}) \leq 0$) regions. The function $g(\mathbf{X}) = 0$ is a measure of a specific system performance condition based on a set of random variables $\mathbf{X}$ and other parameters that are not random.

Thus, a general form of Eq. 2.3 can be written as,

$$R = P(g(\mathbf{X}) > 0) = \int \cdots \int_{g(\mathbf{X}) > 0} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (2.7)$$

where $f_{\mathbf{X}}(\mathbf{x})$ is the joint probability density function of the n -dimensional vector $\mathbf{X}$ of basic variables. Note that neither the resistance nor the demand are explicitly mentioned in this formulation. Equation 2.7 is usually referred to as the *generalized reliability problem* [5].

The solution of Eq. 2.7 is not always an easy task. For instance, there may be a large number of variables involved, the limit state function may not be explicit (i.e., it cannot be described by a single equation), or the solution cannot be found either analytically or numerically. Then, several alternative approaches have been proposed to solve Eq. 2.7; they can be grouped in:

- analytical solutions (e.g., direct integration) or numerical methods;
- simulation methods (e.g., Monte Carlo); or
- approximate methods (e.g., FORM/SORM)

Solving Eq. 2.7 by direct integration or through numerical methods is possible using specialized software such as Matlab®, Mathcad®, or Mathematica®. However, in most cases, this is only possible for simple mechanical problems with few variables and known probability distributions. Therefore, alternative approaches such as simulation and approximate methods have been proposed; they will be briefly discussed in the following subsections.

2.8.3 Simulation

As problems become complex, simulation appears as a good option to estimate reliability. Consider a system whose performance is defined by a set of random variables $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$ with joint probability density function $f_{\mathbf{X}}(\mathbf{x})$. Let us define an indicator function $I[\cdot]$ such that $I[\mathbf{x}] = 0$ for $g(\mathbf{X}) \leq 0$ (failure) and $I[\mathbf{x}] = 1$ for $g(\mathbf{X}) > 0$ (not failure). Then, the reliability can be estimated as the expected value of the indicator function; this is,

$$R = \int \cdots \int I[\mathbf{x}] f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (2.8)$$

The unbiased estimator of the reliability is:

$$R \approx \frac{1}{N} \sum_{i=1}^N I[\mathbf{x}] = \frac{N_F(g(\mathbf{x}) > 0)}{N} \quad (2.9)$$

where N is the number of simulations and $N_F(g(\mathbf{x}) > 0)$ is the number of cases in which the system has not failed.

Although simulation is a very valuable tool, it should be used with care. For instance, an aspect that requires special attention is the case of correlated variables. For correlated normal random variables, methods such as the Cholesky decomposition can be used [8, 23]; for arbitrary correlated variables, there are other methods available; e.g., see [5, 26]. Furthermore, defining the number of simulations necessary to obtain a dependable solution is also a difficult task. It clearly depends on the actual result; for example, if the failure probability is estimated to be about 10^{-4} , the number of simulations required should be larger than 10^4 . Although several statistical models have been proposed to select the number of simulations [8]; the best approach consists of drawing the expected value and the variance of the result as function of the number of simulations; in this case, the solution is reached at convergence.

Clearly the computational cost of simulation is a central issue. The computational cost grows with the number of variables and the complexity of the limit state function. Then, in order to reduce the number of simulations several *variance reduction* techniques have been proposed. Among the most used are importance

sampling, directional simulation, the use of antithetic variables and stratified sampling [5, 27]. Recently, due to the sustained growth of computational capabilities, enhanced simulation methods have gained momentum. Some examples are subset simulation [28, 29], enhanced Monte Carlo simulation [30], methods that use a surrogate of the limit state function based on polynomial chaos expansions and kriging [31, 32], and statistical learning techniques [33].

2.8.4 Approximate Methods

There are some widely used methods to approximate the solution of Eq. 2.7 out of which the most popular is called *First-Order Second Moment* (FOSM) approach. In this case, the information about the distribution of the variables is discarded and only the first two moments are considered. When the information about the distributions is retained and included in the analysis, this method changes the name to *Advanced First-Order Second Moment* (AFOSM). In these case, the limit state, i.e., $g(\cdot) = 0$ is approached using Taylor series facilitating the evaluation. When the method uses a first-order approximation, the method is called *First-Order Reliability Method* (FORM); and when it is based on a second-order approximation it is referred to as *Second-Order Reliability Method* (SORM). Both FORM and SORM are widely used in practical engineering problems [5, 34].

Both FORM and SORM are carried out in the standard or normalized variable space (i.e., $U_i = (X_i - \mu_{X_i})/\sigma_{X_i}$). In FORM, the reliability index, β (see Sect. 2.8.1), is calculated as the minimum distance from the origin to the first-order approximation (using Taylor series) of limit state function [5] (Fig. 2.4). Then, FORM consists on solving the following optimization problem:

$$\begin{aligned} & \text{Minimize } \sqrt{\mathbf{U} \cdot \mathbf{U}^T} \\ & \text{subject to } g(X_1, X_2, \dots, X_n) = 0 \end{aligned} \quad (2.10)$$

where $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$ defines the space of the original variables; and $\mathbf{U} = \{U_1, U_2, \dots, U_n\}$ is the set of normalized independent variables.

