# Chapter 1

# Introduction

Catastrophic tragedies such as

- the Space Shuttle *Challenger* and *Columbia* accidents,
- the sinking of the RMS *Titanic*,
- the Fukushima, Chernobyl, and Three Mile Island nuclear power plant accidents,
- the levee failure in New Orleans during Hurricane Katrina,
- the Bhopal industrial accident,
- the Tacoma Narrows, Minneapolis Interstate 35W, and Francis Scott Key bridge collapses,
- the Versailles railway accident, and
- numerous aircraft accidents

highlight the importance of reliability in design. This textbook describes probabilistic models that are helpful in designing high-reliability products and statistical techniques that can be applied to a data set of lifetimes. Although the majority of the illustrations in this text come from engineering problems, these techniques can be applied to problems that arise in actuarial science and biostatistics.

Reliability engineers, actuaries, and biostatisticians are all interested in lifetimes. A reliability engineer might study the lifetimes of products used in the marketplace. The interest might be in the failure time of a light bulb (a nonrepairable item) or a drill press (a repairable item). An actuary might be interested in the probability distribution of the lifetime of a person in order to determine the appropriate premium for a life insurance policy. A biostatistician might analyze the survival times of cancer patients in order to compare the effectiveness of treatment techniques, such as radiation versus chemotherapy.

Reliability engineers concern themselves with the lifetime of an item, such as a light bulb, a microprocessor, or an airplane wing. More generally, the models and methods described in this book can be applied to a computer program or a supply chain. Reliability engineers usually regard a complex system as a collection of components when performing an analysis. These components are arranged in a structure (simple examples include series and parallel structures) that allows the system state to be determined as a function of the component states. In many applications, component failure times are available. Interest in reliability and quality control has been revived by a more competitive international market, increased consumer expectations, condition-based preventive maintenance, and automation (for example, driverless cars).

Actuaries are usually employed by insurance companies to determine premiums and payment schedules for policies. Actuaries often use *life tables* to determine the appropriate life insurance premium for a particular individual. Life tables are used to determine the probability of death as a function of age for an individual. Factors such as medical history and occupation of an individual affect this probability. These tables need to be updated annually as medical science advances and new diseases arise. Actuaries need to understand the time value of money when determining the appropriate rate schedule for a policy, because payments are typically made over several years. Actuaries working in casualty insurance address analogous questions concerning structures (for example, home insurance), vehicles (for example, motorcycle insurance), etc.

Biostatisticians are interested in the lifetimes of organisms. These may be the lifetimes of patients undergoing treatment for a disease or laboratory animals in an experiment. Biostatisticians often investigate the effects of *covariates*, which are exogenous variables that might influence a lifetime. Simple examples of covariates include the gender, age, and cholesterol level of a patient. To account for these covariates, biostatisticians use probabilistic lifetime models that include covariates and associated regression coefficients that alter the probability distribution of the failure time.

Reliability engineers, actuaries, and biostatisticians use slightly different terminology when describing the object of interest and its lifetime. In particular,

- the reliability literature tends to use *system*, *component*, or *item*,
- the actuarial literature associated with life insurance uses *individual*, and
- the biostatistical literature tends to use *organism* or *patient*

to describe the object of a study. Likewise,

- the reliability literature tends to use *failure*,
- the actuarial and biostatistical literature tends to use *death*, and
- the point process literature tends to use *epoch*

to describe the event at the termination of a lifetime. To avoid switching terms, *failure* of an *item* will be used as much as possible throughout the text because the emphasis here is on reliability. The concept of failure time (or lifetime or survival time) is very general, and the probability models and statistical methods presented here apply to any nonnegative random variable (for example, the response time at a computer terminal).

To complicate matters further, the three different fields have developed their own notation, so the notation used here will not necessarily match the notation in the literature. Whenever possible, existing notation from the reliability literature will be used.

## 1.1 Definition of Reliability

Before considering the analysis of lifetime data (Chapters 7 through 11), probabilistic models for lifetimes will be introduced (Chapters 2 through 6). Many of these models have their basis in reliability theory. As shown in Chapter 3, the concept of reliability generalizes to a survivor function when time dependence is introduced. The definition for reliability is given next. The paragraphs following the definition expand on the italicized words in the definition.

**Definition 1.1** The *reliability* of an *item* is the *probability* that it will *perform* its specified *purpose* without failure for a specified *period of time* under specified *environmental conditions.*

This definition implies that the object of interest is an *item*. The definition of the item depends on the purpose of the study. In some situations, we will consider an item to be an interacting arrangement of components; in other situations, the component level of detail in the model is not of interest. Assume, for example, that a laptop computer (the item of interest) consists of three components: a keyboard, screen, and other electronic components. The object of the study may be the computer (ignoring the individual components), or the item may be modeled as an arrangement of the three components. The modeler needs to carefully define what the item of interest is and the level of detail to be considered when analyzing that item.

In addition to the question of resolution (that is, how to decompose the system into components), the modeler must determine an external boundary for the item. This external boundary determines what is to be considered part of the item and what is to be considered part of the environment around the item for modeling purposes.

Reliability is defined as a *probability*. Thus, the axioms of probability apply to reliability calculations. In particular, this means that all reliabilities must be between 0 and 1, inclusive. Results derived from the axioms are often used in reliability. For example, if two components with independent lifetimes have 1000-hour reliabilities of $p_1$ and $p_2$, and system failure occurs when either component fails, then the 1000-hour system reliability is $p_1 p_2$. This book emphasizes the quantitative aspects of reliability, typically described by probability measures. Other books emphasize the qualitative aspects of reliability, often described by adjectives.

*Performance* for an item must be stated unambiguously. A *standard* is often used to determine what is considered adequate performance. A ball bearing, for example, may be performing adequately if wear has not changed its diameter outside of the range $3 \pm 0.001$ mm. This tolerance is an example of a specification that delineates adequate performance from inadequate performance.

