CHAPTER 14

Reliability Analysis and Life Testing

The terms "survival data," "reliability data," and "lifetime data" are all quite similar. The first term is generally applied to people, such as survival data in years after treatment for cancer patients who follow a specified treatment. The second and third terms refer to electrical or mechanical components, products, or systems but are used in somewhat different ways. For example, we might refer to the reliability of an electrical system, whereas we would be most concerned with the lifetime, not the reliability, of a light bulb since light bulbs function normally until they burn out and are not repaired.

As consumers, we are all interested in product reliability. Similarly, manufacturers are interested in product reliability, which directly relates to warranty costs. Various definitions of reliability have been given, but the one stated below should suffice.

DEFINITION

Reliability is the probability of a product fulfilling its intended function during its life cycle.

A more elaborate definition, which says essentially the same thing, was given by Barlow (1998): "*Reliability* is the *probability* of a device performing its purpose adequately for the period of time intended under operating conditions encountered."

As contrasted with quality control and quality improvement, we might say that a product "has quality" if it is still functioning as intended at a given point in time and the extent of its useful life is what the public considers acceptable. Not everyone has come to view quality as extending beyond the point of manufacture, however, as there are still some outdated views of quality. For example, Rausand and Høyland (2004, p. 6) stated, "Reliability is therefore an extension of quality into the time domain."

Reliability theory and methodology developed as a tool to help nineteenth century maritime and life insurance companies compute profitable rates to charge their customers (*NIST/SEMATECH e-Handbook of Statistical Methods*, Section 8.1.2). Today we would

Modern Engineering Statistics By Thomas P. Ryan

refer to this as survival analysis since people are involved, but the general principles of reliability analysis and survival analysis are certainly related.

To the layperson, product reliability probably means "when can I expect product X to give out completely, at what earlier point in time can I expect to have to start making repairs, and once the repairs begin, how slowly or how rapidly can I expect the product to deteriorate?" This, in essence, is what we mean by reliability. In other words, reliability essentially refers to a continuum of time, whereas there is no mention of "rate of degradation," or similar terminology, in any common definition of quality. The possible use of degradation data in designed experiments is discussed in Section 14.8 and using degradation data in conjunction with acceleration models is discussed in Section 14.3.1.

14.1 BASIC RELIABILITY CONCEPTS

Reliability data are measurements of the time to failure, with possibly other measurements made on sample items. One goal in reliability work is to estimate the percentage of units of a product that will still be functioning after a particular period of time. This could then be translated into a probability statement, for example, "the probability is .24 that a newly designed light bulb will last longer than 25,000 hours." In order to be able to make such a statement, it is necessary to select a probability distribution that will facilitate the construction of reasonably precise probability statements of the type that one wishes to make.

Probability distributions that have frequently been used in reliability work include the Weibull distribution, which was covered in Section 3.4.6. Early work in reliability utilized the exponential distribution to a considerable extent, but it was eventually realized that the Weibull distribution generally provides a much better model. The exponential distribution does have some important applications, however, as was discussed in Section 3.4.5.2. (See the discussion of the Weibull distribution vis-à-vis the exponential distribution in Section 14.5.3.) A good discussion of the exponential distribution and its limited usefulness in reliability work—though it is still widely used—can be found at `http://www.reliasoft.com/newsletter/4q2001/exponential.htm`. Limitations of the exponential distribution in reliability applications are discussed in Section 14.5.2.

One restriction generally imposed on the selection of a distribution is that the distribution must be defined only for positive values, since the random variable is time, which of course cannot be negative. This doesn't mean that distributions for which the random variable can be negative cannot be used in reliability applications, or other applications, however.

In particular, this alone would not rule out the normal distribution as a possible model because there are many random variables (e.g., height, weight, aptitude test scores) that cannot be negative, yet a normal distribution (which of course has a range of $-\infty$ to $+\infty$) is a good model. As long as the mean is well above zero, this may not be a problem. A more important issue is whether the random variable can be expected to be skewed, and we certainly expect failure times to be skewed.

Since reliability data are, at least for the most part, *failure data*, the definition of a "failure" in a given application must be carefully considered. In the example cited previously, is it reasonable to say that a light bulb has failed if it hasn't lasted for 25,000 hours? Often the threshold value for determining failure is chosen subjectively. In many applications, however, an item does not need to have ceased functioning for the item to be declared a

"fail." Specifically, failure might be declared if an item can no longer deliver a specified proportion of its functionality. For example, if a TV picture tube is becoming weak, there may still be a picture, but the picture quality is probably unacceptable. Similarly, a cable TV connection may be poor at the company's end, but the picture may be viewable.

In both cases we would have what is termed a *repairable system*. An item that is repairable, such as a car, may eventually last a long time. Items that are not repairable would include a light bulb that has burned out and an extremely old car that is "beyond repair." These would obviously be termed *nonrepairable systems* and there is a different set of statistical methods applicable to such systems. Most reliability data are collected on nonrepairable populations, rather than on repairable populations, for which units that fail can be repaired and returned to the population.

The term *failure mode* is also used in reliability. A failure mode is simply a cause of failure for a component or unit. There can be different causes of failure and each failure mode has a probability distribution. A *failure mode and effects analysis* (FMEA) is often conducted for the purpose of outlining the expected effects of each potential failure mode. This consists of examining as many components, assemblies, and subsystems as necessary to identify failure modes (i.e., how faults are observed) and the causes and effects of the failures, noting that each failure mode can be caused by several different failure causes.

An example of an FMEA is given in Section 14.7. A related concept is a *failure mode*, *effects*, *and criticality analysis* (FMECA), which occurs if priorities (criticalities) are assigned to the failure mode effects.

Under the *competing risk model*, the component or unit fails when one of the possible causes of failure (i.e., failure mode) occurs. Each failure mode thus has its own reliability, $R(t)$, defined in the next section, with the reliability for the component or system under the competing risk model being the product of the reliabilities for the failure modes, and the component or system failure rate being the sum of the individual failure mode failure rates. The latter holds for models in general as long as independence holds, and as long as component failure results when the first mechanism failure occurs.

These results also hold for a *series system*, for which the system fails if any one of its components fails. In contrast, a *parallel system* is one for which the system continues to operate until *all* of the components fail. In between these two extremes is an "r out of n" system, which will continue to function as long as any r out of n components are functioning. See Section 8.1.8.4 of the *NIST/SEMATECH e-Handbook of Statistical Methods* for additional information about this type of system.

Which of these models is applicable of course depends on how the system was constructed.

In many applications it isn't practical to wait until failures occur. For example, a piece of machinery may have some critical components whose reliability needs to be quickly and precisely assessed, as the reliability of the machinery strongly depends on the reliability of these components. Such a situation would almost dictate the use of *accelerated testing*, which would be performed in a laboratory. In particular, it isn't possible to wait ten years or so until an automobile gives out to determine whether or not it functioned as designed during its useful life. As discussed by Fluharty, Wang, and Lynch (1998), automotive engineers are designing products with anticipated lives that exceed ten years and 150,000 miles, so accelerated testing is used to simulate high mileage and aging. The failure rate in an accelerated test would then have to be converted to a failure rate under normal use conditions, and this would entail the use of models. Clearly, the modeling aspect is important because the results of accelerated tests are of little value if the results cannot be converted to normal use failures.

14.2 NONREPAIRABLE AND REPAIRABLE POPULATIONS

By a "nonrepairable population" we mean one for which units that fail are not repaired. Models that are frequently used for nonrepairable populations include the exponential, Weibull, gamma, lognormal, and extreme value models given in Chapter 3, in addition to a few other models (see Croarkin and Tobias, 2002, Section 8.1.6). Models that are used for repairable systems include the homogeneous Poisson process, with the cumulative number of failures to time T given by the Poisson distribution with parameter λT. (The Poisson distribution was covered in Section 3.3.4.) This model is widely used in industry but the most commonly used model is the Duane model. Duane (1964) presented failure data of different systems during their development programs and showed that the cumulative *mean time between failures* (*MTBF*) plotted against cumulative operating time was close to a straight line when plotted on log–log paper. This model is also often called the *power law model* and has also been referred to as the Army Materials System Analysis Activity model.

