6.3 Availability

A key assumption in point process models is that the amount of time that an item spends out of service being repaired is negligible. This assumption may not always be appropriate. Many settings call for a model which explicitly considers the repair time. This section considers *availability* and a simple maintainability model known as an *alternating renewal process.*

As in the previous section, X_i denotes the *i*th time to failure, for $i = 1, 2, \ldots$. Now define R_i to be the *i*th time to repair, for $i = 1, 2, \ldots$. Figure 6.16 shows a realization of a process that alternates between the functioning and failed states as time passes. An $\times$ denotes a failure time and a $\circ$ denotes a time when the item is returned to service. In most applications, the expected value of X_i is significantly larger than the expected value of R_i, which means that the item spends significantly more time functioning than being repaired. When the times to failure $X_1, X_2, \ldots$ are independent and identically distributed and the times to repair $R_1, R_2, \ldots$ are independent and identically distributed, with two (typically different) probability distributions, the probability model is known as an *alternating renewal process*. The time until the next event, which is either a failure or a repair completion, alternates between the two probability distributions. The alternating renewal process would be an appropriate model for

- a standby model consisting of items with identically distributed lifetimes and identically distributed times to detect failure and replace the failed item with a standby item, or
- an item with perfect repairs and identically distributed times to repair.

The alternating renewal process would *not* be an appropriate model for

- an improving (or deteriorating) item with minimal repairs, or
- an item whose repair times tend to lengthen (or shorten) as the item ages.

The fact that R_i is being modeled as a single random variable can be misleading. As shown in Figure 6.17, R_i may consist of several components, each corresponding to a stage in the repair process. First, when a failure occurs, a period of time may pass before the failure is detected. Second, once the failure is detected, it may take time to diagnose the problem and determine what action should be taken in order to bring the item back into working condition. Third, in some situations there is a time delay in order to obtain parts and labor. Fourth, once the parts and labor are acquired, there is a time delay to perform the repair. Finally, there may be some testing at the repair facility prior to returning the item to service. In the models presented here, all five of these components of the repair time of an item are collapsed into a single random variable, R_i, called the *repair time.*

The state of a binary repairable item at time t indicates the status of the item, either functioning or failed.

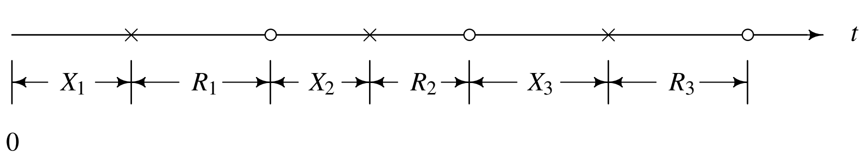

Figure 6.16: Failure and repair process realization for an alternating renewal process.

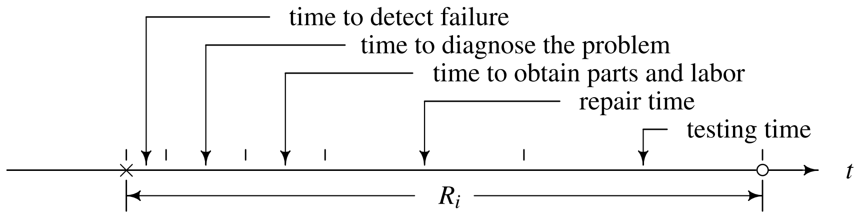

Figure 6.17: Partitioning the repair time.

Definition 6.8 The state of a repairable item at time $t > 0$ is

$$X(t) = \begin{cases} 0 & \text{if the item is failed at time } t \\ 1 & \text{if the item is functioning at time } t. \end{cases}$$

In Chapters 2 and 3, $X(t)$ was simply a random variable that transitioned from state 1 to state 0 at time T as governed by a probability distribution. This probability distribution could be defined by the survivor function $S(t)$ or any of the other lifetime distribution representations given in Section 3.1. As indicated by the realization presented in Figure 6.18, where again an $\times$ denotes a failure time and a $\circ$ denotes the time of completion of a repair, $X(t)$ for a repairable item is a function that alternates between 0 and 1 as time passes. One measure of performance associated with a repairable item is the fraction of time the item spends in state 1, which is often called its *availability*. Tracking the cumulative fraction of time that the item spends operating in the previous realization results in the graph in Figure 6.19. If the distributions of the X_i's and R_i's do not change with time, then the fraction of time that the item spends in the functioning state approaches a constant value as time increases. This is not the case, however, for most repairable items because they either improve or deteriorate. The graph in Figure 6.19 behaves in an analogous fashion to a grade point average. Grades achieved in the freshman and sophomore years have a significant impact on the grade point average. But by the senior year it is much more difficult to impact the grade point average. Likewise, the availability is highly affected by the state of the item early on, but as time passes the current state of the item has decreasing impact.