Specifying the adequate performance of an item implies the ability to recognize failure and identify the time of failure. If the item of interest is an automobile, for example, an appropriate performance level might be that the car is mobile. If this is the case, the automobile is still functioning if the radio is broken or the muffler falls off. The performance of an item is related to the mathematical model used to represent the state of the item. The simplest model for an item is a binary model, in which the item is in either the functioning or failed state. This model is easily applied to an electrical fuse, but more difficult to apply to items that gradually degrade over time, such as a drill bit. To apply a binary model to an item that degrades gradually, a threshold value must be determined to separate the functioning and failed states. To summarize, failure needs to be unambiguously defined in order to build a mathematical model for an item.

The definition of reliability also implies that the *purpose* or intended use of the item must be specified. It is common to produce several grades of one particular consumer good. A drill, for example, might have one grade for a weekend handyman and another for a contractor. Although these two grades of drills have identical functions, their reliabilities differ because they are designed for light and heavy usage.

Definition 1.1 also indicates that *time* is involved in reliability. Including time in the definition has five consequences.

- First, the units for time need to be specified (for example, minutes, hours, or years) by the modeler in order to perform any analysis.

- Second, many lifetime models use the random variable $T$ (rather than $X$, which is common in probability theory) to represent the random failure time of the item.

- Third, time need not be taken literally. The number of miles might represent time for an automobile tire, the number of flips might represent time for a light switch, the number of cycles might represent time for a ball bearing, and the number of whacks might represent time for a stapler.
- Fourth, a time duration must be associated with a reliability. To say that an item has a reliability of 0.8 is meaningless. This reliability for the item is valid for a particular time value, such as 1000 hours. (An exception here are static stress-strength models which specify a stress applied to an item having a particular strength. If the stress exceeds the strength, then the item fails. Dynamic stress-strength models address random loads applied to an item over time.)
- Finally, determining what should be used to measure the lifetime of an item may not be obvious. Many analysts have considered whether continuous operation or on/off cycling is more effective for items such as motors or computers. If a light bulb, for example, is to operate continuously, the number of hours burned should be used for the lifetime. If the light bulb is turned on and off, as most are, the number of hours burned may not be appropriate. If a comparison of the number of hours burned by bulbs burned continuously and bulbs burned intermittently showed a shorter burning lifetime for those burned intermittently, the modeler can assume that the shock of lighting and extinction decreases the light bulb's life. The modeler must account for these differences in order to compare light bulbs used continuously with those used intermittently. One way to model the age of the light bulb is to use a function of the number of hours burned and the number of illuminations so that all light bulbs will have equivalent age regardless of the number of illuminations. This example may be oversimplified, but it indicates that there may be more to defining a lifetime than just the amount of time spent functioning.

The final aspect of the definition of reliability is that *environmental conditions* must be specified. Conditions such as temperature, humidity, and turning speed all affect the lifetime of a machine tool. The 20,000-mile reliability of a subcompact car is different if it is used for highway driving compared to towing a trailer down city streets. Environmental conditions associated with the lifetimes of people might be the city in which they live and whether they smoke. Included in environmental conditions is the preventive maintenance to be performed on the item. Although preventive maintenance can be costly, it is usually effective in prolonging the lifetime of an item that is wearing out and hence increasing its reliability. In some settings, environmental conditions for an item cannot be specified in advance. In this case, the definition of reliability can be modified to be the probability of adequate performance under *encountered conditions*. Lab conditions typically do not provide a perfect match to encountered conditions.

Reliability is often misunderstood. A single grenade, for example, that explodes when it should might erroneously be called 100% reliable. This is inaccurate because a reliability of 100% implies that each grenade of this type will explode when it should. The true reliability of these grenades might be 0.9999 (or 99.99%), and we just happened to test one that worked.

Reliability is also often confused with quality. The primary difference between these two terms is that reliability incorporates the passage of time, whereas quality does not, because it is a static descriptor of an item. Two transistors of equal quality sit side by side on a shelf. One of these transistors will be used in a television set, the other in a missile launch environment. Both transistors are of identical quality, but the first one has a higher reliability because it will operate in a less stressful environment.

High reliability implies high quality, but the converse is not necessarily true. Consider two automobile tires, each of high quality. One was produced in 1957, the other in 2024. Although each was produced with the most stringent quality control procedures available when it was manufactured, their reliabilities will be different due to technology changes introduced between 1957 and 2024, such as steel-belted radials. The 60,000-mile reliability of the tire produced in 2024 will be higher than the 60,000-mile reliability of the 1957 tire. Technology advances in the years between the manufacture of the two tires may come in the form of improved design (for example, treads or steel belts), components (rubber), or processes (manufacturing advances). *Some* quality improvements (such as improved tread design) improve the reliability of the tire, while others (such as improved white wall design) do not.

## 1.2 Case Study: The *Challenger* Accident

This section consists of a brief case study involving the O-rings on the solid rocket motors on the Space Shuttle. Large, complex systems such as the Space Shuttle are often divided into *subsystems* to simplify reliability calculations. The Space Shuttle system consists of four subsystems: the *orbiter* that contains the crew and the controls, the *external liquid-fuel tank* that contains fuel for the main engines on the orbiter, and *two solid rocket motors* that are used to boost the orbiter into orbit. The emphasis here is on the reliability of the solid rocket motors because the Rogers Commission Report concluded that O-ring failure on the joints in these motors caused the tragic *Challenger* accident on January 28, 1986.