The *pdf* of the model can be written

$$P(N(T) = k) = \frac{a^k T^{bk} e^{-\lambda t}}{k!}$$

with $N(T)$ denoting the cumulative number of failures from time 0 to time T, a and b are constants, and λ is the parameter to be estimated. [See Croarkin and Tobias (2002, Section 8.1.7.2) for further details.]

Most books on reliability are books for nonrepairable systems, which is understandable because most reliability data that are collected are on nonrepairable systems. A notable exception is Rigdon and Basu (2000), which is devoted to repairable systems. See also Section 8.1.7 of Croarkin and Tobias (2002), which discusses some of the models that are used for repairable systems.

14.3 ACCELERATED TESTING

An accelerated failure model relates the time-to-failure distribution to the stress level. The general idea is that the level of stress is only compressing or expanding time and not changing the shape of the time-to-failure distribution. With this assumption, projecting back to stress under normal use is not as difficult as it might seem as one has to simply handle the change of scale in regard to the time variable. That is, changing stress is equivalent to transforming the time scale used to record the time at which failures occur, so it is a matter of "decelerating" back to normal use by using an appropriate acceleration factor. This is accomplished by using *acceleration models*. These models are mechanistic models that are based on the physics or chemistry that underlies a particular failure mechanism. Acceleration models of course use time as a variable but generally do not contain interaction terms. Models without such terms seem to work well.

The general idea is that for two stress levels, $S_2 > S_1$, the time to failure at stress level S_1 can be represented as

$$T(S_1) = a(S_2, S_1)T(S_2) \tag{14.1}$$

with $T(S_i)$ denoting the time to failure at stress S_i, $i = 1, 2$. In order to be able to project back to obtain $T(S_1)$, the acceleration rate $a(S_2, S_1)$ must obviously be known. In order to easily solve for $a(S_2, S_1)$ for a specified model, such as the Arrhenius equation presented in Section 14.3.1, it is necessary to use the equivalent form

$$\log[T(S_1)] = \log[a(S_2, S_1)] + \log[T(S_2)] \tag{14.2}$$

If we let $\log[T(S_1)]$ correspond to the probability of failure by time t with parameters $\mu(S_1)$ and $\sigma(S_1)$, it follows that

$$\mu(S_1) = \log[a(S_2, S_1)] + \mu(S_2) \tag{14.3}$$

and $\sigma(S_1) = \sigma(S_2) = \sigma$. In words, Eq. (14.3) states that the mean time to failure at stress S_1 is equal to the log of the acceleration rate plus the mean time to failure at stress S_2.

In order for accelerated testing to be successful, an acceleration factor must be identified, if one exists. Unfortunately, acceleration factors don't always exist. For example, Wu and Hamada (2000, p. 548) pointed out that a factor for accelerating the failure of hybrid electronic components has not been found. There are some acceleration factors that do work well, however, including temperature and voltage.

There are also some major obstacles to the successful use of accelerated testing. The first is that the relationship between time to failure and stress must be modeled, and it may be difficult to check the adequacy of the model. In essence, what is needed is a mechanistic model (Chapter 10), and as Hooper and Amster (2000) pointed out, one must understand the physics of failure.

An obvious question (or two) to ask is: Why is a model necessary? Can't we just obtain all of the data that we need from accelerated testing? The answer is "no" because most of the data will come from high-stress testing, so it is necessary to extrapolate the failure results at high-stress levels to lower levels, including "no stress."

Since accelerated testing is performed because of relative scarcity of failure data under normal stress conditions, we would like to have some assurance that accelerated testing will be successful. A simulation procedure for arriving at a test design that will increase the chances of success is given in Chapter 8 of the *NIST/SEMATECH e-Handbook of Statistical Methods* (Croarkin and Tobias, 2002), and specifically at `http://www.itl.nist.gov/div898/handbook/apr/section3/apr314.htm`.

Even though selecting an appropriate accelerated test model can be difficult and model validation even more so, there is some "history" that can be relied on in zeroing in on an appropriate model. For example, Spence (2000) points out that the Arrhenius model has been used to calculate acceleration factors for semiconductor life testing for many years. This model is considered in the next section.

14.3.1 Arrhenius Equation

We would expect that as temperature is increased, a reaction that is affected by temperature, such as a chemical reaction, will occur more rapidly. The Arrhenius equation embodies this idea and is named after S. A. Arrhenius (1859–1927). The equation when applied to

accelerated testing is given by

$$\mu(S) = \log(R_0) + \frac{E_a}{k(S + 273)} \tag{14.4}$$

with R_0 denoting the rate constant, E_a is the activation energy, k is Boltzmann's constant, S is the stress (temperature in degrees Celsius in the Arrhenius equation), and $\mu(S)$ is a function of the time to failure at stress S. Equation (14.4) does not by itself give the acceleration rate, which is obtained by using two temperatures, S_1 and S_2. That rate must be solved for by using Eqs. (14.3) and (14.4), and doing so produces

$$a(S_2, S_1) = \exp\left[\left(\frac{E_a}{k}\right)\left(\frac{1}{(S_1 + 273)} - \frac{1}{(S_2 + 273)}\right)\right] \tag{14.5}$$

as the reader is asked to verify in Exercise 14.3. The value obtained from this equation after specifying S_1 and S_2 would then be used in Eq. (14.1) to solve for (estimate) $T(S_1)$.

Regarding computation, Section 8.4.2.2 of the *NIST/SEMATECH e-Handbook of Statistical Methods* shows how to use JMP for Arrhenius parameter estimation.

14.3.2 Inverse Power Function

The Arrhenius equation is for temperature as the accelerating variable. When voltage is the acceleration variable, the inverse power function is frequently used, and the use of the function leads to

$$\mu(S) = \beta_0 + \beta_1 \log(S)$$

and an acceleration factor of

$$a(S_2, S_1) = \left(\frac{S_1}{S_2}\right)^{\beta_1}$$

Since $S_2 > S_1$, β_1 must obviously be negative since the acceleration rate must be greater than one. Of course, β_0 and β_1 would have to be estimated, and the method of maximum likelihood is a common choice for this purpose.

As in regression analysis, once the parameters are estimated, residual analysis would be used to check the fit of the model. Of course, in regression analysis normality is assumed, but normality is generally not part of reliability work. Therefore, a probability plot of the residuals would be constructed using the distribution that is assumed under the model that is fit, with the objective to see if the points plot on a straight line. If they do so, this would validate the implied distributional assumption, which in turn would essentially validate the selected stress function.

14.3.3 Degradation Data and Acceleration Models

It is possible to predict when failures will occur if failure can be related directly over time to changes in the values of a measurable product random variable. Then degradation data can be used without having to wait for failures, with failure predicted by extrapolating from the degradation data. In order for this to work, a variable must be identified that changes slowly over time and eventually reaches a value that causes the unit to fail. Unfortunately,

it is often difficult or impossible to find such a variable. Consequently, although the use of degradation data is an attractive alternative to having to wait for failures, it could be difficult to implement the idea.

If such a variable is found and the approach is used, it is highly desirable to obtain an actual failure or two to compare the actual failure time with what is predicted using the acceleration model, and thus obtain a "residual" or two that can be used for model criticism.

For more information on the use of degradation data, see Section 8.4.2.3 of Croarkin and Tobias (2002). For additional information on accelerated testing the reader is referred to Nelson (1990).

14.4 TYPES OF RELIABILITY DATA

Answers to questions about the reliability of a particular product are complicated by the fact that we generally cannot take, say, 100 units of product X and see how long it takes for each unit to become inoperable because of practical considerations. For example, let's say we start with 100 light bulbs and after a certain length of time, 98 of these have burned out. We wait for the other two bulbs to burn out . . . and wait . . . and wait, but the bulbs defy us and continue functioning. So we grow weary of this waiting game and terminate the test. What have we done? For one thing, we have created *censored data*.

Now what do we do since we don't have data on all 100 light bulbs? We could compute the average lifespan for the 98 bulbs, but we would be underestimating what the average would have been if we had the time and patience to wait for the other two light bulbs to burn out.