There are four different definitions for availability. The mathematical definitions of the four varieties are given next.

Definition 6.9 The *availability* of an item is the probability that it is functioning. Four types of availability measures are

- point availability: $A(t) = P\big[X(t) = 1\big] = E\big[X(t)\big]$, for any $t > 0$,
- limiting availability: $A = \lim_{t \to \infty} A(t)$,
- average availability on $(0, c]$: $A_c = \frac{1}{c}\int_0^c A(t)\,dt$, for any $c > 0$,
- limiting average availability on $(0, c]$: $A_\infty = \lim_{c \to \infty} A_c$.

The first type of availability, *point availability*, is the probability that the item is operating at

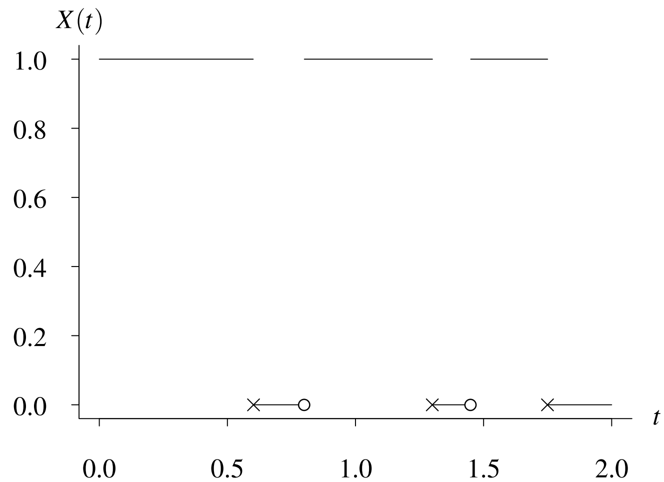

Figure 6.18: A realization of $X(t)$.

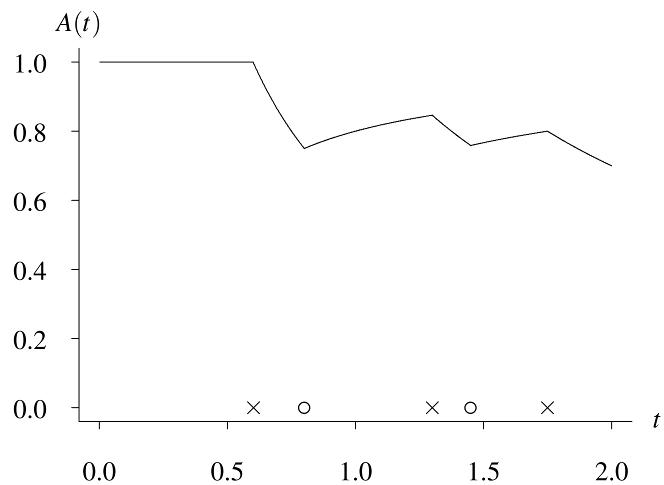

Figure 6.19: A realization of the point availability function $A(t)$.

time t. If repair is not possible, then $A(t)$ simplifies to a survivor function. If the item has stationary times to failure X_i and stationary times to repair R_i, then the point availability approaches a constant value as time increases. This value is known as the *limiting availability* and represents the long-term fraction of time that an item is available. If, for example, the item of interest is a jet aircraft and $A = 0.96$, then the conclusion is that in the long run the aircraft will be available 96% of the time. The *average availability* on $(0, c]$ is the expected fraction of the time the item is operating in the first

c time units it is put into service:

$$A_c = \frac{1}{c} E\left[\int_0^c X(t)\,dt\right] = \frac{1}{c}\int_0^c E[X(t)]\,dt = \frac{1}{c}\int_0^c A(t)\,dt.$$

Finally, the *limiting average availability* is the expected fraction of the time that the item is operating. If the limiting availability exists, then the limiting availability and the limiting average availability are equal.

The simplest instance of an alternating renewal process occurs when X_i and R_i are exponential random variables. More complicated situations, such as nonstationarity or nonexponential failure and repair times, are often evaluated by Monte Carlo simulation rather than using the analytic approach which is presented here. Let $X_1, X_2, \ldots$ be independent and identically distributed exponential random variables with failure rate λ_0, and let $R_1, R_2, \ldots$ be independent and identically distributed exponential random variables with repair rate λ_1. (Some authors use μ, rather than λ_1, for the repair rate. The notation λ_1 is used here to avoid possible confusion with the expected value of a random variable.) The repairs are assumed to be perfect repairs. The *transition diagram* for this alternating renewal process is shown in Figure 6.20. The failure rate λ_0 is the rate at which the item transitions from the operating state to the failed state (that is, from state 1 to state 0). The repair rate λ_1 is the rate at which the item transitions from the failed state to the operating state (that is, from state 0 to state 1). This model is a special case of an alternating renewal process in which the time to every other renewal has the same exponential distribution.