Frequently, the limit state function is not linear. In this cases, FORM can be used only to approximate the solution and the quality of the results depends on the nonlinearity of the limit state function g (Fig. 2.4); i.e., as g becomes highly nonlinear the FORM approximation is less accurate. SORM is an alternative to deal with this problem since it uses a second-order approximation to the limit state function; however, the mathematical complexity of the solution increases significantly for high-dimensional variable problems. Another important difficulty of this approach arises when the random variables are not normally distributed. In this case, FORM cannot be applied directly. To manage this problem, Fiessler and Rackwitz [35] proposed a solution that approximate the tail of nonnormal distributions to normal distributions; this method has been used widely used with rather good results.

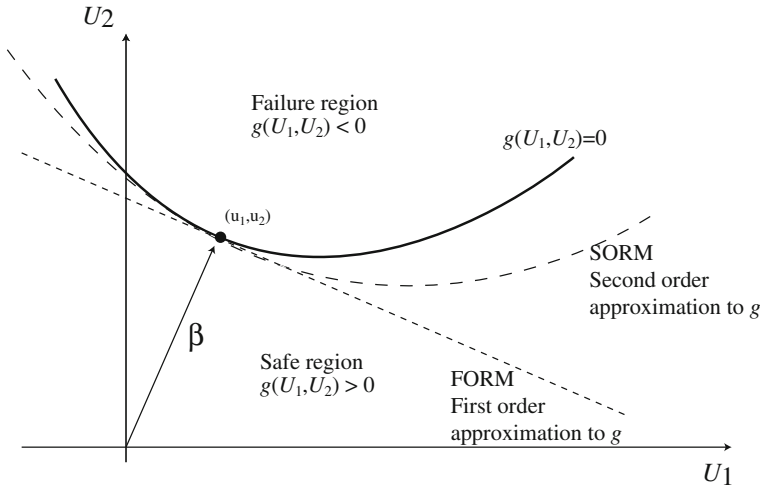


Fig. 2.4 Definition of the *reliability index* as the distance to the limit state function for the case of two random variables

The details of these methods are beyond the scope of this book and have been widely discussed elsewhere; e.g., [3, 5, 8, 23, 36].

2.9 Notation and Reliability Measures for Nonrepairable Systems

The static approach shown above lends itself very well for design studies and when the mission length of the system is fixed in advance. However, the primary focus of this book is on systems that evolve over time and which have an indeterminate mission length. Thus, it is important to distinguish between systems that are nonrepairable (that is, they are abandoned after a failure occurs), and systems that can be maintained operational through some external actions. In the latter, the system may experience a sequence of failures, repairs, replacements, and other maintenance activities.

The purpose of this section is to introduce the notation and basic notions of reliability that will be used later on in the book. Initially, we consider the case of a system that terminates upon failure, but in the later sections, we will extend this framework to include repairable systems. For these systems, we require a somewhat more general (although completely consistent) approach. These definitions are all quite standard and can be found in many reliability texts; e.g., [1, 2, 37–39].

2.9.1 Lifetime Random Variable and the Reliability Function

The study of reliability revolves around the idea that the time at which a system fails cannot be predicted with certainty. We define the *lifetime*, or *time to failure* (these are equivalent concepts) as a nonnegative random variable L , measured in units of time and described by its cumulative distribution function:

$$F_L(t) = P(L \leq t), \quad t \in [0, \infty] \quad (2.11)$$

We will typically assume that the lifetime is continuous, and thus has density f_L , where

$$f_L(t) = \frac{dF_L(t)}{dt}. \quad (2.12)$$

When the context is clear, we will drop the subscript and refer to the distribution function of the lifetime simply as F ; with density f .

The reliability of the system at time t , $R(t)$, is defined as the probability that the system is operational at time t ; i.e.,

$$R(t) = P(L > t) = 1 - F(t) = \bar{F}(t) \quad (2.13)$$

Clearly, the reliability function $R(\cdot)$ is simply the complement of the distribution function of the lifetime evaluated at time t . Also known as the *survivor function*, $R(t)$ represents the probability that the system operates satisfactorily up to time t . Then, it follows that

$$R(t) = 1 - \int_0^t f(\tau) d\tau = \int_t^\infty f(\tau) d\tau \quad (2.14)$$

and the density of the time to failure can be expressed in terms of the reliability as:

$$f(t) = -\frac{d}{dt} R(t) \quad (2.15)$$

2.9.2 Expected Lifetime (Mean Time to Failure)

The mean system lifetime (also known as mean time to failure or MTTF) is simply the expectation of L ; i.e.,

