5-1

One of the most important factors which influences the design process of a product or a system is the reliability functions of its components. For example, a piece of equipment, such as a personal computer, often uses a large number of integrated circuits in a single system. Clearly, a small percentage of the integrated circuits that exhibit early life failures can have a significant impact on the reliability of the system. It is commonly observed that the failure rate of semiconductor devices (components of the integrated circuits) is initially high, and then decreases to a steady state value with time. There are many other examples that illustrate that the reliability of the system is highly dependent on the reliability of the individual components that comprise the system. Submarine optical fiber transmission systems, weather satellites, telecommunication networks, air traffic control, and supercomputers are typical systems that require “highly reliable” components in order to achieve the reliability goals of the total system. Designers of such systems usually set the long-term reliability of the systems to unusually high values, for example, 0.999999 or higher (note that reliability is a monotonically decreasing function with time, and availability is indeed the long-term measure of reliability for repairable systems). Such values may appear unrealistic, but they are achievable when appropriate subsystems are designed and highly reliable components are incorporated in the systems. In order to estimate the reliability of the individual components or the entire system, we may follow one or more of the following approaches.

5-2

The failure data for components can be found in data banks such as GIDEP (GovernmentIndustry Data Exchange Program), MIL-HDBK-217D (which includes failure data for components as well as procedures for reliability prediction), AT&T Reliability Manual (Klinger et al., 1990), and Bell Communications Research Reliability Manual (Bell Communications Research, 1986, 1995). In such data banks and manuals, the failure data are collected from different manufacturers and are presented with a set of multiplying factors that relate the failure rates to different manufacturer’s quality levels and environmental conditions. For example, the general expression used to determine steady-state failure rate λSS of an electronic or electrical component (most of, if not all, electronic components exhibit constant failure rates) that includes different devices is

5-3

An operational life test (OLT) is one in which prototypes of a product—whether it is a single component product such as a telephone pole, or multicomponent products such as cars and computers—are subjected to stresses and environmental conditions at typical normal operating conditions. The duration of the test is determined by the number of products under test (sample size) and the expected number of failures. In all cases, the test should be terminated when its duration reaches the expected life of the product. Clearly, this test requires extensive durations especially when the product’s life is rather long, the case of many electronic devices. An example of the operational life testing is the testing of utility poles by taking a sample and placing it under the same environmental and weather conditions and observing the failure times over an extended period that ranges from 1 year to several years. Similar operational life testing is performed on electric switching systems and mechanical testing machines. Usually, the OLT equipment is designed to be capable both of operating components and of testing them on a scanning basis. As mentioned earlier, the test conditions are not accelerated but rather designed to simulate the field operating conditions (such as temperature fluctuations and power on/off). Analyses of test results are used to monitor and estimate the reliabilities and failure rates of products in order to achieve the desired specifications. Although the results obtained from OLT are the most useful among other tests, the duration of the test is relatively long and the costs associated with the tests may make them prohibitive to run. Indeed, this test is not classified as an accelerated life testing (ALT) since no real acceleration of time or stress is performed

5-4

It is often found that in a large population of components (or products), some individual components have quality defects which considerably affect the component's life. In order to "weed out" these individual components, a burn-in test is performed at stress conditions—that is, the time or applied stresses are accelerated. It is important to note that the test conditions must be determined such that the majority of failures are detected without significantly overstressing the remaining components. Additionally, an optimal burn-in period should be estimated such that the total cost to the producer and the user of the product is minimized. There are two cost elements that should be considered in estimating the optimal burn-in period. They are: (1) cost per unit time of the test (long test periods are costly to the producer), and (2) cost of premature failures since short test periods may not completely "weed out" the defective components which in turn results in significant costs for both producers and consumers. Mathematical models for estimating the optimal burn-in period are given in the literature of Jensen and Petersen (1982), Bergman (1985), Kuo et al. (2001), and Wu and Su (2002).