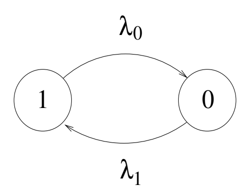

Figure 6.20: Transition diagram for an alternating renewal process.

There are numerous ways to find the point availability $A(t)$, the probability that the item will be available at time t, in the case of exponential times to failure and exponential times to repair. Perhaps the simplest is to use conditioning. For a small time increment Δt,

$$\begin{aligned} P\big[\text{item available at } t+\Delta t\big] = {} & P\big[\text{item does not fail in } (t, t+\Delta t]\big]\, P\big[\text{item available at } t\big] + \\ & P\big[\text{item repaired in } (t, t+\Delta t]\big]\, P\big[\text{item not available at } t\big]. \end{aligned}$$

In terms of $A(t)$, λ_0, and λ_1, this is

$$A(t+\Delta t) = (1-\lambda_0\Delta t)A(t) + \lambda_1\Delta t\big(1-A(t)\big).$$

Rearranging terms,

$$\frac{A(t+\Delta t)-A(t)}{\Delta t} = -(\lambda_0+\lambda_1)A(t)+\lambda_1.$$

Taking the limit of both sides of this equation as Δt approaches 0 yields

$$A'(t) = -(\lambda_0+\lambda_1)A(t)+\lambda_1$$

for $t > 0$ via the definition of the derivative. This is an exact ordinary differential equation that can be solved with the aid of an integrating factor. Using the initial condition $A(0) = 1$, which means that the item is operating initially, the solution is

$$A(t) = \frac{\lambda_1}{\lambda_0 + \lambda_1} + \frac{\lambda_0}{\lambda_0 + \lambda_1} e^{-(\lambda_0 + \lambda_1)t} \qquad t > 0.$$

Several features of this point availability function make it particularly easy to analyze. First, note that the limiting availability is

$$A = \lim_{t \to \infty} A(t) = \frac{\lambda_1}{\lambda_0 + \lambda_1}.$$

Equivalently, since the mean time to failure (MTTF) is $1/\lambda_0$ and the mean time to repair (MTTR) is $1/\lambda_1$,

$$A = \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}}.$$

A graph of $A(t)$ for an item with failure rate $\lambda_0 = 1$ and repair rate $\lambda_1 = 2$ is shown in Figure 6.21. For these parameter settings, the limiting availability, illustrated by the horizontal dashed line in Figure 6.21, is

$$A = \frac{2}{2+1} = \frac{2}{3},$$

which indicates that the item will be available $2/3$ of the time in the long run. Second, the average availability on $(0, c]$ can easily be determined:

$$\begin{aligned} A_c &= \frac{1}{c} \int_0^c A(t)\, dt \\ &= \frac{1}{c} \left[\int_0^c \left(\frac{\lambda_1}{\lambda_0 + \lambda_1} + \frac{\lambda_0}{\lambda_0 + \lambda_1} e^{-(\lambda_0 + \lambda_1)t} \right) dt \right] \\ &= \frac{1}{c} \left[\frac{\lambda_1 c}{\lambda_0 + \lambda_1} - \frac{\lambda_0}{(\lambda_0 + \lambda_1)^2} e^{-(\lambda_0 + \lambda_1)c} + \frac{\lambda_0}{(\lambda_0 + \lambda_1)^2} \right]. \end{aligned}$$

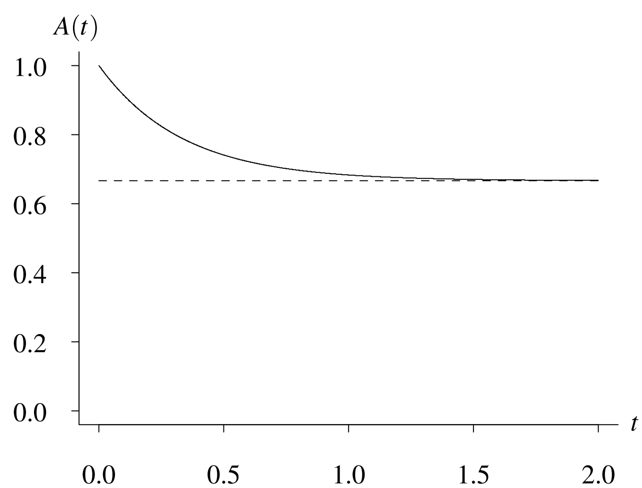

Figure 6.21: Point availability function.