$$\mathbb{E}[L] = MTTF = \int_0^\infty \tau f(\tau) d\tau. \quad (2.16)$$

Because the lifetime is a nonnegative random variable, the *MTTF* can be expressed (using integration by parts) in terms of the reliability function as

$$MTTF = \int_0^{\infty} R(\tau) d\tau. \quad (2.17)$$

2.9.3 Hazard Function: Definition and Interpretation

The (unconditional) probability of failure of a device in the time interval $[t_1, t_2]$ is given by $F(t_2) - F(t_1)$ (or $R(t_1) - R(t_2)$). To compute the (conditional) probability of failure of a device in a certain time interval *given* that the device is working at the beginning of the time interval involves the concept of the *hazard function*, also called the *hazard rate*, $h(t)$. The hazard function can be interpreted as the instantaneous failure rate (i.e., failure in the next small instant of time) of a system of age t ; in terms of conditional probability

$$h(t)\Delta t \approx P(L \leq t + \Delta t \mid L > t), \quad (2.18)$$

for small values of Δt . Therefore, the hazard function $h(t)$ is defined by

$$\begin{aligned} h(t) &= \lim_{\Delta t \rightarrow 0} \frac{P(L \leq t + \Delta t \mid L > t)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} \frac{P(t < L \leq t + \Delta t)}{\Delta t P(L > t)} \\ &= \frac{f(t)}{R(t)} \end{aligned} \quad (2.19)$$

Consequently, the cumulative hazard function, denoted by Λ , is defined by:

$$\Lambda(t) = \int_0^t h(s) ds. \quad (2.20)$$

It is easy to show that [6]

$$\Lambda(t) = -\ln\{R(t)\}, \quad (2.21)$$

or put differently,

$$R(t) = \exp \left\{ - \int_0^t h(s) ds \right\} = \exp\{-\Lambda(t)\}. \quad (2.22)$$

This relationship establishes the link between the cumulative hazard function, i.e., $\Lambda(t)$, and the reliability function. Inserting Eq. 2.22 in 2.19 and solving for $f(t)$, we can also obtain an expression for the lifetime density in terms of the hazard function:

$$f(t) = h(t)\exp\{-\Lambda(t)\}. \quad (2.23)$$

A constant hazard function ($h(t) \equiv \lambda$ for all t and some $\lambda > 0$) holds if and only if the lifetime L has an exponential distribution with parameter $\lambda > 0$; i.e.,

$$f(t) = \lambda e^{-\lambda t} \quad (2.24)$$

and the reliability function can be expressed as

$$R(t) = e^{-\lambda t} \quad (2.25)$$

Exponentially distributed lifetimes have the “memoryless” property; that is, failures are neither more likely early in a system’s life nor late in a system’s life, but are in some sense “completely” random.

The hazard function has been used to study the performance of a wide variety of devices [6]. Generally, the hazard function will vary over the life cycle of the system, particularly as the system ages. A conceptual description of the hazard function that proves useful for some engineered systems is the so-called “bathtub” curve shown in Fig. 2.5.

The bathtub curve proposes an early phase, characterized by a decreasing hazard function (i.e., DFR), that reflects early failures due to manufacturing quality or design defects. This phase is commonly termed the *infant mortality* phase and is followed by a period of constant hazard, where failures are due to random external factors,

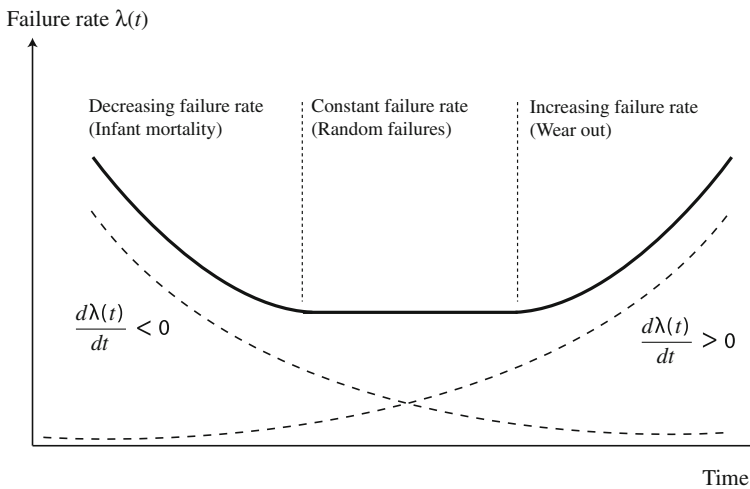


Fig. 2.5 Time-dependent failure rate: the *bathtub* curve

such as high vibrations, over-stresses, unexpected changes in temperature, and other extreme conditions. Finally, if units from the population remain in use long enough, the failure rate begins to increase as materials wear out and degradation failures occur at an ever increasing rate (i.e., IFR); this is known as the *wear out failure* period. Wear out is the result of aging due to, for instance, fatigue or depletion of materials (such as lubrication depletion in bearings).