The two solid rocket motors are each shipped to the Kennedy Space Center in four pieces. Each assembled solid rocket motor contains three joints, for a total of six joints for the entire Space Shuttle. These joints are referred to as field joints, and they are critical to seal the small gaps that are left between each of the parts of the motor. O-rings, which measure 37.5 feet in circumference and are 0.28 inch thick, are used to seal the gaps in the field joints. All six O-rings must operate in order to avoid having the propellant escape from the solid rocket motor, so the O-rings form a six-component *series* system, as indicated by the *block diagram* in Figure 1.1. Series system arrangements for components as critical as O-rings are risky because failure of any one component results in catastrophic failure. A common technique used by design engineers to improve the reliability of a series system is to include a redundant component (also known as a secondary or backup component) that will take over if the primary component fails. The designers of the solid rocket motors added a secondary O-ring to back up the primary O-ring. This resulted in a total of 12 O-rings arranged as shown in Figure 1.2. This technique is highly effective if the components that are being placed in parallel fail independently of one another. The redundant component must become active and perform its function with a high probability upon failure of the initial component. Positively correlated failure times degrade the effectiveness of the technique.

In 1977, before the first shuttle flight in 1981, NASA discovered a problem known as *field joint rotation* that indicated that the failure of the primary and secondary O-rings might not be independent. Soon after ignition, the pressure and temperature inside the solid rocket motor increase rapidly. This causes the O-rings, along with some putty used to protect them, to seat and seal the field joint. The sides of the motor's metal casing also tend to bulge under the heat and pressure, and it was discovered that this bulge caused a rotation between the pieces of the motor. This rotation causes a

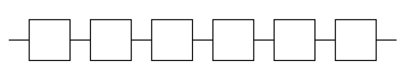

Figure 1.1: Six-component series arrangement of O-rings.

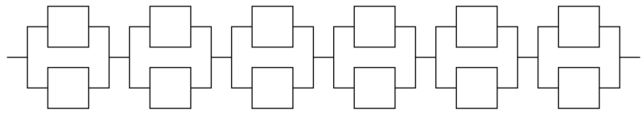

Figure 1.2: Twelve-component arrangement of O-rings.

widening of the gap that the primary and secondary O-rings were intended to seal. Unfortunately, this wider gap affected both the primary and secondary O-rings, so the assumption of independent failures was not justified. Thus, the reliability of each field joint was not as high as would be calculated assuming independence.

Prior to the *Challenger* accident, there were 24 shuttle flights. On 23 of these flights, which are listed in Table 1.1, the solid rocket motors were recovered from the ocean for inspection and possible reuse. (The O-rings from Flight STS-4 were lost at sea.) There was concern that an environmental variable, temperature at launch, might influence the reliability of the field joints. This was particularly important because the forecast of 26–29°F for the morning of the launch of the *Challenger* was by far the coldest launch temperature to date. Table 1.1 gives the flight number, joint temperature at launch, and number of field joint failures, sorted by temperature, for the 23 flights. More thorough analyses than that given here categorize the field joint failures into *erosion incidents*

| Flight | Temperature ($^\circ F$) | Failures |
|---|---|---|
| 51-C | 53 | 3 |
| 41-B | 57 | 1 |
| 61-C | 58 | 1 |
| 41-C | 63 | 1 |
| STS-1 | 66 | 0 |
| STS-6 | 67 | 0 |
| 51-A | 67 | 0 |
| 51-D | 67 | 0 |
| STS-5 | 68 | 0 |
| STS-3 | 69 | 0 |
| STS-2 | 70 | 1 |
| STS-9 | 70 | 0 |
| 41-D | 70 | 1 |
| 51-G | 70 | 0 |
| STS-7 | 72 | 0 |
| STS-8 | 73 | 0 |
| 51-B | 75 | 0 |
| 61-A | 75 | 2 |
| 51-I | 76 | 0 |
| 61-B | 76 | 0 |
| 41-G | 78 | 0 |
| 51-J | 79 | 0 |
| 51-F | 81 | 0 |

Table 1.1: Launch temperature and number of O-ring field joint failures.

and *blow-by incidents*. Furthermore, these incidents had been combined into what was known to the investigators as a *damage index*. In a three-hour teleconference between the manufacturer of the solid rocket motors and NASA officials on the evening prior to the accident, it was concluded that the previous launch temperature data were not conclusive on predicting primary O-ring failure. The Rogers Commission, appointed by President Reagan after the accident, believed that a mistake was made in the analysis of the O-ring failure data. The original analysis left out the flights with zero failures because it was believed that these flights did not contribute any information about the effect of temperature on failure. Figure 1.3 shows a plot of the data in Table 1.1, revealing that there is a "U" shape when the flights with zero failures are excluded, but a trend is more apparent when the flights with zero failures are included. Tied data pairs have been offset.

After the accident, an analysis by three statisticians concluded that the effect of temperature was indeed significant. They estimated the launch reliability to be as high as 0.87 when the temperature was 31°F and concluded that postponement of the flight until the temperature reached 60°F would have increased the launch reliability to as high as 0.98. The analysis presented here concerns the O-rings only, and, of course, thousands of other electrical, mechanical, and software system components needed to also function for the Space Shuttle to operate properly.

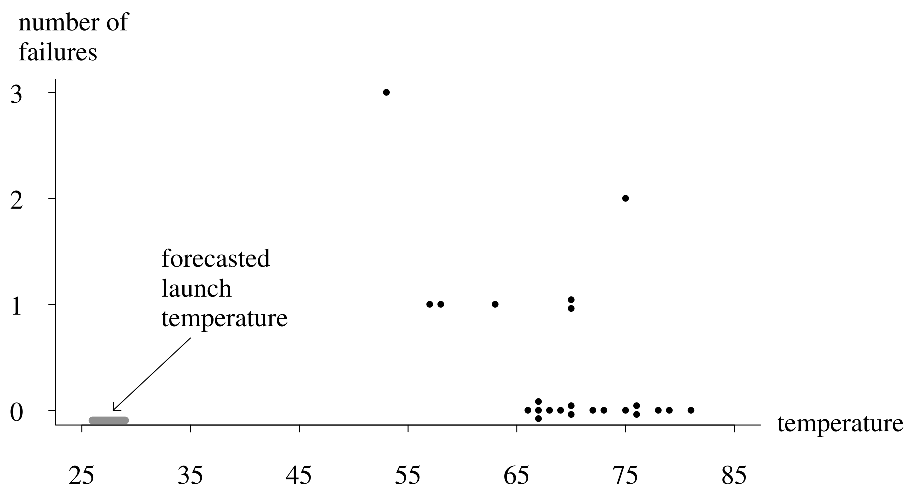


Figure 1.3: Launch temperature ($^\circ F$) versus number of field joint failures.