This scenario is typical of what occurs with life testing in that the data are censored. That is, we do not have measurements on every item in the sample. We thus have to estimate "what would have been" so we need statistical techniques that are different (and more involved) than what has been presented in previous chapters.

Not surprisingly, software must be used to handle this situation. We recognize the added complication of censoring at this point in the chapter, but we will wait until near the end of the chapter to return to it after first considering more simplified scenarios.

Much has changed during the last forty years regarding views toward product reliability and desirable product lifespans. It was argued decades ago that it would be bad for the U.S. economy for products to last a long time as then consumers wouldn't need to purchase replacements. So it was viewed as desirable to have "built-in obsolescence" and to have products with not-too-long lifespans.

That view changed when it became apparent that Japanese cars, in particular, had a long lifespan, so consumers began to demand this. Now, in the early years of the twenty-first century, consumers have a much different view of product reliability and the desired lifespan of products. In particular, wouldn't it be nice if light bulbs lasted "almost forever" so that those of us who live in houses with high ceilings and light fixtures near the ceiling wouldn't have to risk life and limb to replace them? Perhaps with such considerations in mind, companies like General Electric, Cree Research, and Emcore have worked on building a better light bulb—one whose lifespan is a dramatic improvement over the light bulbs that we have used for decades. Subsequent research accomplishments will present new challenges for modeling product lifespan.

Unfortunately, product reliability generally commands the most attention when there is loss of life, as happened when the combination of Firestone tires on Ford Explorer vehicles

gained national attention during 2001 because of the deaths that were apparently caused by either the vehicle or the tires (or both), and which led to the dissolution of the long-standing business relationship between Firestone and Ford.

In this chapter we survey reliability and life testing, focusing more on concepts and applications than on theory. We also want to recognize that reliability is inextricably linked to nonconformities, as units with nonconformities at the beginning of use will likely have a shorter lifespan than those units without any nonconformities.

14.4.1 Types of Censoring

Since censored data predominate in reliability studies, it is important to understand the different types of censored data. There are three basic types of censoring: right censored, left censored, and interval censored, in addition to the exact failure time.

With *exact failure* data, the time that an item fails is recorded, and the testing does not terminate until the item has failed. Thus, there is "zero censoring" and obviously exact failure data is the easiest data with which to work. Unfortunately, we will not often encounter such data (exclusively) as this would suggest that the quality of what is being examined is probably poor.

With *right censoring*, an item is removed from test when it is still functioning, so the time that it would have failed if testing on the item had continued is unknown. When this type of censoring is used, one must be able to assume that items that are removed from test at a given time are representative of all of the items that are on test at that time. This type of censoring is also called *Type I censoring*.

With *left censoring*, items are removed from the test at a certain time if they are not functioning. The time of failure is of course unknown; all that is known is that it failed sometime between the start of the test and the time at which it was removed. This type of censoring is also called *Type II censoring*. This type of censoring is rarely used, however.

With *interval censoring*, an item that is put on test is found to be functioning at a certain time thereafter (say, t_l), and then found to be not functioning when checked at a later time (say, t_u). All that is known is that the item failed sometime between t_l and t_u. This type of data is sometimes referred to as "readout data."

14.5 STATISTICAL TERMS AND RELIABILITY MODELS

We will let $f(t)$ denote an arbitrary failure distribution, with t denoting time (of failure). Since time is continuous, we will consider only continuous distributions. Therefore, we will not be concerned with $f(t)$ for a given value of t, but rather will be concerned with $F(t)$, the probability of failure by time t, and simple functions of it. Although the data that are collected are failure data, we will naturally focus attention on the positive side of the coin: nonfailure. Accordingly, we define $R(t) = 1 - F(t)$ as the *reliability function*, which is the probability that a unit is still functioning at time t. (In survival analysis this is referred to as the *survival function*.) The *failure rate* is defined as $h(t) = f(t)/R(t)$, which is sometimes referred to as the instantaneous failure rate.

Depending on the situation, the mean time to failure (*MTTF*) can also be of interest. It is defined as $MTTF = \int_0^\infty t f(t)\, dt$. Since $f(t)$ in a particular application may be highly skewed, the median life, defined as t_m such that $R(t_m) = .50$, may sometimes be preferable. If *MTTF* is to be estimated, which of course would be necessary if $f(t)$ were not known, care

should be exercised in using sample data for the estimation as early failures could greatly overestimate *MTTF* if the time interval for the sample is not reasonably long. Conversely, the *MTTF* could be underestimated if early failures are rare but failures frequently occur just beyond the sampling interval.

The mean time between failures (*MTBF*), mentioned in Section 14.2, is also a useful measure for repairable systems, although its usefulness is limited by the fact that most systems are nonrepairable, as stated previously. It is estimated by the total time that a system has been in operation divided by the total number of failures. The general form of a confidence interval for the *MTBF* is given in many books on reliability [e.g., see, Von Alven (1964) or Dhillon (2005)].

The mean residual life—*MRL*(*t*)—is also sometimes used. It is defined as

$$MRL(t) = \int_o^\infty R(x|t)\, dx = \frac{1}{R(t)} \int_t^\infty R(x)\, dx$$

with $R(x|t)$ denoting the *conditional survivor function* at time t. The mean residual life can be written in terms of the failure rate, as has been shown by Gupta and Bradley (2003).

These are all important functions in reliability work.

DEFINITION

The *reliability function* is given by $R(t) = 1 - F(t)$ and is the probability that a unit is still functioning at time t. The failure rate, also called the *hazard function*, is given by $h(t) = f(t)/R(t)$ and measures susceptibility to failure as a function of the age of the unit.

In words, $h(t)$ is the probability of failure at time t divided by the probability that the unit is still functioning at time t. It has been stated that the failure rate is the rate at which the population survivors at any given instant are "falling over the cliff." As mentioned previously, however, we technically cannot speak of probability of failure at a particular point in time because we use only continuous time-to-failure distributions. Therefore, borrowing a calculus argument, we should view $h(t)dt$ as approximately equal to the probability of a failure occurring in the interval $(t, t + dt)$, given that the unit has survived to time t, with $f(t) = (d/dt)F(t)$.

This is somewhat analogous to saying that, at a particular point in time, a car is traveling 50 miles per hour. If the driver doesn't drive for an hour, she will not travel 50 miles. But she *would* have traveled 50 miles if she had maintained that constant speed for an hour, so her interval of driving time would be $(t, t + 1$ hour$)$.

We illustrate these concepts in the next few sections for some commonly used failure distributions.

14.5.1 Reliability Functions for Series Systems and Parallel Systems

As stated in Section 14.1, with a series system, failure occurs when one of the components of the system fails. If independence of the components is assumed, which may or may not be a plausible assumption, depending on the application, the system reliability is the

product of the component reliabilities: that is, $R_s(t) = \prod_i R_i(t)$, with $R_s(t)$ denoting the system reliability at time t and $R_i(t)$ denotes the reliability for the ith series component at time t.

Some series systems consist of identical parts, such as cells of a battery. For k identical, independent components, $R_s(t) = [R_i(t)]^k$.

Series systems do not always have independent components, however, and determining the system reliability for dependent components is generally complicated and will not be pursued here. The interested reader is referred to Nelson (1982, p. 172). One approach that will sometimes work for a complex multicomponent system is to try to reduce it to a single-component system with a known reliability function, or at least a known cumulative distribution function.

The system reliability for independent components can be illustrated as follows.

■ EXAMPLE 14.1

Light-emitting diodes (LEDs) are part of the move toward brighter and more efficient lights. Assume that a system consists of two different types of LEDs, one whose lifetime is believed to be about 12,000 hours with a standard deviation of 900 hours and to be (approximately) normally distributed, and for the other LED a normal distribution with a mean of 11,000 hours and a standard deviation of 800 hours is believed to provide a good fit. (The question of whether or not a normal distribution would seem to be a viable model for this application is addressed in Exercise 14.29.)