Despite the fact that the bathtub curve is presented and discussed in almost all reliability books, some caveats on its practical applicability are in order. Its use as a conceptual device may be appropriate for some product *populations*, and in particular, the decreasing hazard part of the curve corresponds to the elimination through failure of relatively weaker members of the population (i.e., those of poor quality). There has been little published empirical evidence for the bathtub curve as a general model for the hazard function over a product's life, and a number of authors [40–42] have cautioned against its indiscriminate use in practice.

Statistical information about failure rates is usually fitted to a probability model. The numerical methods used for this purpose can be found elsewhere [4, 6, 23].

2.9.4 Conditional Remaining Lifetime

Another important concept in reliability analysis is the conditional remaining life distribution $H(t|x)$, defined as follows (Fig. 2.6):

$$H(t|x) = P(L \leq x + t | L > x) = \frac{F(x + t) - F(x)}{1 - F(x)}, \quad t, x \geq 0 \quad (2.26)$$

where L is the time to failure with distribution $F(t)$, and $H(t|x)$ is a conditional distribution, which can be interpreted as the distribution of the remaining life of a system of age x . If L is continuous, with density f , the conditional remaining life density is given by

$$h(t|x) = \frac{f(x + t)}{1 - F(x)}, \quad (2.27)$$

which is basically the density function of the time to failure truncated in x . The mean of this distribution gives the conditional expected remaining life $\mathbb{E}[L|x]$ of a system of age x :

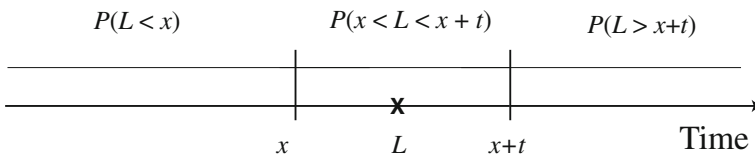


Fig. 2.6 Conditional remaining life

$$\mathbb{E}[L|x] = \mathbb{E}[L - x|L > x] = \int_0^\infty (1 - H(\tau|x))d\tau = \int_0^\infty \tau h(\tau|x)d\tau, \quad (2.28)$$

where the last equality holds if the lifetime distribution is continuous.

Example 2.2 According to field reports, the mean time to failure of a specific type of component was found to be $\mu = 12$. Because there is not clear information about the distribution of the time to failure, it is required to compute the basic reliability quantities for the following three distributions: lognormal (mean $\mu = 12$ and COV = 0.25), uniform [43, 44], and exponential with $\lambda = 1/12$.

Equation 2.19 was used to evaluate the hazard rate for the three distributions; the results are shown in Fig. 2.7. Note that for the particular and important case of the exponential distribution:

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{\lambda \cdot \exp(-\lambda t)}{\exp(-\lambda t)} = \lambda = \frac{1}{12}. \quad (2.29)$$

which is time-independent and reflects the memoryless property of the exponential distribution. The corresponding reliability functions were evaluated using Eq. 2.25, for $T_0 = 0$, the results are presented in Fig. 2.7b.

On the other hand, the conditional survival probability density (Eq. 2.27) for a value of $x = 3$ is shown in Fig. 2.8a. Note that the x -axis represents the time t after $x = 3$; for instance $h(t = 5|x = 3)$ means the density at a time $t = 8$. Finally, the evolution of the conditioned survival density function, for various x and for the lognormal case only, is presented in Fig. 2.8b. It can be observed that larger values of x shift the function to the left. This is caused by the fact that as x becomes larger, $1 - F(x)$ becomes smaller.

2.9.5 Commonly Used Lifetime Distributions

Among the most commonly used distribution functions in reliability and survival analysis are the exponential (described above), Weibull, lognormal, and gamma (although this list is by no means complete; for a more comprehensive list see) [45]. These distributions can be represented as special cases of the *generalized gamma* family. The generalized gamma is a three-parameter distribution; its density and cumulative distribution functions are given below [45]:

$$f(t; \theta, \beta, \kappa) = \frac{\beta}{\Gamma(\kappa)\theta} \left(\frac{t}{\theta}\right)^{\kappa\beta-1} e^{-(t/\theta)^\beta}, \quad t > 0 \quad (2.30)$$

$$F(t; \theta, \beta, \kappa) = \Gamma_1 \left[\left(\frac{t}{\theta}\right)^\beta; \kappa \right]. \quad (2.31)$$