## 1.3 Case Study: Airline Safety

The previous section considered a reliability case study which ended in a catastrophe—the *Challenger* accident. The case study considered in this section concerns reliability improvement over time in the airline industry. The purpose of presenting this second case study is to emphasize that the mathematically-based probability models and statistical methods presented in this book do not capture all that is necessary to design and deploy a high-reliability system. These methods often need to be combined with management, organizational, and cultural factors that also play a crucial role in improving reliability.

This section contains a case study in reliability growth concerning worldwide airline travel, with a particular emphasis on the effect of a voluntary safety reporting program that accelerated in the United States in the mid-1990s.

Measuring reliability in airline travel is nontrivial. Examples of questions that need to be asked when quantifying reliability include the following.

- Are only commercial flights being considered? Should non-commercial flights also be included in the analysis?
- Is the measured quantity simply death on a flight or is injury included?
- What is the time window under consideration?
- If a flight with 200 passengers aboard has a crash, how do we factor in whether 1 person dies or all 200 die?
- Should deaths of people on the ground from a crash be considered?
- Is the scale of the measure of performance stated in terms of the number of miles traveled or in fraction of flights with no fatalities?
- Should a deliberate act against an aircraft, such as a terrorist attack, be included in the statistics?

Professor Arnold Barnett from the Massachusetts Institute of Technology has spent many years thinking about how to quantify aviation safety, and in a 2020 journal article (referenced in the Further Reading section of this chapter) he concluded that, although imperfect, the *death risk per boarding* is one of the best metrics to capture airline safety. We will state this metric in terms of reliability per boarding in order to be consistent with Definition 1.1. The metric can be thought of as an answer to the following question: if a boarding pass to a scheduled commercial airline flight under the time window of interest is selected at random, what is the probability that the bearer of this boarding pass will survive the air journey? The air journey begins upon arrival to the airport of origination and ends upon departure from the destination airport.

Barnett chose a time window of a decade in his analysis. This time window is long enough so that the usual month-to-month or year-to-year statistical random sampling variability is damped considerably, and yet short enough so that reliability growth patterns due to technology and safety improvements that occur over time are apparent. Figure 1.4 shows the reliability per boarding metric associated with a passenger in each decade. There is substantial reliability growth that has occurred decade-by-decade. The death per boarding decreased during the time period under consideration by a factor of 23: it decreased from 1 in 350,000 between 1968 and 1977 to 1 in 7.9 million between 2008 and 2017. Stated in another fashion, the last point plotted in Figure 1.4 is 23 times closer to the elusive reliability of 1 than the first point plotted. These numbers are based on worldwide data. The death per boarding statistic varies considerably by country; Barnett found that the death per boarding between 2008 and 2017 ranged from 1 in 28.8 million for traditional first-world countries to 1 in 1.3 million for less-developed countries.

What factors account for such a dramatic increase in reliability? Andy Pasztor covered aviation for the *Wall Street Journal* and wrote an article that describes the factors associated with the spectacular reliability growth achieved in the commercial airline industry in the United States. After some high-profile airline crashes in the mid-1990s (for example, ValuJet Flight 592 which crashed in the Florida Everglades on May 11, 1996 from a fire in a cargo compartment and TWA Flight 800 which crashed in the Atlantic Ocean on July 17, 1996 most likely from an explosion in a fuel tank), federal regulators, industry executives, trade associations, and pilots-union leaders came together to

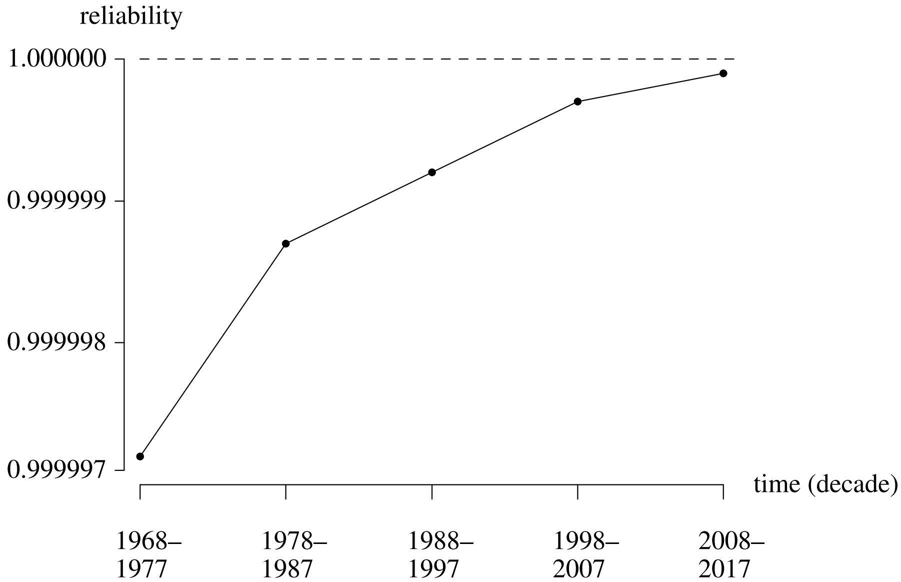


Figure 1.4: Worldwide commercial airline reliability per boarding by decade.