The reliability for the two-component system is given by $R_s(t) = [1 - F_1(t)][1 - F_2(t)]$. For example, the system reliability at $t = 13{,}000$ is given by $R_s(13{,}000) = [1 - F_1(13{,}000)][1 - F_2(13{,}000)] = (1 - .8665)(1 - .9938) = .0008$. Thus, it is highly unlikely that the system will be functioning after 13,000 hours. ■

For parallel systems, the system fails when the last component fails, so the probability that the system is still functioning at time t is equal to one minus the probability that all components have failed by time t: that is, $R_s(t) = 1 - [F_1(t)][F_2(t)] \cdots [F_k(t)]$ for a parallel system of k independent components, and thus $R_s(t) = 1 - [F_i(t)]^k$ when the distributions are identical.

14.5.2 Exponential Distribution

Although the exponential distribution is no longer as heavily used in reliability applications as was once the case, it is a base distribution that has illustrative value. From Section 3.4.5.2, if we let $\lambda = 1/\beta$ in Eq. (3.14), we obtain $f(t) = \lambda \exp(-\lambda t))$ for the exponential distribution. It follows that $R(t) = \exp(-\lambda t)$, so that $h(t) = \lambda$. Thus, the hazard rate is constant for the exponential distribution and does not depend on t. (The exponential distribution is the only distribution that has a constant failure rate.)

Thus, an unused item would have the same reliability as a used item. Does this sound reasonable? It obviously doesn't apply to cars and people, for example. If it did, an infant would be just as likely to die while an infant as a 95-year-old would be to die at that age. Similarly, a new car isn't going to fail at the same rate as a 15-year-old car. And so on.

The "bathtub" curve is generally considered to be a reasonable representation of the failure rate for many, if not most, products, and is shown in Figure 14.1. That is, if a

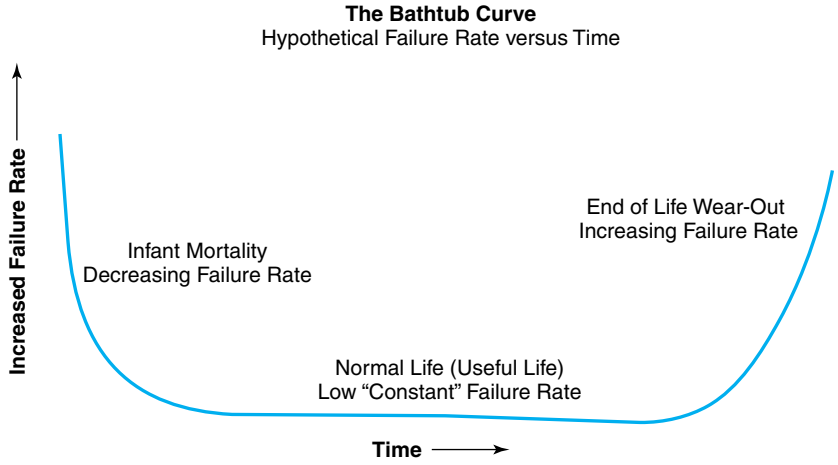

Figure 14.1 "Bathtub" curve of product life: hypothetical failure rate versus time.

very large sample of failure times was available for a particular product, a graph of the frequencies of the different failure times would at least roughly resemble the form of a bathtub. In words, the failure rate would be high when the product is first used (since the product could be defective); then the rate would level off and be flat for a long period of time, and finally the rate would increase as the product neared the end of its life.

Since this obviously sounds reasonable, we might restrict our attention to probability models whose failure rate would graphically correspond roughly to a bathtub. Such models are discussed in the following sections. (The bathtub model applies to both nonrepairable and repairable systems.)

14.5.3 Weibull Distribution

The two-parameter Weibull distribution, given in Section 3.4.6 (there is also a three-parameter Weibull distribution), is generally a better model for reliability applications than the exponential distribution and indeed is one of the most frequently used distributions in reliability work. This is partly due to the flexibility in fitting data that is afforded by a distribution that has two or three parameters. The distribution is named after Walodie Weibull, a Swedish engineer who proposed the distribution in 1939 as an appropriate distribution for the breaking strength of materials, although the distribution is used just about as often in analyzing lifetime data.

Extreme value theory can be used to justify the use of the Weibull distribution to model failure when there are many similar "competing" failure processes and the first failure causes the mechanism to fail. With $f(t) = (\beta/\alpha)(t/\alpha)^{\beta-1}\exp[-(t/\alpha)^{\beta}]$, it can easily be shown, as the reader is asked to do in Exercise 14.1, that $R(t) = \exp[-(t/\alpha)^{\beta}]$. It follows that $h(t) = (\beta/\alpha)(t/\alpha)^{\beta-1}$. Notice that $h(t)$ depends on t unless $\beta = 1$. (Recall from Section 3.4.6 that the Weibull distribution reduces to the exponential distribution when $\beta = 1$.) The shape of the curve when $h(t)$ is graphed against t will depend on the combination of values for α and β.

Therefore, for simplicity assume that $\alpha = 1$. We see that $h(t)$ is a constant (namely, $1/\alpha$) when $\beta = 1$. When $\beta > 1$, $h(t)$ is an increasing function of t and is a decreasing function of t when $\beta < 1$. Since β, if not equal to one, must obviously be either greater than one or less than one, we cannot capture the entire "bathtub." This is clearly true for any value of α. Therefore, what is frequently done is to use $\beta < 1$ and concentrate on modeling the portion of the bathtub that corresponds to the major portion of a product's life. As such, the Weibull distribution has been successfully used in modeling failure rates for material strengths, capacitors, and ball bearings, just to name a few.

The natural logarithm (base e) of a Weibull random variable has an extreme value distribution (Section 14.5.5), and the log of data well-fit by a Weibull is often used as that simplifies the analysis somewhat.

14.5.4 Lognormal Distribution

Another model worthy of consideration in reliability applications is the lognormal distribution. Recall from Section 3.4.8 that this distribution arises if the natural logarithm of a random variable has a normal distribution. Unlike the exponential and Weibull distributions, the lognormal distribution does not have a failure rate that can be written in a simple form. The lognormal has been used successfully in many applications to model the failure rates of semiconductor failure rate mechanisms such as corrosion, diffusion, metal migration, electromigration, and crack growth. It is also commonly used as a distribution for repair time, as well as being commonly used in the analysis of fatigue failures caused by fatigue cracks.

14.5.5 Extreme Value Distribution

Assume that a system consists of n identical components that are aligned in a series, and the system fails when the first of these components fails. System failure times then are the minimum of n random component failure times. Since the n components are identical, they would have identical distributions.

The *pdf* of the smallest extreme (i.e., the minimum) is given by

$$f(x) = \frac{1}{b} \exp\left(\frac{x-a}{b}\right) \exp\left[-\exp\left(\frac{x-a}{b}\right)\right] \qquad -\infty < x < \infty$$

with a the location parameter and b the scale parameter.

Since the extreme value distribution is the limiting distribution of the minimum of a large number of unbounded identically distributed random variables, this distribution would be a logical choice for modeling system failures. If n is large, the Weibull distribution might also be considered since the Weibull is the limiting distribution of the extreme value distribution. (An extreme value variant of the Weibull distribution is referred to as a Type III extreme value distribution, with the extreme value distribution itself labeled Type I.)

14.5.6 Other Reliability Models

The *gamma distribution* (Section 3.4.5) is infrequently used as a reliability model, although it is appropriate when standby backup units are in place that have exponential lifetimes. Although the proportional hazards model of Cox (1972) is frequently used in survival

analysis, it has very limited applicability in engineering and thus won't be covered in detail here.

The Birnbaum–Saunders distribution (not covered in Chapter 3) results from a physical fatigue process for which crack growth (e.g., a crack in cement) causes failure. For more information on the model, see Birnbaum and Saunders (1969) and Section 8.1.6.6 of Croarkin and Tobias (2002).

14.5.7 Selecting a Reliability Model

As with modeling in general, there are various approaches that can be used to select a reliability model. As implied previously, there may be physical aspects that could suggest the choice of a particular model. If not, probability plots (illustrated in Section 6.1.2) and other plots can be performed when the data are failure data from nonrepairable populations, although probability plots for distributions other than the normal distribution are not available in some statistical software, and this is especially true for distributions such as the extreme value distribution. These plots are available in MINITAB, however, and Figure 14.2 is a panel of probability plots for some of the distributions that are available under MINITAB's reliability/survival menu.