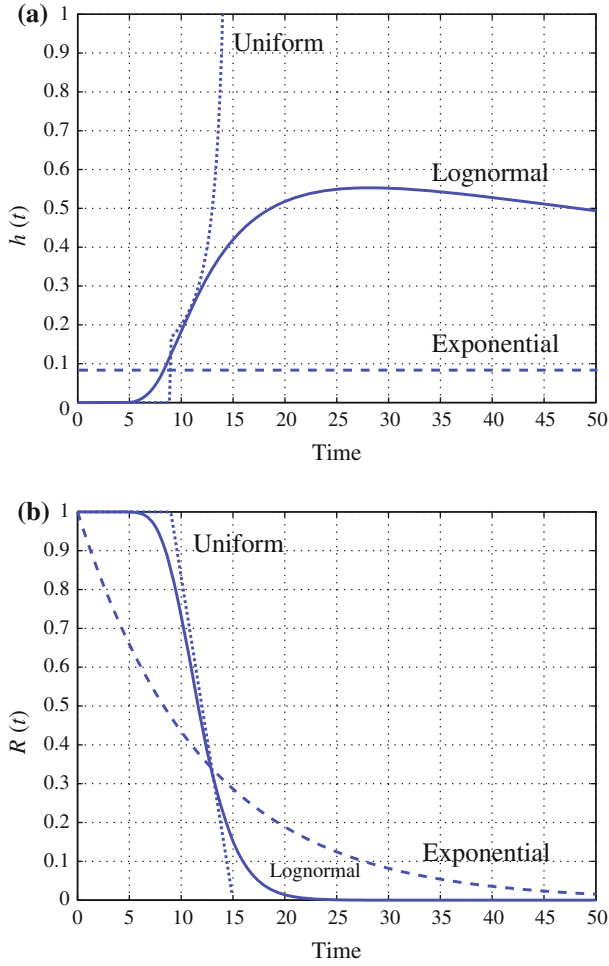


Fig. 2.7 **a** Failure rate and **b** reliability function for the three distributions

where $\theta > 0$ is a scale parameter, and $\beta > 0$ and $\kappa > 0$ are shape parameters; Γ is the gamma function and Γ_1 is the incomplete gamma function; i.e.,

$$\Gamma(\kappa) = \int_0^{\infty} z^{\kappa-1} e^{-z} dz, \quad z > 0 \quad (2.32)$$

$$\Gamma_1(z; \kappa) = \frac{\int_0^z y^{\kappa-1} e^{-y} dy}{\Gamma(\kappa)}, \quad z > 0. \quad (2.33)$$

Table 2.1 shows the parameter selection for the special cases of the generalized gamma mentioned above.

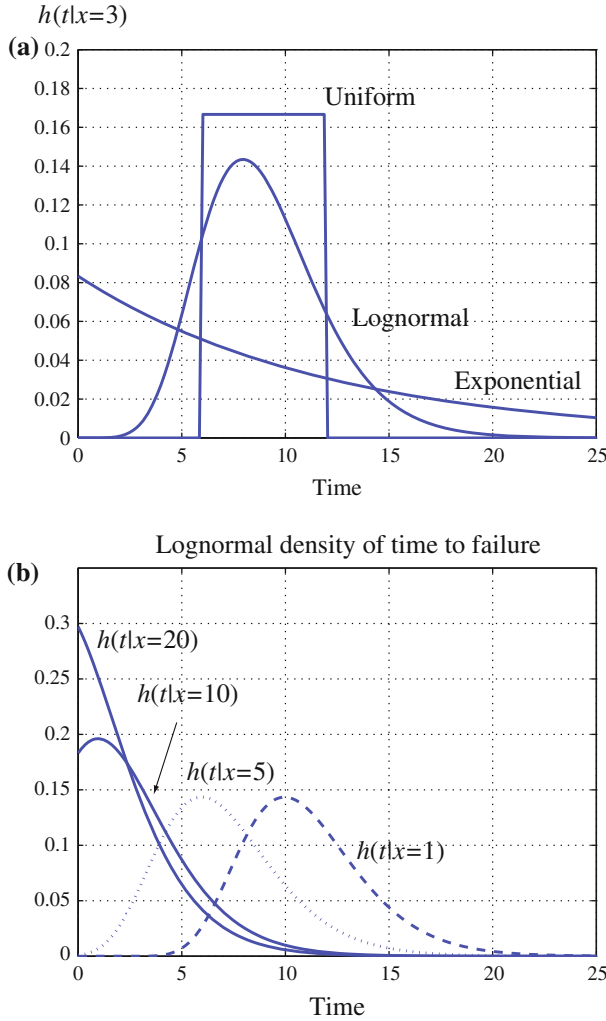


Fig. 2.8 Conditional density function for **a** $x = 3$ and all three failure time distributions; and **b** for $x = \{1, 5, 10, 20\}$ and the lognormal failure time distribution

2.9.6 Modeling Degradation to Predict System Lifetime

Based on the discussion in Sect. 2.4, L is realized when the degradation accumulated by the system meets or exceeds its nominal life (or more generally, the performance threshold or limit state); see Fig. 2.9.