address the common problem of airline safety. They joined an existing voluntary incident reporting program in which all carriers would share data and there would be no punishment for airlines or pilots for good-faith mistakes or procedural violations. This approach is in direct opposition to human nature. It is our nature to attempt to cover up an error that we have made or to deflect the blame to someone else. It has been this way since the beginning (Genesis 3:6–13). This is particularly acute in the airline industry because of the extreme consequences of an error. Nevertheless, the approach worked. In the early days of the program, there were doubts that mistakes (assuming no willful disregard for safety) that were identified would be accepted without any disciplinary consequences. After a few years, however, all parties involved were on board and the necessary steps were taken to implement the program. Such a program requires trust among all participating parties, and the sharing of information on crashes, near-crashes, procedural violations, etc. resulted in a database that allowed experts to determine factors that would result in reliability growth such as enhanced pilot training, improved aircraft design for reliability and maintainability, identifying and improving key high-reliability aircraft components, and enhanced screening procedures. Their conclusions were based on analyzing flight data and reports filed by pilots, mechanics, air-traffic controllers, flight attendants, etc. Personnel were praised for identifying and reporting mistakes. Alternatively, if a mistake was not reported, you could lose your job. Although air safety has improved worldwide because of aircraft safety improvements, reliability gains differed from one country to the next. The voluntary reporting system developed in the United States aided their rapid gain in reliability.

Examples of factors that came into play over the years of voluntary reporting that resulted in reliability growth include the following.

- Nearly every large commercial aircraft contains two nearly-indestructible “black boxes” (they are actually orange and are named the flight data recorder and the cockpit voice recorder) that record flight information that can be helpful in determining mechanical and/or personnel errors that led to problems.

- The NTSB (National Transportation Safety Board) scrutinizes every crash under its jurisdiction; near-crashes may also provide equally-valuable data.
- After the terrorist attacks on the United States on September 11, 2001, passenger screening was federalized, leading to more uniform screening procedures from one airport to the next. In addition, the door between the cockpit and the cabin was reinforced on all commercial aircraft.
- Cockpit technology improvements, such as a terrain awareness and warning system that alerts pilots when they are flying dangerously close to terrain, increased reliability.
- Realistic ground-based simulators result in improved ability of pilots to encounter and react to difficult situations.
- Parallel redundancy in procedures came into play. As an example, two pilots were often required to physically point to cockpit computers while they double-checked out loud that flight-control information had been entered properly.

Technology improvements have played a significant role in the reliability growth in the airline industry. Automated systems can now determine if an aircraft is approaching an incorrect runway, help prevent ice buildup, help avoid midair collisions, warn a pilot if the aircraft is descending too rapidly, or is landing too far down the runway to brake properly. Cockpit radars can be used to avoid turbulence. Automation can also produce potential problems. Officials are concerned that overreliance on autopilots can result in a degradation of pilot skills in case automated systems fail.

One instance of the growth in reliability because of the voluntary data-driven approach came as a result of deadly crash on August 17, 2006. Comair Flight 5191 was assigned to take off from Runway 26 but inadvertently took off from the shorter Runway 22, which did not give the aircraft enough speed to become airborne. The deadly crash killed everyone but one crew member. The National Transportation Safety Board concluded that the crash was due to pilot error. The Federal Aviation Administration inspected databases of comparable hazards of this nature from pilot reports of similar runway confusion at other fields, and ordered improved signage and tarmac warnings.

The voluntary safety reporting program has been central to the reliability growth of the dozens of factors required for safe airline transportation. Other industries should follow this template.

## 1.4 Text Overview

This text is divided into two sections. Chapters 2 through 6 discuss probability models for lifetimes, and the remaining chapters discuss statistical methods related to data collection and inference.

Chapter 2 defines the object of interest in a reliability study as a *system*, which is assumed to be a collection of *n components*. The structure function determines whether the system is functioning, given the states of the components. Two of the simplest arrangements of components are series and parallel systems. A series system functions if all of its components function. A parallel system functions if one or more of its components function. The design for the O-rings on the Space Shuttle described in the previous section is a combination of series and parallel arrangements. Block diagrams of three-component series, parallel, and complex arrangements of components are shown in Figure 1.5. The reliability function, on the other hand, is used to determine the system reliability, given the component reliabilities. These calculations are often known as *static* reliability calculations because the item is considered at only one particular point in time (for example, the mission time).

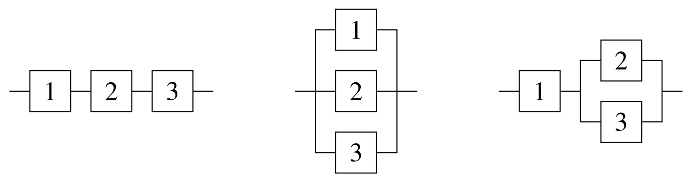


Figure 1.5: Block diagrams of three-component series, parallel, and complex systems.

Chapter 3 generalizes the concepts introduced in Chapter 2 as random lifetimes are considered. In particular, five different representations for the distribution of the failure time of an item are considered: the survivor, density, hazard, cumulative hazard, and mean residual life functions. The survivor and hazard functions are arguably the two most popular of the five because they express the probability of surviving to time $t$ and the risk at time $t$, respectively. A sample survivor function and corresponding hazard function are shown in Figure 1.6. This particular hazard function shape is known as a *bathtub-shaped hazard function* and is appropriate for modeling human lifetimes. In addition, many mechanical items also follow a bathtub-shaped hazard function. These five representations apply to individual components, as well as systems of components. These calculations are known as *dynamic* reliability calculations because system reliability is now a function of time.