Figure 14.2 shows that a Weibull distribution gives by far the best fit of the four fitted distributions, and in fact the data were simulated from a two-parameter Weibull distribution. (For the other distributions for which MINITAB automatically produces a probability plot, a three-parameter Weibull distribution gave a slightly better fit, as did a three-parameter lognormal distribution. This is not surprising as a model other than the model used in generating the data will often give a slightly better fit. Here, however, model parsimony

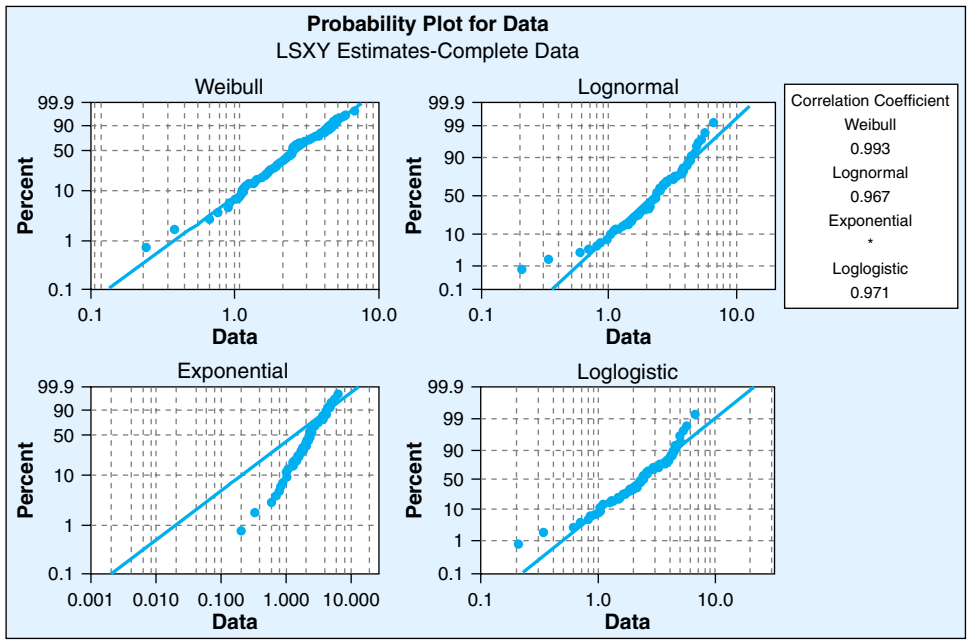

Figure 14.2 Panel of probability plots of simulated data.

would mandate that the simpler two-parameter model be fit since there is hardly any difference in the fit of the three models.)

Various statistical tests have also been given for determining whether or not a particular distribution provides a good fit to reliability data, and in particular, certain tests have been developed for the exponential distribution because of the central role that it has played in reliability work.

For repairable populations, trend plots and Duane plots can be used (see Croarkin and Tobias, 2002, Section 8.2.2.3), provided that appropriate software or computer paper is available, although use of the latter is an antiquated approach.

Selection of a life distribution model can be made using theoretical considerations or for other reasons such as a model simply providing a good fit. In general, the more parameters a distribution has, the greater modeling flexibility the distribution provides. Accordingly, we would expect the two-parameter Weibull distribution to be particularly useful, and the three-parameter Weibull distribution to provide even greater flexibility.

Furthermore, use of the Weibull distribution can be supported by extreme value theory, as was stated previously. That is, if multiple (competing) failure mechanisms exist and failure of the unit occurs at the occurrence of the first failure mechanism, the Weibull distribution should be a prime candidate for the life distribution model.

The lognormal distribution will, in general, be a good candidate when failures result from chemical reactions. As stated previously, a gamma distribution will rarely give a good fit to failure time data as its failure rate generally does not represent the failure rates of products.

14.6 RELIABILITY ENGINEERING

Reliability engineering is an activity rather than a branch of engineering like mechanical engineering or civil engineering. As such, reliability engineering can be performed by design engineers, quality engineers, or systems engineers, in addition to reliability engineers. The objective is to determine the expected or actual reliability of a product, process, or service and of course to try to improve reliability. The American Society for Quality (ASQ) offers certification in reliability engineering in the form of its certified reliability engineer (CRE) designation. Various other certification programs are offered, including the one offered by the University of Tennessee. Reliability engineering is viewed as consisting of reliability, maintainability, and availability. (Maintainability refers to the field/activity of estimating the distribution of time to repair, and availability is a commonly used measure of performance of repairable systems.)

14.6.1 Reliability Prediction

Reliability prediction is part of reliability engineering, and by this we mean predicting future failures. Kvam and Miller (2002) stated "Engineering practitioners have been apprehensive about applying predictive inference to problems that involve the estimation of future failures." Reasons for this could certainly be debated, but it is inarguably true that prediction intervals outside the area of regression analysis have been almost completely ignored in courses and texts on engineering statistics. One impediment to the widespread use of prediction and prediction intervals in reliability engineering is that such intervals

(e.g., for multiple failure models) lack the simplicity of the prediction interval form given in Eq. (7.1) and can be computer-intensive.

14.7 EXAMPLE

Here is a typical failure mode effects analysis (FMEA).

Failure Mode of Circuit Breakers, Percentage of Total Failures in Each Failure Mode

Failure Characteristics	Percentage of Total Failures, All Voltages
Back-up protective equipment required, failed while opening	9
Other circuit breaker failures:	
Damaged while successfully opening	7
Failed while in service	32
Failed to close when it should	5
Damaged while closing	2
Opened when it shouldn't	42
Failed during testing or maintenance	1
Damage discovered during testing or maintenance	1
Other	1
Total Percentage	100

[*Source*: J. C. Das, "Industrial Power Systems," in *Wiley Encyclopedia of Electrical and Electronics Engineering Online* (J. Webster, ed.). Used with permission of John Wiley & Sons, Inc.]

14.8 IMPROVING RELIABILITY WITH DESIGNED EXPERIMENTS

Just as the goal should be to improve quality, as emphasized throughout Chapter 11, improving reliability should also be a goal. Measuring reliability and predicting reliability are obviously important, but that isn't enough if product reliability is judged unacceptable. Wu and Hamada (2000, Chapter 12) discuss the use of experimental design for improving reliability. This does not require the use of esoteric experimental designs, as designs presented in Chapter 12 of this book can be used, but the manner in which the values of the dependent variable are determined does dictate how the analysis is to be performed.

Specifically, if the data are not censored, standard analysis methods can be used, unless one is concerned over the fact that the error terms in the linear model with failure time as the dependent variable cannot have a normal distribution since the failure times cannot be normally distributed because they are nonnegative. To adjust for this, Wu and Hamada (2000) advocated using a transformation, such as a log transformation, so that the transformed times can be negative. Although theoretically proper, regression models are usually applied without transforming the dependent variable when the latter can only

assume positive values. Since normality doesn't exist in practice anyway, one might simply check to see if the residuals appear to be approximately normally distributed. The residuals will usually have closer to a normal distribution than will the errors when the errors are not normally distributed, but approximate normality for the residuals may be good enough. Otherwise, a log transformation would be advisable.

■ EXAMPLE 14.2 (CASE STUDY)

Problem Statement

Textbook problems are often much simpler than problems encountered in real life, with the consequence that students may fail to understand that the use of statistics to solve problems will not necessarily be easy. A good example of this is a problem that was faced by the Controls Division of the Eaton Corporation in Athens, Alabama. During 1988, the failure rate of an industrial thermostat assembly inexplicably rose to as high as 7% in certain batches. The employees decided to conduct an experiment, and subsequently conducted a second experiment that was guided by what they learned from the first experiment. Bullington, Lovin, Miller, and Woodall (1993) described only the results of the second experiment.

Experimental Design

Eleven factors were chosen for study and the experimenters decided to use the Taguchi L_{12} array, which is better known as a 12-run Plackett–Burman design (see Section 12.9.2). There were 10 observations for each of the 12 treatment combinations, and the 120 thermostats were cycled until failure, with testing stopped after 7,342,000 cycles. Thus, the data were right censored. The data are in `CH14ENGEATON.MTW` at the text website.