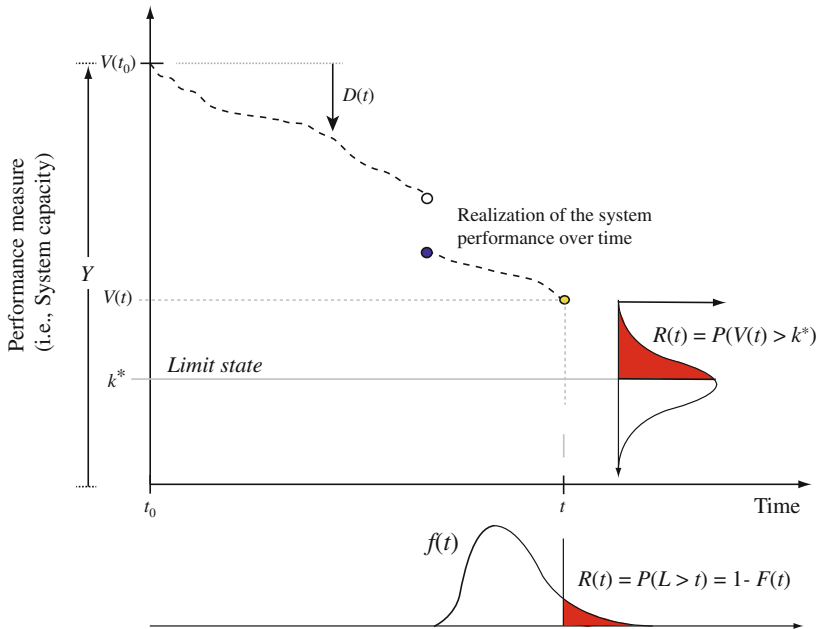
To formalize this idea, let us define Y as a positive random variable that measures nominal capacity of a system (in physical units); i.e., initial capacity. Let us further define $V(t)$ to be a system performance indicator at time t ; for example, the structural

Table 2.1 Special cases of the generalized gamma distribution [45]

Parameters of the generalized gamma	Distribution	$F(t)$
$\beta = 1$	Gamma (θ, κ)	$\Gamma_1\left(\frac{t}{\theta}; \kappa\right)$
$\kappa = 1$	Weibull ($\ln(\theta), 1/\beta$)	$1 - \exp\left(-\left(\frac{t}{\ln(\theta)}\right)^{1/\beta}\right)$
$\beta = 1; \kappa = 1$	Exponential (θ)	$1 - \exp(-\theta t)$
$\kappa \rightarrow \infty$	Lognormal	$\Phi\left(\frac{\ln(t) - \frac{\ln(\theta) + \ln(\kappa)}{\beta}}{1/(\beta\sqrt{\kappa})}\right)$

capacity of a bridge after t years. To allow generality, we will henceforth refer to $V(t)$ simply as “remaining capacity” of the system at time t and $D(t)$ as the *total* degradation by time t . Then, if the remaining capacity decreases over time as a result of the process of degradation, the random variable that describes the systems lifetime can be viewed as the length of time required for the remaining capacity to reach a threshold k^* , with $k^* \leq Y$. Therefore, for $t \geq 0$,

$$V(t) = \max(Y - D(t), k^*) \quad (2.34)$$

**Fig. 2.9** Illustration of the definition of reliability

and

$$L = \inf\{t \geq 0 : V(t) \leq k^*\}, \quad (2.35)$$

or equivalently,

$$L = \inf\{t \geq 0 : D(t) \geq Y - k^*\}. \quad (2.36)$$

where k^* is the minimum performance threshold for the system to operate successfully; i.e., *limit state* (see Fig. 2.9). So we can interpret the device lifetime L as a first passage time of the total degradation process to a random threshold $Y - k^*$. As we mentioned earlier, this characterization allows, at least conceptually, for us to model the fact that random environmental effects “drive” system degradation. However, we should note at the outset that first passage problems are, in general, somewhat difficult to analyze for general degradation processes. The later chapters of this book will be devoted to these types of problems.

Note also that the relationship between reliability evaluated in terms of the system life, L , and as a static condition at a given point in time t is shown also in Fig. 2.9; this complementarity can be observed as well in Eqs. 2.35 and 2.36.

2.10 Notation and Reliability Measures for Repairable Systems

The previous section presented notation and reliability measures for systems consisting of a single lifetime; that is, systems that are abandoned upon failure. Most systems of interest, however, are not discarded (or replaced) upon failure, but rather made operational again by some type of maintenance or repair. Maintenance activities may be scheduled *prior* to failure as well (preventively), in an attempt to avoid failures at inopportune times (see Chap. 10). Repairable systems are studied with a variety of outcomes in mind, such as to minimize overall life-cycle costs, to develop effective inspection/maintenance strategies, to estimate warranty costs, and to decide when an aging system should be replaced (completely overhauled) rather than simply repaired. A sample path of a repairable system is shown in Fig. 2.10.

We will assume that failures render the system inoperable for a random amount of time during which the repair (or replacement) is made. In the simplest case, we might consider a sequence of successive lifetimes $\{L_1, L_2, \dots\}$ and a sequence of repair times $\{R_1, R_2, \dots\}$, where each lifetime is followed by a repair time.