5-5

Accelerated life testing (ALT) is used to obtain information quickly on life distributions, failure rates, and reliabilities. ALT is achieved by subjecting units and components to test conditions such that failures occur sooner. Thus, prediction of the long-term reliability can be made within a short period of time. Results from the ALT are used to extrapolate the unit characteristic at any future time t and at given normal operating conditions. There are two methods used for conducting an accelerated life test. In the first method, it is possible to accelerate the test by using the product more intensively than in normal use. For example, in evaluating the life distribution of a light bulb of a telephone set which is used on the average 1 h a day, a usage of the bulb during its expected life of 40 years can be compressed into 18 months by cycling the power on/off continuously during the test period. Another example, the endurance limit of a crankshaft of a car with an expected life of 15 years (3 h of driving per day), can be obtained by compressing the test into 2 years. However, such time compression (accelerating time)

may not be possible for a product that is in constant use, such as a mainframe computer. Moreover, in such cases, the prediction of reliability must consider the aging effect on the component's life.

When time cannot be compressed, the test is usually conducted at higher stress levels than those at normal use. For example, assuming the normal operating temperature of a computer is 25°C, we may accelerate the test by subjecting the critical components of the computer to a temperature of 100°C or higher. This causes the failure of the components to occur in a shorter time. Obviously, the higher the stress, the shorter the time needed for the failures to occur. Such accelerated testing should be carefully designed in order not to induce different failure modes than those that occur at normal operating conditions. The types of stresses, stress levels, test durations, and others are discussed in details in Chapter 6.

5-6-1

Suppose we place n units under test for a period of time T. We record the failure times of r failed units as $t_1, t_2, \ldots, t_r \leq T$. The test is terminated at time T with $n - r$ surviving (nonfailed) units. The number of failures, r, is a random variable that depends on the duration of the test and the applied stress level and stress type. Analysis cannot be performed about the reliability and failure rate of the units if no failures occur during T. Therefore, it is important to determine T such that at least some units fail during the test. The time T at which the test is terminated is referred to as the test censoring time, and this type of censoring is referred to as Type 1 censoring.

5-6-2

Suppose we place n units under test and the exact failure times of failed units are recorded. The test continues to run until exactly r failures occur. The test is terminated at tr. Since we specify r failures in advance, we know exactly how much data will be obtained from the test. It is obvious that this type of testing guarantees that failure times will occur and reliability analysis of the data is assured. Of course, the accuracy of reliability analysis is dependent on the number of failure times recorded. The test duration, T, is a random variable which depends on the value of r and the applied stress level. In this type of test, the censoring parameter is the number of failures, r, during the test. It is usually preferred to Type 1 censoring.

5-6-3

Random censoring arises when, for example, n units (devices) are divided among two or more independent test equipment. Suppose after time tf has elapsed, we observe a failure of one of the test equipment. The units placed on this test equipment are removed from the test while the remaining units on the other test equipment continue until the test is completed. The time at which we observe the failure of the test equipment is called the censoring time of units. Since the failure time of the test equipment is a random variable, we refer to this type of censoring as random censoring. There are other types of censoring which are used for specific purposes. Suppose for example that n units are placed on a test at the same time and at predetermined time τ1, r1 surviving units are randomly removed from the test and (n − r1) units continue on test. At a second predetermined time τ2, r2 surviving units are randomly removed from the test and the remaining (n − r1 − r2) continue on test. This process continues until a predetermined time point τs is reached (test termination time) or when all the units fail. When r1 = r2 = . . . = rs−1 = 0, the progressive censoring becomes Type 1 censoring. If the r1 . . . rs−1 surviving units are removed from the test following the first to the (s − 1)th failure, the test is generalized to progressive Type 2 censoring. In the following sections, we analyze the failure data obtained from reliability testing assuming that the testing is conducted at normal conditions. We start by using parametric

fittings for the data when failure times of all units under test are known and when censoring exists.

5-6-4

As discussed in Chapter 1, the hazard rate for a time interval is the ratio between the number of units failed during the time interval and the number of surviving units at the beginning of the interval divided by the length of the interval. Censored units during an interval should not be counted as part of the failed units during that interval. Otherwise, the hazard rate will be inflated. The following example illustrates the necessary calculations for both the hazard rate and the cumulative hazard under censoring.

5-8

The Rayleigh distribution exhibits a linearly increasing hazard function with time. This implies that when the time to failure follows the Rayleigh distribution, an intense aging of the equipment takes place, and the failures do not satisfy the conditions of a stationary random process. More importantly, during the early life of a product where the hazard rate is small, the probability of failure-free operation of the product (or system) decreases with time more slowly than in the case of the exponential distribution. However, as the time increases, the probability of failure-free operation decreases with time at a faster rate than the exponential distribution. Rayleigh distribution is useful in modeling rapidly fading communication channels where the amplitude of the signal can be described by such a distribution.