It should be noted that this is *not* a Plackett–Burman design with 10 replications, as each set of 10 thermostats manufactured at each treatment combination was made in the same batch of production.

This is an important point, as we would expect the experimental error to be less with this manufacturing arrangement than if, say, 120 different batches were used. The standard analysis of the data would be based on the assumption that the experiment was replicated, and of course the estimate of the error variance helps determine what effects are significant when a replicated design is used. So an underestimate of the error variance could cause certain effects to be falsely declared significant. (This relates to the discussion in Chapter 12 of multiple readings versus replication and the importance of determining the correct error term when trying to ascertain the significance of effects.)

Data Analysis

Indeed, the analysis given by Bullington et al. (1993) showed all 11 of the main effects (the only effects that are estimable) to be significant, with the largest p-value being .0025. Even though the selection of these 11 factors was based on the outcome of a previous experiment, it is unlikely that all of the effects are really significant. In an attempt to circumvent the analysis problem caused by each set of 10 observations being from the same batch, Bullington et al. (1993) used the average of the 10 numbers as a single response value. The data must then be analyzed as having come from an unreplicated design.

Analysis Problems

The Eaton experimenters decided to use only the two shortest failure times for each treatment combination in their analysis. Doing so discards 80% of the data and biases the results, however. In particular, for several treatment combinations the two shortest failure times differed considerably from some of the other eight failure times. Thus, this analysis is clearly objectionable.

The fact that all 11 main effects are significant when the 120 observations are used suggests that the error variance has probably been underestimated, but we can't draw that conclusion without first considering the (complex) alias structure of a Plackett–Burman design. Each main effect is partially aliased with every two-factor interaction that does not involve the main effect letter. For example, the main effect of A is partially aliased with every two-factor interaction that does not contain the letter A. That's $\binom{11}{2} = 55$ interactions! (The term "partially aliased" means that the coefficient of the effect in the alias structure is less than one.)

The main effect estimates are given in Table 14.1.

TABLE 14.1 Main Effect Estimates for Thermostat Experiment Data

Effect	Estimate
A	−0.62
B	0.44
C	−0.64
D	0.57
E	−2.05
F	0.46
G	−0.78
H	−1.11
I	−0.66
J	−0.55
K	−0.71

Clearly, E and H stand out, with all of the other effect estimates being below 1.00 in absolute value. [Factor E was the grain size of the beryllium that was used in the diaphragm and factor H was the length of time of the heat treatment. The various factors are described in detail in Bullington et al. (1993).] Wu and Hamada (2000) analyzed these data and noted that the coefficients (i.e., half the effect estimates) for the other nine effects are in a narrow range (.22 to .39 in absolute value). This suggested to them that these nine effects being significant could be due to their partial aliasing with the EH interaction. To check this, they removed factor B from the analysis (one factor had to be removed since the model was saturated). Their subsequent analysis showed, interestingly, that E, H, and EH were the only significant effects. Thus, their suspicion does seem to have been correct, and a meaningful analysis resulted despite the shaky nature of the multiple values that were not true replicates.

In MINITAB, reliability data are analyzed using the `LREGRESSION` command, which stands for "life regression." The sequence of MINITAB commands that produced Table 14.1 and other statistics are given at the textbook website. A distribution would have to be assumed for the error distribution, and of course this assumption would have to be checked.

One way to do so would be to construct a probability plot of the standardized residuals. We are immediately faced with the question of what should be the distribution of the standardized residuals. In this case study, the failure times were believed to follow a lognormal distribution. A distribution must be specified because the parameters are estimated using maximum likelihood. The standardized residuals could not have a lognormal distribution, even approximately, however, as this would mean that the failure times were never underestimated since the values of a lognormal random variable can never be negative. Clearly, this would be nonsensical.

We can avoid this dilemma by taking the logarithm of the failure times and using those transformed values in the analysis. Then the error distribution is normal, under the initial assumption of a lognormal distribution, and a normal probability plot of the standardized residuals could be constructed. The reader is asked to do this in Exercise 14.2 and interpret the plot.

Whether we construct a normal probability plot or simply look at the list of the standardized residuals, there are certain observations that stand out. In particular, the first observation is much smaller than the other observations for that treatment combination, and the same can be said for observation #51. Similarly, observations #101and #102 stand out, as does #120. Since there is considerable variability among the failure times for a given treatment combination, these may be valid numbers, but they should definitely be checked out. ■

14.8.1 Designed Experiments with Degradation Data

With improvement in quality, as has been the objective particularly during the past fifteen years or so, failure data is harder to come by because there may be very few failures during normal testing intervals. This can even be a problem with accelerated testing since a good accelerating factor may be hard to find.

Consequently, it may be necessary to use product degradation data in designed experiments rather than failure data. Wu and Hamada (2000, Section 12.8) gave a procedure that might be used with a designed experiment. As stated in Section 14.3.3 of this book, a detailed discussion of the use of degradation data, including its advantages and disadvantages, can be found in Section 8.4.2.3 of Croarkin and Tobias (2002). See also Nelson (1990, pp. 521–544) and Tobias and Trindade (1995, pp. 197–203).

14.9 CONFIDENCE INTERVALS

Various confidence intervals can be constructed in a reliability analysis, including confidence intervals on parameter values, but we are generally most interested in confidence intervals on percentiles. In confidence intervals such as those given in Chapter 5, the width of the confidence interval varies inversely with $\sqrt{n}$. This does not hold true for confidence intervals constructed in reliability analysis, however. Rather, the width of the interval varies inversely with the square root of the number of *failures*. Thus, having a large sample size isn't necessarily helpful, as we also need at least a moderate number of failures.

For the Bullington et al. (1993) data analyzed in Example 14.2, we would certainly be interested in constructing confidence intervals on selected percentiles. For example, we might be interested in the interval for the 50th percentile of the number of cycles before failure, and also, say, the 80th percentile, as well as the 95th and 99th percentiles, for a

given combination of (coded) values of factors E and H, with the EH interaction also in the model. The results when factor E is at the high level (i.e., $+1$) and factor H is at the low level (i.e., -1) are as follows.

Table of Percentiles

Percent	Predictor Row Number	Percentile	Standard Error	95.0% Normal CI Lower	Upper
50	1	145.6535	20.6414	110.3296	192.2870
80	1	279.9207	41.9363	208.6951	375.4549
90	1	393.8559	62.9653	287.9100	538.7881
95	1	522.1653	89.0778	373.7653	729.4862

From casual inspection of the data we can see that the best settings would be both factors at the low level, so when those settings are used we obtain the following results:

Table of Percentiles

Percent	Predictor Row Number	Percentile	Standard Error	95.0% Normal CI Lower	Upper
50	1	6860.876	1166.750	4916.156	9574.885
80	1	13185.41	2490.279	9106.057	19092.24
90	1	18552.22	3760.132	12470.27	27600.44
95	1	24596.12	5304.854	16116.87	37536.39

Both the percentile estimates and the standard errors are very large numbers, with the latter resulting partly from the fact that most of the data for this treatment combination were censored, so this creates considerable uncertainty. Since the censoring value was 7342 and the estimate of the 90th percentile value is more than twice this value, there is obviously considerable extrapolation beyond the censoring value.

It is best that tables of this type be produced in menu mode in MINITAB because of the number of subcommands that would have to be entered if command code were used.

14.10 SAMPLE SIZE DETERMINATION

A simple formula for determining sample size, based on minimal assumptions, was given by Meeker, Hahn, and Doganaksoy (2004). Analogous to the Section 14.5 notation, let $R(t) = R$ denote the desired reliability for a specified value of t, and let α have the usual designation, as in a $100(1 - \alpha)\%$ confidence interval. Then

$$n = \frac{\log(\alpha)}{\log(R)} \tag{14.6}$$

■ EXAMPLE 14.3

To illustrate the application of the formula, assume that the reliability of a certain type of light-emiting diode (LED) is to be .99 for a specified number of hours of operation. If

the degree of confidence is 95%, $n = \log(\alpha)/\log(R) = \log(.05)/\log(.99) = 298$, so this number of units would need to be used in the test.