Let us define the *system state* at time t , $Z(t)$, as operational ($Z(t) = 1$) or failed ($Z(t) = 0$); then we can define *point availability* $A(t)$ as the probability that the system is operational at time t . That is,

$$A(t) = P(Z(t) = 1) = P(V(t) > 0). \quad (2.37)$$

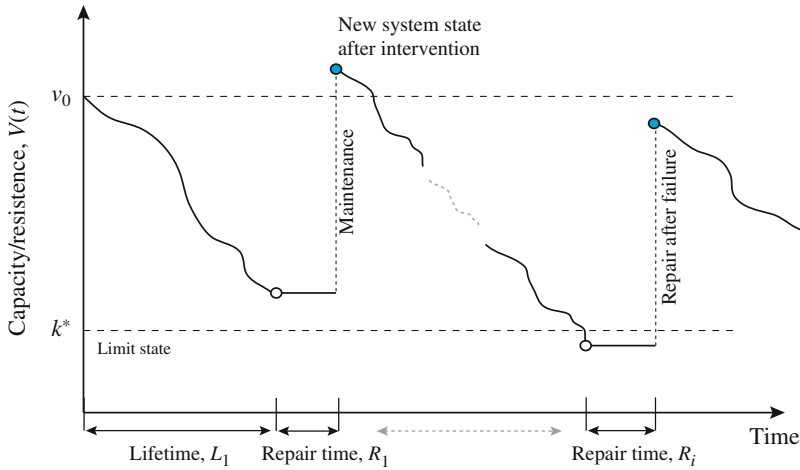


Fig. 2.10 Sample path of repairable system

Let us make note of the obvious—namely, that point availability is a time-dependent quantity that will typically depend on the *initial conditions*, that is, what is going on at the origin.

In addition to point availability, we will also be interested in the *limiting* availability A ; i.e.,

$$A = \lim_{t \rightarrow \infty} A(t). \quad (2.38)$$

In order to work with limiting availability, we will first need to make sure that this quantity exists. For the models we will work with, the limiting availability will typically also be a *stationary* availability; that is, for certain initial conditions, the limiting availability will describe the time-dependent availability for all t . Later in the book, we will discuss the problem of availability in more detail. Moreover, we will make some assumptions about the probability laws associated with lifetimes and repair times in order to calculate availability.

2.11 Summary and Conclusions

Reliability, the probability that the system performs as conceived, is a key concept in the design and operation of any engineered system. In structures and infrastructure, reliability methods have been traditionally classified in four levels (I to IV) depending of their complexity when modeling uncertainty; and according to the type and extent of information used in the analysis. Reliability models can be

organized also based on the relevance of the information that they provide for the decision making process.

Overall decisions about the performance of the system use models based on failure observations. On the other hand, decisions about specific system components require models that carefully describe their performance in time. In this chapter, we discussed and presented existing models to manage these types of problems. Since the theoretical aspects presented here have been widely discussed elsewhere, the chapter is intended only as a conceptual summary of the main ideas and techniques behind reliability modeling.

References

1. R.E. Barlow, F. Proschan, *Mathematical theory of reliability* (Wiley, New York, 1965)
2. T.J. Aven, U. Jensen, *Stochastic Models in Reliability*. Series in Applications of Mathematics: Stochastic Modeling and Applied Probability, vol. 41 (Springer, New York, 1999)
3. H.O. Madsen, S. Krenk, N.C. Lind, *Methods of Structural Safety* (Prentice Hall, Englewood Cliffs, 1986)
4. J.R. Benjamin, C.A. Cornell, *Probability, Statistics, and Decisions for Civil Engineers* (McGraw Hill, New York, 1970)
5. R.E. Melchers, *Structural Reliability-Analysis and Prediction* (Ellis Horwood, Chichester, 1999)
6. E.E. Lewis, *Introduction to Reliability Engineering* (Wiley, New York, 1994)
7. M.G. Stewart, R.E. Melchers, *Probabilistic Risk Assessment of Engineering Systems* (Chapman & Hall, Suffolk, 1997)
8. A. Haldar, S. Mahadevan, *Probability, Reliability and Statistical Methods in Engineering Design* (Wiley, New York, 2000)
9. A.M. Freudenthal, The safety of structures. Trans. ASCE **112**, 125–180 (1947)
10. A.I. Johnson, *Strength, Safety and Economical Dimensions of Structures*, vol. 22 (Statens Kommitte for Byggnadsforskning, Meddelanden, Stockholm, 1953)
11. E. Basler, Analysis of structural safety. In *Proceedings of the ASCE Annual Convention*, Boston MA, June 1960
12. C.A. Cornell, Bounds on the reliability of structural systems. ASCE-J. Struct. Div. **93**, 171–200 (1967)
13. C.A. Cornell, Probability-based structural code. J. Am. Concr. Inst. (ACI) **66**(12), 974–985 (1969)
14. J. Ferry-Borges, Implementation of probabilistic safety concepts in international codes, *Proceedings of the International Conference on Structural Safety and Reliability* Verlag, Dusseldorf, Aug 1977, pp. 121–133
15. A. Pugsley, *The Safety of Structures* (Edward Arnold, London, 1966)
16. A.M. Hasofer, N.C. Lind, Exact and invariant second moment code format. ASCE J. Eng. Mech. Div. **100**, 111–121 (1974)
17. D. Veneziano, Contributions to second moment reliability theory. Research Report R-74-33, Department of Civil Engineering, MIT, Cambridge, MA, 1974
18. Canadian Standard Association (CSA), Standards for the design of cold-formed steel members in buildings. CSA-S-136, Canada, 1974
19. D. Paez-Pérez, M. Sánchez-Silva, A dynamic principal-agent framework for modeling the performance of infrastructure. Eur. J. Oper. Res. (2016) (in press)
20. D. Paez-Pérez, M. Sánchez-Silva, Modeling the complexity of performance of infrastructure (2016) (under review)

