Introduction to Stochastic Processes

8

Hoang Pham

Contents

Abstract

This chapter briefly discusses stochastic processes, including Markov processes, Poisson processes, renewal processes, quasi-renewal processes, and nonhomogeneous Poisson processes. The chapter also provides a short list of books for readers who are interested in advanced topics in stochastic processes.

Keywords

Markov processes · Poisson processes · Renewal processes · Stochastic processes · Nonhomogeneous Poisson processes

H. Pham (✉)
Department of Industrial and Systems Engineering, Rutgers University, Piscataway, NJ, USA
e-mail: hopham@soe.rutgers.edu

8.1 Introduction

Stochastic processes are used to describe the operation of a system over time. There are two main types of stochastic processes: continuous and discrete. A complex continuous process is a process describing a system transition from state to state. The simplest process that will be discussed here is a Markov process. In this case, the future behavior of the process does not depend on its past or present behavior. In many systems that arise in practice, however, past and present states of the system influence the future states, even if they do not uniquely determine them.

8.2 Markov Processes

In this section, we will discuss discrete stochastic processes. As an introduction to the Markov process, let us examine the following example.

Example 8.1 Consider a parallel system consisting of two components (see Fig. 8.1). From a reliability point of view, the states of the system can be described by

State 1:	Full operation (both components operating);
State 2:	One component is operating and one component has failed;
State 3:	Both components have failed.

Define

$$\begin{aligned} P_i(t) &= P\left[X(t) = i\right] \\ &= P\left[\text{system is in state } i \text{ at time } t\right] \end{aligned}$$

H. Pham (ed.), *Springer Handbook of Engineering Statistics*, Springer Handbooks,
https://doi.org/10.1007/978-1-4471-7503-2_8

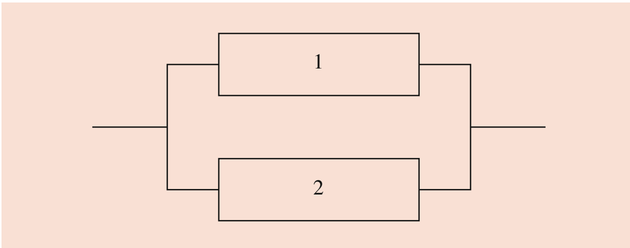

Fig. 8.1 A two-component parallel system

and

$$P_i(t+\mathrm{d}t) = P[X(t+\mathrm{d}t) = i]$$
$$= P[\text{system is in state } i \text{ at time } t+\mathrm{d}t].$$

Define a random variable $X(t)$ which can assume the values 1, 2, or 3 corresponding to the states mentioned above. Since $X(t)$ is a random variable, one can discuss $P[X(t) = 1]$, $P[X(t) = 2]$ or the conditional probability $P[X(t_1) = 2 \mid X(t_0) = 1]$. Again, $X(t)$ is defined as a function of time t, while the conditional probability $P[X(t_1) = 2 \mid X(t_0) = 1]$ can be interpreted as the probability of being in state 2 at time t_1, given that the system was in state 1 at time t_0. In this example, the "state space" is discrete, i.e., 1, 2, 3, etc., and the parameter space (time) is continuous. The simple process described above is called a stochastic process: a process that develops over time (or space) in accordance with some probabilistic (stochastic) laws. There are many types of stochastic processes.

Here we emphasize the Markov process, which is a special type of stochastic process. Let the system be observed at discrete moments of time t_n, where $n = 0, 1, 2, \ldots$, and let $X(t_n)$ denote the state of the system at time t_n.

Definition 8.1 Let $t_0 < t_1 < \ldots < t_n$. If

$$P\Big[X(t_n) = x_n \mid X(t_{n-1}) = x_{n-1}, X(t_{n-2}) = x_{n-2}, \ldots, X(t_0) = x_0\Big] = P[X(t_n) = x_n \mid X(t_{n-1}) = x_{n-1}] \tag{8.1}$$

then the process is called a *Markov process*.

From the definition of a Markov process, given the present state of the process, its behavior in the future does not depend on its behavior in the past.

The essential characteristic of a Markov process is that it is a process that has no memory; its future is determined by the present and not the past. If, in addition to having no memory, the process is such that it depends only on the difference $(t + \mathrm{d}t) - t = \mathrm{d}t$ and not the value of t – in other words $P[X(t + \mathrm{d}t) = j \mid X(t) = i]$ is independent of t – then the process is Markov with stationary transition probabilities or is homogeneous in time. This is the same property noted in exponential event times; in fact, referring back to the graphical representation of $X(t)$, the times between state changes are exponential if the process has stationary transition probabilities.