If a Weibull distribution is assumed, Meeker and Escobar (1998) gave the sample size expression as

$$n = \frac{1}{k^{\beta}}\left(\frac{\log(\alpha)}{\log(R)}\right) \tag{14.7}$$

with β the Weibull shape parameter and k a constant to be selected. It is worth noting that Eq. (14.7) reduces to Eq. (14.6) if $\beta = 1$ with k chosen to be 1. ■

It is also possible to use software such as MINITAB for such purposes as determining the sample size for a confidence interval on the reliability after a certain length of time and a specified value for the distance between the estimate of the reliability and a lower or upper confidence bound. Below we use the same time and parameters as in the preceding example and specify a distance of 0.05. This results in a sample size of $n = 44$, as can be seen.

```
Estimation Test Plans

Uncensored data

Estimated parameter: Reliability at time = 48
Calculated planning estimate = 0.965571
Target Confidence Level = 95%

Planning distribution: Weibull
Scale = 448.3, Shape = 1.5

                          Actual
             Sample    Confidence
Precision      Size         Level
     0.05        44       95.1233
```

14.11 RELIABILITY GROWTH AND DEMONSTRATION TESTING

Reliability growth testing is performed for the purposes of (1) assessing current reliability, (2) identifying and eliminating faults, and (3) forecasting future reliability. This is somewhat similar to the stages in the use of control charts, as discussed in Chapter 11. Reliability demonstration testing is used to see if a desired reliability level has been achieved after faults have been identified and removed.

For example, assume that a Weibull distribution with a scale parameter of 448.3 and a shape parameter of 1.5 (α and β, respectively, in the *pdf* for the Weibull distribution in Section 14.5.3 and also in Section 3.4.6) is used as the reliability model and no failures will be allowed. How many items must be tested for a period of 48 hours to demonstrate a reliability of 90% with a confidence level of 95%? This can be determined using appropriate software and only using software because the hand computations would be quite

complicated. The MINITAB output for this is given below and we see that $N = 86$ items must be tested.

```
Demonstration Test Plans

Substantiation Test Plan
Distribution: Weibull, Shape = 1.5
Scale Goal = 448.3, Target Confidence Level = 95%


                           Actual
Failure Testing Sample Confidence
   Test    Time   Size      Level
      0      48     86    95.0859
```

Reliability demonstration testing is related to demonstrating a mean time to failure (*MTTF*).

14.12 EARLY DETERMINATION OF PRODUCT RELIABILITY

As discussed by Mann, Schafer, and Singpurwalla (1974) and Meeker and Escobar (1998), traditional reliability methods fail when data are sparse. When test data are sparse, PREDICT (Performance and Reliability Evaluation with Diverse Information Combination and Tracking) can be used to estimate reliability. This is a formal, multidisciplinary process that combines all available sources of information. It is described in detail by Booker, Bement, Meyer, and Kerscher (2003).

14.13 SOFTWARE

14.13.1 MINITAB

MINITAB has excellent reliability capabilities, some of which were illustrated in the chapter. Capabilities not illustrated include accelerated life testing, with the Arrhenius equation being one of four equations from which the user can select. There are many other methods available, including parametric and nonparametric distribution analyses and growth curve analyses.

14.13.2 JMP

JMP has very limited capability for reliability analyses, and what is there is really survival analysis methods playing a dual role. Although it is true that the survival of a patient is roughly tantamount to a product still functioning after an extended period of time, obviously patients aren't "accelerated" so there are certain tools in reliability analysis that are not used in survival analysis.

14.13.3 Other Software

Other general purpose statistical software also have reliability analysis capability, in addition to software that is specifically for reliability analysis. One list of such software is given at `http://www.asq.org/reliability/resources`.

14.14 SUMMARY

Reliability engineering and reliability analysis are important parts of engineering. Reliability is distinct from quality (improvement) in that with reliability we are essentially trying to model and depict time to failure, whereas with quality improvement the objective is to prevent bad product from occurring very often and to reduce the variability of the values of process characteristics over the units of production. There is no attempt to try to determine how long an item will last.

Although the definition of quality implies quality over time, quality control/improvement methods do not require monitoring when the product is in customers' hands. Such monitoring would be of no value anyway in terms of quality improvement since the quality is set once it leaves the production process.

In order to model failures, there of course must generally be failures. If product reliability is so good that reliability testing doesn't produce practically any failures, we then have a problem similar to the problem caused by an extremely small percentage of nonconforming units when a *p*-chart (Section 11.9.1) is used.

The usual approach when regular testing is insufficient is to use accelerated testing. The success of accelerated testing depends very strongly on whether or not accelerating factors can be found that lead to failure. Such factors don't always exist.

A model that ideally fits the failure times well is necessary for both normal stress testing and accelerated testing. The Weibull distribution and lognormal distribution often work well as a model for normal testing, with the model for accelerated testing determined by the choice of the acceleration factor(s).

A good introductory-level presentation of the analysis of reliability data is Nelson (1983). Another useful reference not previously mentioned is Meeker and Escobar (1998), and Blischke and Murthy (2003) is a book that consists entirely of case studies, 26 of them. A good, short (199-page) introduction to engineering reliability, although at a higher level than this chapter, is Barlow (1998).

REFERENCES

Barlow, R. E. (1998). *Engineering Reliability*. Philadelphia: Society for Industrial and Applied Mathematics and Alexandria, VA: American Statistical Association.

Birnbaum, Z. W. and S. C. Saunders (1969). A new family of life distributions. *Journal of Applied Probability*, **6**, 319–327.

Blischke, W. R. and D. N. Prabhakar Murthy (2003). *Case Studies in Reliability and Maintenance*. Hoboken, NJ: Wiley.

Booker, J. M., T. R. Bement, M. A. Meyer, and W. J. Kerscher (2003). PREDICT: A new approach to product development and lifetime assessment using information integration technology. Chapter 11 in *Handbook of Statistics 22: Statistics in Industry* (R. Khattree and C. R. Rao, eds.). Amsterdam: Elsevier Science B.V.

Bullington, R. G., S. Lovin, D. M. Miller, and W. H. Woodall (1993). Improvement of an industrial thermostat using designed experiments. *Journal of Quality Technology*, **25**(4), 262–270.

Cox, D. R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society, Series B*, **34**, 187–220.

Croarkin, C. and P. Tobias, technical editors (2002). *NIST/SEMATECH e-Handbook of Statistical Methods* (`http://www.itl.nist.gov/div898/handbook/mpc/section3/mpc33.htm`).

Dhillon, B. S. (2005). *Reliability, Quality and Safety for Engineers*. Boca Raton, FL: CRC Press.

Duane, J. T. (1964). Learning curve approach to reliability monitoring. *IEEE Transactions on Aerospace*, **2**, 563–566.

Fluharty, D. A., Y. Wang, and J. D. Lynch (1998). A simplified simulation of the impact of environmental interference on measurement systems in an electrical components testing laboratory. Chapter 15 in *Statistical Case Studies: A Collaboration Between Academe and Industry* (R. Peck, L. D. Haugh, and A. Goodman, eds.). Philadelphia: Society for Industrial and Applied Mathematics and Alexandria, VA: American Statistical Association.

Gupta, R. C. and D. M. Bradley (2003). Representing the mean residual life in terms of the failure rate. *Mathematical and Computer Modeling*, **37**, 1271–1280.

Hooper, J. H. and S. J. Amster (2000). Analysis and presentation of reliability data. In *Handbook of Statistical Methods for Engineers and Scientists*, 2nd ed. (H. M. Wadsworth, ed.). New York: McGraw-Hill.

Kvam, P. H. and J. G. Miller (2002). Discrete predictive analysis in probabilistic safety assessment. *Technometrics*, **34**(1), 106–117.

Mann, N. R., R. E. Schafer, and N. D. Singpurwalla (1974). *Methods for Statistical Analysis of Reliability and Life Data*. New York: Wiley.

Meeker, W. Q. and L. A. Escobar (1998). *Statistical Methods for Reliability Data*. New York: Wiley.

Meeker, W. Q., G. J. Hahn, and N. Doganaksoy (2004). Planning life tests for reliability demonstration. *Quality Progress*, **37**(8), 80–82.

Nelson, W. (1982). *Applied Life Data Analysis*. New York: Wiley.