21. D.I. Blockley, *Engineering Safety* (McGraw Hill, New York, 1992)
22. T. Bedford, R. Cooke, *Probabilistic Risk Analysis: Foundations and Methods* (Cambridge University Press, Cambridge, 2001)
23. A.S. Nowak, K.R. Collins, *Reliability of Structures* (McGraw Hill, Boston, 2000)
24. K.C. Kapur, L.R. Lamberson, *Reliability in Engineering Design* (Wiley, New York, 1977)
25. M. Ghosn, B. Sivakumar, F. Moses, Infrastructure planning handbook: planning engineering and economics. NCHRP Report 683: Protocols for Collecting and Using Traffic Data in Bridge Design. National Academy Press (National Academy of Science), Washington, 2011
26. P.-L. Liu, A. Der Kiureghian. Optimization algorithms for structural reliability analysis. Report UCB SESM-86 09, Department of Civil Engineering, University of California at Berkeley, 1986
27. S.M. Ross, *Simulation*, 4th edn. (Elsevier, Amsterdam, 2006)
28. S.K. Au, J. Beck, Estimation of small failure probabilities in high dimensions by subset simulation. *Prob. Eng. Mech.* **16**(4), 263–277 (2001)
29. S.K. Au, Reliability-based design sensitivity by efficient simulation. *Comput. Struct.* **83**, 1048–1061 (2005)
30. A. Naes, B.J. Leira, O. Batsevych, System reliability analysis by enhanced monte carlo simulation. *Struct. Saf.* **31**, 349–355 (2009)
31. B. Sudret, Global sensitivity analysis using polynomial chaos expansions. *Reliab. Eng. Syst. Saf.* **93**, 964–979 (2008)
32. B. Sudret, Meta-models for structural reliability and uncertainty quantification. In *Proceedings of the 5th Asian-Pacific Symposium on Structural Reliability and its Applications—Sustainable infrastructures*, ed. by K.K. Phoon, M. Beer, S.T. Quek, S.D. Pang (Reserch Publishing, Chennai, 2012), Singapore, 23–25 May 2012
33. J.E. Hurtado, *Structural Reliability: Statistical Learning Perspectives* (Springer, New York, 2004)
34. A. Haldar, S. Mahadevan, *Reliability Assessment Using Stochastic Finite Element Analysis* (Wiley, New York, 2000)
35. R. Rackwitz, B. Fiessler, Structural reliability under combined random load sequences. *Struct. Saf.* **22**(1), 27–60 (1978)
36. M. Sánchez-Silva, Introducción a la confiabilidad y evaluacin de riesgos: teoría y aplicaciones en ingeniera. Segunda Edición (Ediciones Uniandes, Bogotá, 2010)
37. E. Çinlar, *Introduction to Stochastic Processes* (Prentice Hall, New Jersey, 1975)
38. M. Finkelstein, *Failure Rate Modeling for Risk and Reliability* (Springer, New York, 2008)
39. I.B. Gerstbakh, *Reliability Theory with Applications to Preventive Maintenance* (Springer, New York, 2000)
40. G.-A. Klutke, P.C. Kiessler, M.A. Wortman, A critical look at the bathtub curve. *IEEE Trans. Reliab.* **52**(1), 125–129 (2003)
41. D. Kececioglu, F. Sun, *Environmental Stress Screening: Its Quantification, Optimization, and Management* (Prentice Hall, New York, 1995)
42. W. Nelson, *Applied Life Data Analysis* (Wiley, New York, 1982)
43. A.H.-S. Ang, W.H. Tang, *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering* (Wiley, New York, 2007)
44. S. Asmussen, F. Avram, M.R. Pistorius, Russian and American put options under exponential phase-type levy models. *Stoch. Process. Appl.* **109**, 79–111 (2004)
45. W.Q. Meeker, L.A. Escobar, *Statistical Methods for Reliability Data* (Wiley, New York, 1998)

Reliability and Life-Cycle Analysis of Deteriorating
Systems

Sánchez-Silva, M.; Klutke, G.-A.

2016, XXIV, 355 p. 106 illus., Hardcover

ISBN: 978-3-319-20945-6