Thus, a Markov process which is homogeneous in time can describe processes with exponential event occurrence times. The random variable of the process is $X(t)$, the state variable rather than the time to failure used in the exponential failure density. To illustrate the types of processes that can be described, we now review the exponential distribution and its properties. Recall that, if $X_1, X_2, \ldots, X_n$, are independent random variables, each with exponential density and a mean of $1/\lambda_i$, then $\min\{X_1, X_2, \ldots, X_n\}$ has an exponential density with a mean of $(\sum \lambda_i)^{-1}$.

The significance of this property is as follows:

1. The failure behavior of components operated simultaneously can be characterized by an exponential density with a mean equal to the reciprocal of the sum of the failure rates.
2. The joint failure/repair behavior of a system where components are operating and/or undergoing repair can be characterized by an exponential density with a mean equal to the reciprocal of the sum of the failure and repair rates.
3. The failure/repair behavior of a system similar to that described in (2) above but further complicated by active and dormant operating states and sensing and switching can be characterized by an exponential density.

The above property means that almost all reliability and availability models can be characterized by a time-homogeneous Markov process if the various failure times and repair times are exponential. The notation for the Markov process is $\{X(t), t > 0\}$, where $X(t)$ is discrete (state space) and t is continuous (parameter space). By convention, this type of Markov process is called a continuous-parameter Markov chain.

From a reliability/availability viewpoint, there are two types of Markov processes. These are defined as follows:

1. *Absorbing process:* Contains an "absorbing state," which is a state that, once entered, the system can never leave (e.g., a failure which aborts a flight or a mission).
2. *Ergodic process:* Contains no absorbing states, meaning that $X(t)$ can move around indefinitely (e.g., the operation of a ground power plant where failure only temporarily disrupts the operation).

Figure 8.2 presents a summary of Markov processes broken down into absorbing and ergodic categories. Both the reliability and the availability can be described in terms of the probability of the process or system being in defined "up" states, e.g., states 1 and 2 in the initial example. Likewise,

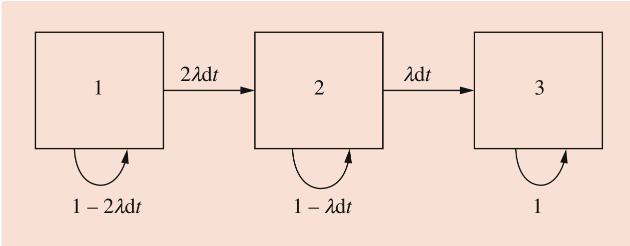

Fig. 8.2 State transition diagram for a two-component system

the mean time between failures (MTBF) can be described as the total time spent in the "up" states before proceeding to the absorbing state or failure state.

Define the incremental transition probability as

$$P_{ij}\,(\mathrm{d}t) = P\left[X\,(t + \mathrm{d}t) = j | X(t) = i\right].$$

This is the probability that the process [random variable $X(t)$] will move to state j during the increment t to $(t + \mathrm{d}t)$, given that it was in state i at time t. Since we are dealing with time-homogeneous Markov processes (exponential failure and repair times), the incremental transition probabilities can be derived from an analysis of the exponential hazard function. It was shown that the hazard function for an exponential with a mean of $1/\lambda$ was just λ. This means that the limiting (as $\mathrm{d}t \to 0$) conditional probability of an event occurring between t and $t + \mathrm{d}t$, given that an event had not occurred at time t, is simply λ, in other words:

$$h(t) = \lim_{\mathrm{d}t \to 0} \frac{P\left[t < X < t + \mathrm{d}t | X > t\right]}{\mathrm{d}t} = \lambda.$$

The equivalent statement for the random variable $X(t)$ is

$$h(t)\;\mathrm{d}t = P\left[X\,(t + \mathrm{d}t) = j | X(t) = i\right] = \lambda\;\mathrm{d}t.$$

Now, $h(t)$ dt is in fact the incremental transition probability, so $P_{ij}(\mathrm{d}t)$ can be stated in terms of the basic failure and/or repair rates. Define

$P_{\mathrm{i}}(t)$: the probability that the system is in state i at time t
$r_{ij}(t)$: transition rate from state i to state j

In general, the differential equations can be written as follows:

$$\frac{\partial P_i(t)}{\partial t} = -\sum_j r_{ij}(t) P_i(t) + \sum_j r_{ji}(t) P_j(t). \qquad (8.2)$$

Solving the above differential equations, one can obtain the time-dependent probability of each state.