Chapter 4 investigates several popular parametric models for the lifetime distribution of an item. The *exponential distribution* is examined first due to its importance as the only continuous distribution with the *memoryless property*, which implies that $(a)$ a used item that is functioning has the same failure time distribution as a new item, and $(b)$ failures are due to random loads or stresses which are beyond the item's intended design in terms of strength or capacity. Just as the normal distribution plays a central role in classical statistics due to the central limit theorem, the exponential distribution is central to the study of the distribution of lifetimes because it is the only continuous distribution with a constant hazard function. Figure 1.7 shows the survivor function and hazard function for a lifetime of an item with the exponential distribution with failure rate $\lambda = 1$. The rightmost survivor function is that of the remaining lifetime of the item conditioned on survival to

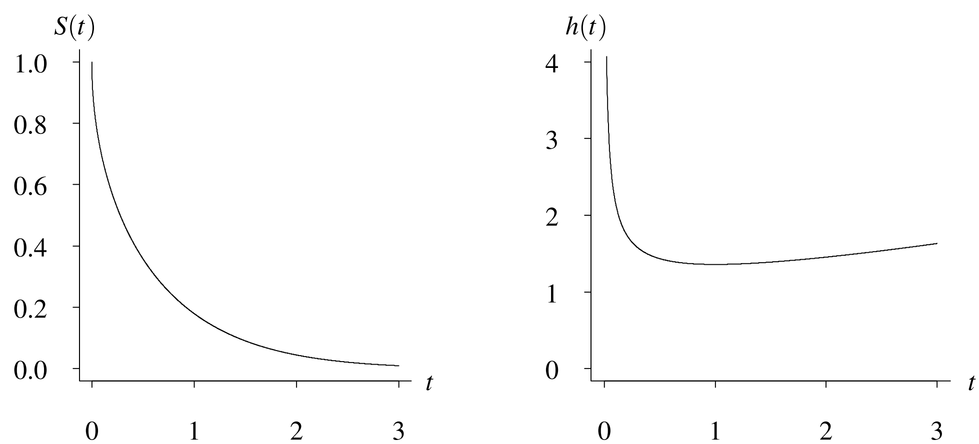


Figure 1.6: Survivor and hazard functions.

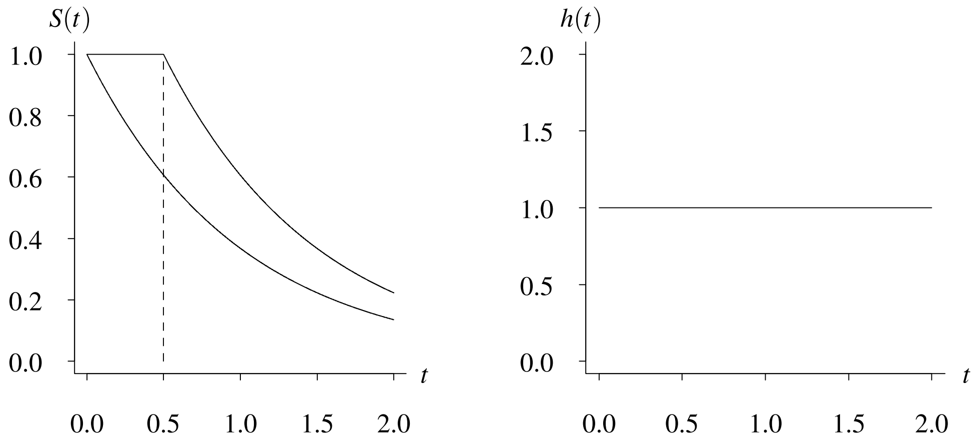


Figure 1.7: Survivor and hazard functions for the exponential distribution.

time 0.5. The fact that the *conditional survivor function* of the item that has survived to time 0.5 has the same shape as that of a new item is graphical evidence that the exponential distribution has the memoryless property. The exponential distribution has the drawback that its ability for modeling the lifetime of a component or system is limited. Its application in reliability is mainly limited to electrical and electronic components. More flexible lifetime distributions, such as the Weibull and gamma distributions, are also outlined in this chapter. Each of these distributions has mathematical properties that make it appropriate for modeling certain types of failure mechanisms, such as fatigue or wear out. These distributions have multiple parameters that make them useful in a wide variety of lifetime modeling situations.

Chapter 5 presents probabilistic models that lack many of the limitations inherent in the parametric lifetime models discussed in Chapter 4. In particular, most of the lifetime distributions presented in Chapter 4 have hazard functions that are monotonic (that is, increasing or decreasing for all $t$). Competing risks and mixtures are two models that can easily achieve a nonmonotonic hazard function; in addition, they can be applied to situations involving multiple causes of failure and multiple populations of items. The hazard functions in Figure 1.8 for a calculator show three competing risks for the lifetime of the calculator: manufacturing defects, dropping the calculator, and wear-out failures. Examples of components of a calculator that experience wearout include the buttons, the display, and the battery. Covariates are exogenous variables associated with the item of interest (for example, the age or gender of a person in a medical setting) that might influence the distribution of a lifetime. A single lifetime probability distribution should not be used to model the failure time of a circular saw blade that operates at one of three different turning speeds. Two models that are useful ways to include these covariates in a lifetime model are the accelerated life and proportional hazards models. In the Space Shuttle *Challenger* case study, the temperature at the launch of the Space Shuttle was determined to be a covariate that influenced the survival of the original design of the O-rings.

Chapter 6 introduces mathematical models for *repairable systems*. Many systems, such as aircraft, have repairable components. Failure or degradation of these components might $(a)$ be detected by a condition-based monitor, $(b)$ be detected in routine inspection and maintenance, $(c)$ be detected during an overhaul of the aircraft, or $(d)$ be the cause of catastrophic failure. These models are ef-

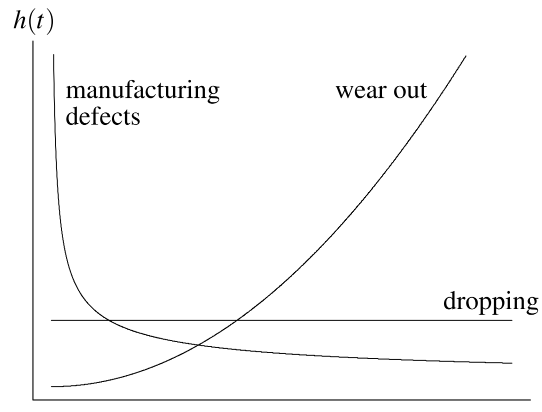


Figure 1.8: Competing risks for a calculator.