Nelson, W. (1983). *How to Analyze Reliability Data*, ASQC Basic References in Quality Control: Statistical Techniques, Vol. **6**. Milwaukee, WI: American Society for Quality Control.

Nelson, W. (1990). *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*. New York: Wiley.

Rausand, M. and A. Høyland (2004). *System Reliability Theory: Models, Statistical Methods, and Applications*. Hoboken, NJ: Wiley.

Rigdon, S. E. and A. P. Basu (2000). *Statistical Methods for the Reliability of Repairable Systems*. New York: Wiley.

Spence, H. (2000). Application of the Arrhenius equation to semiconductor reliability predictions. Manuscript.

Tobias, P. A. and D. C. Trindade (1995). *Applied Reliability*, 2nd ed. New York: Chapman and Hall.

Von Alven, W. H. ed. (1964). *Reliability Engineering*. Englewood Cliffs, NJ: Prentice Hall.

Wu, C. F. J. and M. Hamada (2000). *Experiments: Planning, Analysis, and Parameter Design Optimization*. New York: Wiley.

EXERCISES

14.1. Derive the reliability function for the Weibull distribution that was given in Section 14.5.3.

14.2. Referring to Example 14.2, construct a normal probability plot of the standardized residuals, after first taking the (base e) log of the failure values. (If you use MINITAB, you can use the `SPplot` subcommand with the `LREG` main command in MINITAB to produce the plot.) Interpret the plot. In particular, are there any unusual values? If the standardized residuals appear to be approximately normally distributed, then the untransformed failure times must be approximately lognormally distributed. Does the latter appear to be the case? Explain.

14.3. Verify Eq. (14.5).

14.4. A complex electronic system is built with a certain number of backup components in its subsystems. One subsystem has three identical components, each of which has a probability of .25 of failing within 1100 hours. The system will operate as long as at least one of the components is operating (i.e., a parallel system). What is the probability that the subsystem is operational after 1100 hours?

14.5. Suppose that the time to failure, in minutes, of a particular electronic component subjected to continuous vibrations can be approximated by a Weibull distribution with $\alpha = 50$ and $\beta = 0.40$. What is the probability that such a component will fail in less than 5 hours?

14.6. A freight train requires two locomotives for a one-day run. It is important that the train not arrive late, so the reliability of the locomotives should be determined and will hopefully be adequate. Past experience suggests that times to failure for such locomotives can be approximated by an exponential distribution with a mean of 42.4 days. Assume that the locomotives fail independently and determine the reliability, $R(t)$, for the train. Should the company be concerned about the reliability number? Explain.

14.7. It is known that the life length of a certain electronic device can be modeled by an exponential distribution with $\lambda = .00083$. Using this distribution, the reliability of the device for a 100-hour period of operation is .92. How many hours of operation would correspond to a reliability of .95?

14.8. Although a normal distribution is not one of the most frequently used distributions for modeling failure, it has been used to model certain types of failure behavior. Assume that the time to failure of an item is approximately normally distributed with a mean of 90 hours and a standard deviation of 10 hours.

(a) What is the numerical value of $R(t)$ for $t = 105$?

(b) What value of t corresponds to a reliability of .95?

14.9. Assume that an electronic assembly consists of three identical tubes in a series system and the tubes are believed to function (and fail) independently. It is believed that the function $f(t) = 50t \exp(-25t^2)$ adequately represents the time to failure for each tube. Determine $R_s(t)$. Looking at your answer, does $f(t)$ appear to be a realistic function? Explain.

14.10. The design of an electrical system is being contemplated, with an eye toward using enough components in a parallel system to provide acceptable reliability. Assume that the time to failure for each component is approximately normally distributed with a mean of 4500 hours and a standard deviation of 400 hours, and that the components are independent. If the probability of failure within the first 5,000 hours is to be less than .005, how many components should be used?

14.11. What is the failure rate for the time to failure distribution given in Exercise 14.10, if it can be obtained? If it cannot be obtained, explain why that is the case.

14.12. Determine the failure rate when the time to failure for a particular type of component is believed to be reasonably well-represented by a Weibull distribution with $\alpha = 40$ and $\beta = 0.35$. Graph the failure rate. Does this resemble a bathtub curve, as in Figure 14.1? Does the curve look reasonable relative to what we might expect for the hazard rate for a company whose products have a good reputation for good reliability? Explain.

14.13. Obtain the expression for the failure rate for a normal distribution if it is possible to do so. If it is not possible, explain why not.

14.14. Obtain the general expression for $R(t)$ when time to failure is assumed to have the extreme value distribution given in Section 14.5.5.

14.15. Show that the failure rate for the extreme value distribution given in Section 14.5.5 is not a constant.

14.16. A parallel system consists of three nonidentical components with failure probabilities during the first 100 hours of .052, .034, and .022, respectively. What is the probability that the system is still functioning at 100 hours?

14.17. Consider the following statement: "For a series system comprised of independent components, the reliability of the system is less than or equal to the lowest component reliability." Does that seem counterintuitive? Show why the statement is true.

14.18. Show that the reliability for a parallel system is greater than or equal to the highest component reliability.

14.19. If possible, give three examples of engineering applications for which a constant failure rate seems appropriate, and thus the exponential model would seem appropriate as a model for time to failure. If you have difficulty doing so, state how the failure rates would deviate from being constant in engineering applications with which you are familiar.

14.20. Critique the following statement: "Making normal stress estimates using accelerated testing is as risky as extrapolation in regression since in both cases the user is extrapolating from the range of the data."

14.21. A system need not be strictly a parallel or a series system, as many other combinations are possible and are in use. Assume that the Weibull model is being used and the system is such that three of the components are arranged as in a parallel system but the fourth component stands alone and must function for the system to function. Give the general expression for $R_s(t)$, assuming independence for the components.

14.22. The time to failure of a choke valve can be modeled by a Weibull distribution with $\beta = 2.25$ and $\alpha = 0.87 \times 10^4$ hours.

(a) Determine $R(4380)$, the reliability at 4380 hours.

(b) Determine the *MTTF*.

14.23. Determine the *MTTF* for the lognormal distribution in terms of the parameters of the distribution.

14.24. A machine with constant failure rate λ has a probability of .50 of surviving for 125 hours without failure. Determine the failure rate λ if $R(t) = \exp(-\lambda t)$.

14.25. Consider an exponential distribution with $\lambda = 2.4 \times 10^{-5}$ (hours) and determine the *MTTF*.

14.26. Consider $R(t) = 1 - (x/\beta)^{\alpha}$, $0 \leq x \leq \beta$, $\alpha, \beta > 0$. Determine $f(t)$ and the failure rate.

14.27. Determine the failure rate for a continuous uniform distribution defined on the interval (a, b). Does this seem to be a reasonable failure rate? Are there any restrictions on t, relative to a and b? Explain.

14.28. Using the results of Exercise 14.26, determine the *MTTF*.

14.29. Consider Example 14.1 on LEDs. Do you think a normal distribution would generally be a viable model for such applications? Why or why not?

14.30. A computer has two independent and identical central processing units (CPUs) operating simultaneously. At least one CPU must be operating properly for the computer to function successfully. If each CPU reliability is .94 at a certain point in time, what is the computer reliability at that point in time? Does this reliability seem acceptable? Explain.

14.31. A safety alarm system is to be built of four identical, independent components, with the reliability of each component being p. Two systems are under consideration: (1) a 3-out-of-4 system, meaning that at least 3 of the components must be functioning properly, and (2) a 2-out-of-4 system. Determine the system reliability for each as a function of p. If the cost of the second system is considerably less than the cost of the first system, would you recommend adoption of the second system if $p \geq .9$? Explain.

14.32. Assume that a system has four active, independent, and nearly identical units that form a series system. If the reliability of two of the units after 1,000 hours is .94 and .95 for the other two units, what is the system reliability at 1,000 hours?

14.33. The first ten failure times for a repairable system are (in hours), 12,150, 13,660, 14,100, 11,500, 11,750, 14,300, 13,425, 13,875, 12,950, and 13,300. What is the estimate of the *MTBF* for that system?

14.34. Could a *MTBF* be computed for a nonrepairable system? Why or why not?