Returning to Example 8.1, it is easy to construct a state transition showing the incremental transition probabilities between all possible states for the process:

State 1:	Both components operating
State 2:	One component up and one component down
State 3:	Both components down (absorbing state)

The loops in Fig. 8.2 indicate the probability of remaining in the present state during the dt increment

$$\begin{array}{ll} P_{11}\,(\mathrm{d}t) = 1 - 2\lambda\;\mathrm{d}t & P_{12}\,(\mathrm{d}t) = 2\lambda\;\mathrm{d}t \\ P_{21}\,(\mathrm{d}t) = 0 & P_{22}\,(\mathrm{d}t) = 1 - \lambda\;\mathrm{d}t \\ P_{31}\,(\mathrm{d}t) = 0 & P_{32}\,(\mathrm{d}t) = 0 \\ P_{13}\,(\mathrm{d}t) = 0 & \\ P_{23}\,(\mathrm{d}t) = \lambda\;\mathrm{d}t & \\ P_{33}\,(\mathrm{d}t) = 1 & \end{array}$$

The zeros for P_{ij}, $i > j$ show that the process cannot go backwards: this is not a repair process. The zero on P_{13} shows that, for a process of this type, the probability of more than one event (e.g., failure, repair, etc.) occurring in the incremental time period dt approaches zero as dt approaches zero.

Except for the initial conditions of the process (the state in which the process starts), the process is completely specified by incremental transition probabilities. The reason that this is useful is that assuming exponential event (failure or repair) times allows us to characterize the process at any time t, since the process depends only on what happens between t and $(t + \mathrm{d}t)$. The incremental transition probabilities can be arranged into a matrix in a way that depicts all possible statewide movements. Thus, for parallel configurations,

$$\left[P_{ij}\,(\mathrm{d}t)\right] = \begin{pmatrix} 1 - 2\lambda\mathrm{d}t & 2\lambda\mathrm{d}t & 0 \\ 0 & 1 - \lambda\mathrm{d}t & \lambda\;\mathrm{d}t \\ 0 & 0 & 1 \end{pmatrix}$$

for $i, j = 1, 2$, or 3. The matrix $[P_{ij}(\mathrm{d}t)]$ is called the incremental, one-step transition matrix. It is a stochastic matrix (the rows sum to 1.0). As mentioned earlier, this matrix, along with the initial conditions, completely describes the process.

Now, $[P_{ij}(\mathrm{d}t)]$ gives the probabilities of remaining or moving to all of the various states during the interval t to $t + \mathrm{d}t$; hence,

$$\begin{aligned} P_1\,(t + \mathrm{d}t) &= (1 - 2\lambda\;\mathrm{d}t)\,P_1(t) \\ P_2\,(t + \mathrm{d}t) &= 2\lambda\;\mathrm{d}t P_1(t)\,(1 - \lambda\;\mathrm{d}t)\,P_2(t) \\ P_3\,(t + \mathrm{d}t) &= \lambda\;\mathrm{d}t P_2(t) + P_3(t) \end{aligned}$$

By algebraic manipulation, we have

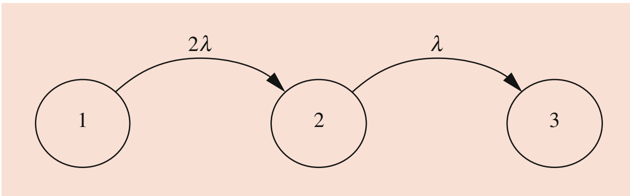

Fig. 8.3 Markov transition rate diagram for a two-component parallel system

$$\frac{[P_1\,(t+\mathrm{d}t)-P_1(t)]}{\mathrm{d}t}=-2\lambda\;P_1(t),$$
$$\frac{[P_2\,(t+\mathrm{d}t)-P_2(t)]}{\mathrm{d}t}=2\lambda\;P_1(t)-\lambda\;P_2(t),$$
$$\frac{[P_3\,(t+\mathrm{d}t)-P_3(t)]}{\mathrm{d}t}=\lambda\;P_2(t).$$