fective for determining the *availability* of a system. The uptime and downtime associated with a repairable system are now explicitly modeled. Figure 1.9 shows graphs associated with the availability of an aircraft; state 1 denotes available and state 0 denotes unavailable. The left-hand graph is associated with a single realization of the repair record for one particular aircraft, and the right-hand graph is a probability model that might be associated with this type of aircraft. The left-hand graph shows the times when the aircraft is available. For this particular realization, there are three points in time when the aircraft goes from state 1 to state 0. The graph of the availability function $A(t)$ shows the point availability approaching a steady-state limiting availability of 0.615 as $t$ increases. This implies that, although the availability of a new aircraft is typically higher initially, it eventually approaches a constant value.

The emphasis changes from developing probabilistic models for lifetimes to analyzing lifetime

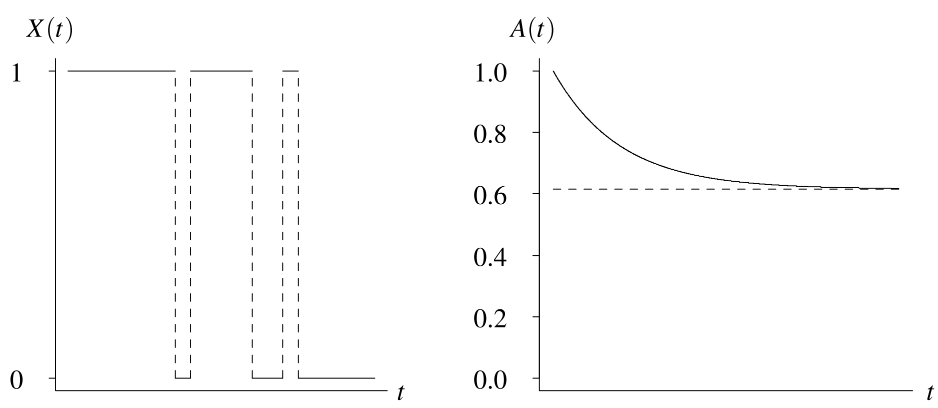


Figure 1.9: Aircraft availability graphs.

data sets in Chapter 7. One problem associated with these data sets is that of *censored data*. Data are censored when only a bound on the lifetime is known. This would be the case, for example, when conducting an experiment with light bulbs, and half of the light bulbs are still functioning at the end of the experiment. Figure 1.10 shows the results of a life test of five light bulbs. Failure times were observed for the first, second, and fifth light bulbs, indicated by an $\times$, and the other two are *right-censored* data values, indicated by a $\circ$. The third and fourth light bulbs will produce a failure time eventually, but all we know from this data set is that their failure times exceed a threshold. Censored observations complicate the analysis of a data set. The properties of point and interval estimates for a parameter are reviewed in this chapter.

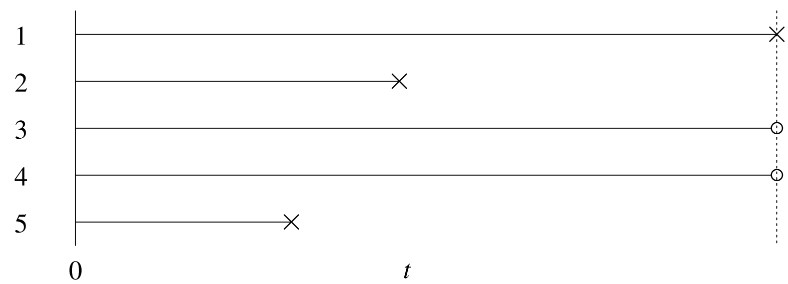


Figure 1.10: A right-censored data set.

Chapter 8 surveys methods for fitting parametric distributions to data sets. Maximum likelihood parameter estimates are emphasized because they have certain desirable statistical properties. Figure 1.11 shows two graphs that are useful in the analysis of a data set of failure times. The graph on the left shows the empirical survivor function and the fitted Weibull survivor function for an uncensored data set. The empirical survivor function [also known as the nonparametric estimator of $S(t)$] is a step function that decreases by $1/n$ at each data value, where $n$ is the sample size. The data

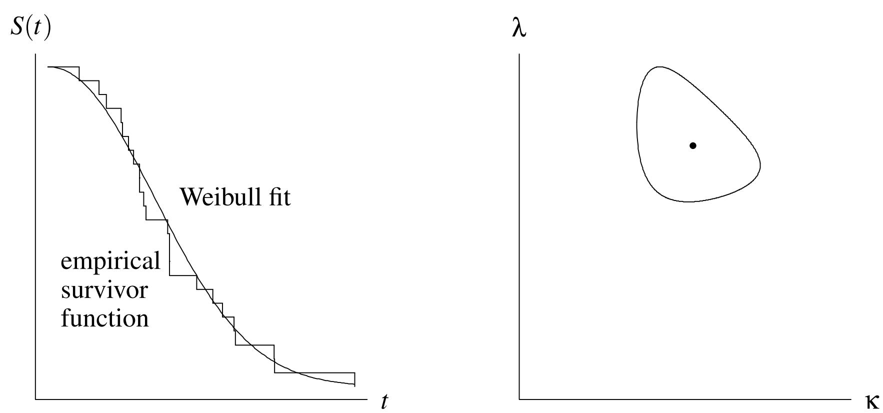


Figure 1.11: Fitting a parametric distribution to a lifetime data set.

are fitted to the Weibull distribution with scale parameter $\lambda$ and shape parameter $\kappa$ by maximum likelihood and the fitted survivor function is the smooth curve. In the graph on the right, a 95% confidence region for the parameters $\lambda$ and $\kappa$ indicates the precision of the estimates. The point in the center of the confidence region denotes the maximum likelihood estimators $\hat{\kappa}$ and $\hat{\lambda}$.

Chapter 9 presents techniques that can be used to estimate the effects of covariates on a particular lifetime distribution. In particular, the accelerated life and proportional hazards models have regression coefficients that can be estimated for data sets containing covariates. Maximum likelihood methods are used to estimate these coefficients. Figure 1.12 shows the hazard functions for the lifetime of a drill bit operating at three different turning speeds. The single covariate in this case is the turning speed of the drill bit. As expected, the risk is higher at increased turning speeds. These estimation techniques can be used for several covariates, and interactions between covariates can also be modeled.

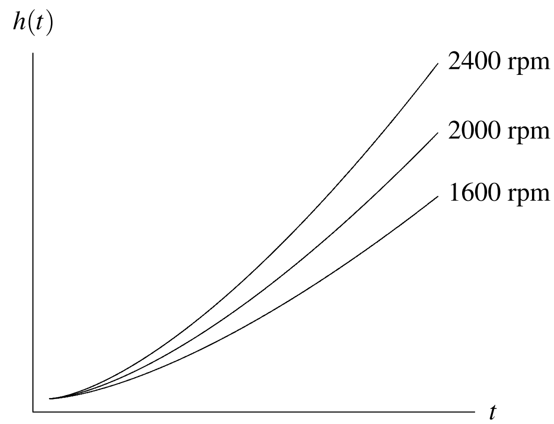


Figure 1.12: Hazard functions for several turning speeds of a drill bit.

Chapter 10 introduces nonparametric methods for estimating the distribution of the lifetime of an item. Nonparametric methods are often used when there is no parametric distribution to accurately model the lifetime. The two methods considered in this chapter are *life tables*, which are used by actuaries, and the Kaplan–Meier product–limit estimate, which is used to estimate the survivor function from a right-censored data set. An example of a product–limit estimate for a survivor function from a censored data set is shown in Figure 1.13. The downward steps in the survivor function correspond to observed failure times and the hash marks on the survivor function denote right-censoring times. The dotted lines correspond to point-wise approximate two-sided 95% confidence limits for the survivor function for all time values. Nonparametric methods for repairable systems are also presented.

Chapter 11 discusses statistical techniques that can be used to assess model adequacy. Once a parametric model has been chosen to represent the failure time distribution for a particular item, the adequacy of the model should be assessed. Goodness-of-fit tests assess how well a fitted lifetime distribution models the lifetime of the item. One popular test for continuous lifetime models is the Kolmogorov–Smirnov test, in which the largest vertical difference between the fitted and empirical survivor functions is used as the test statistic. Figure 1.14 shows the empirical survivor function estimate for a data set of $n = 23$ ball bearing failure times (in millions of revolutions) as a decreasing step function. The smooth curve is the survivor function associated with fitting the exponential

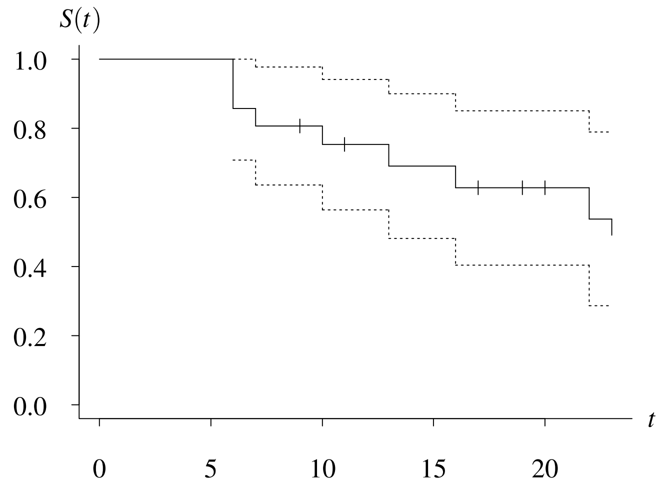


Figure 1.13: Kaplan–Meier product–limit survivor function estimate.

distribution to the data. The largest vertical difference between the empirical and fitted distribution, $D_{23}$, is indicated in the figure. The exponential distribution does a poor job of modeling the ball bearing failure times, as evidenced by the large value of $D_{23}$. A two-parameter distribution, such as the Weibull or gamma distribution, provides a much better fit. The Kolmogorov–Smirnov test can be used to test the adequacy of any fitted probability distribution to a data set of lifetimes. A goodness-of-fit test associated with a particular lifetime distribution, such as a test for exponentiality, often has higher power.

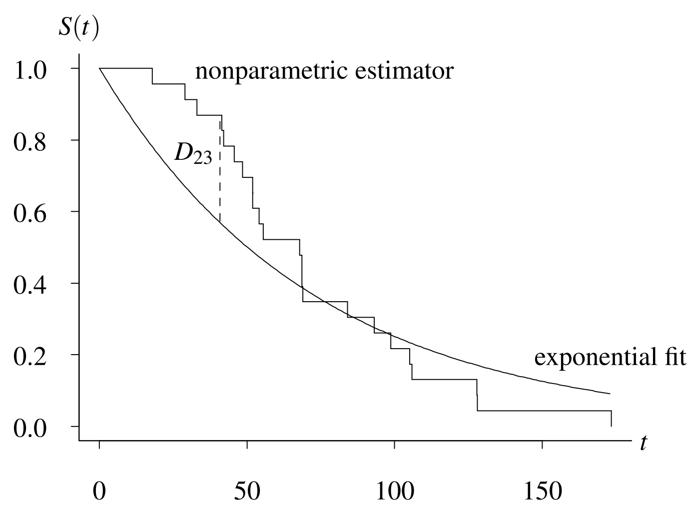


Figure 1.14: Kolmogorov–Smirnov test statistic for the ball bearing failure times.