Taking limits of both sides as $\mathrm{d}t\rightarrow 0$, we obtain (also see Fig. 8.3):

$$\begin{aligned}P_1'(t)&=-2\lambda P_1(t),\\ P_2'(t)&=2\lambda P_1(t)-\lambda P_2(t),\\ P_3'(t)&=\lambda P_2(t).\end{aligned}\qquad(8.3)$$

The above system of linear first-order differential equations can be easily solved for $P_1(t)$ and $P_2(t)$, meaning that the reliability of the configuration can be obtained:

$$R(t)=\sum_{i=1}^{2}P_i(t).\qquad(8.4)$$

Actually, there is no need to solve all three equations, only the first two, because $P_3(t)$ does not appear and also $P_3(t)=[1-P_1(t)]-P_2(t)$. The system of linear, first-order differential equations can be solved by various means, including both manual and machine methods. We use manual methods employing the Laplace transform (see Appendix A) here.

$$\begin{aligned}L\,[P_i(t)]&=\int_0^\infty \mathrm{e}^{-st}P_i(t)\mathrm{d}t=f_i(s),\\ L\,[P_i'(t)]&=\int_0^\infty \mathrm{e}^{-st}P_i'(t)\mathrm{d}t=sf_i(s)-P_i(0).\end{aligned}\qquad(8.5)$$

Application of the Laplace transform will allow us to transform the system of linear, first-order differential equations into a system of linear algebraic equations that can easily be solved, and solutions of $P_i(t)$ can be determined via the inverse transforms.

Returning to the example, the initial condition of a parallel configuration is assumed to be "fully up", such that

$$P_1\,(t=0)=1,P_2\,(t=0)=0,P_3\,(t=0)=0.$$

Transforming the equations for $P_1'(t)$ and $P_2'(t)$ gives

$$\begin{aligned}s\,f_1(s)-P_1(t)|_{t=0}&=-2\lambda\,f_1(s),\\ s\,f_2(s)-P_2(t)|_{t=0}&=2\lambda\,f_1(s)-\lambda\,f_2(s).\end{aligned}$$

Evaluating $P_1(t)$ and $P_2(t)$ at $t=0$ gives

$$\begin{aligned}s\,f_1(s)-1&=-2\lambda\,f_1(s),\\ s\,f_2(s)-0&=2\lambda\,f_1(s)-\lambda\,f_2(s).\end{aligned}$$

from which we obtain

$$\begin{aligned}(s+2\lambda)\;f_1(s)&=1,\\ -2\lambda\,f_1(s)+(s+\lambda)\;f_2(s)&=0.\end{aligned}$$

Solving the above equations for $f_1(s)$ and $f_2(s)$, we have

$$\begin{aligned}f_1(s)&=\frac{1}{(s+2\lambda)},\\ f_2(s)&=\frac{2\lambda}{[(s+2\lambda)\,(s+\lambda)]}.\end{aligned}$$

From the inverse Laplace transforms in Appendix A,

$$\begin{aligned}P_1(t)&=\mathrm{e}^{-2\lambda t},\\ P_2(t)&=2\mathrm{e}^{-\lambda t}-2\mathrm{e}^{-2\lambda t},\\ R(t)&=P_1(t)+P_2(t)=2\mathrm{e}^{-\lambda t}-\mathrm{e}^{-2\lambda t}.\end{aligned}\qquad(8.6)$$

The example given above is that of a simple absorbing process where we are concerned about reliability. If a repair capability were added to the model in the form of a repair rate μ, the methodology would remain the same, with only the final result changing. With a repair rate μ added to the parallel configuration (see Fig. 8.4), the incremental transition matrix would be

$$[P_{ij}\,(\mathrm{d}t)]=\begin{pmatrix}1-2\lambda\;\mathrm{d}t & 2\lambda\;\mathrm{d}t & 0\\ \mu\;\mathrm{d}t & 1-(\lambda+\mu)\,\mathrm{d}t & \lambda\;\mathrm{d}t\\ 0 & 0 & 1\end{pmatrix}.$$

The differential equations would become (see Fig. 8.4)

$$\begin{aligned}P_1'(t)&=-2\lambda P_1(t)+\mu P_2(t),\\ P_2'(t)&=2\lambda P_1(t)+(\lambda+\mu)\,P_2(t),\end{aligned}$$

and the transformed equations would become

$$\begin{aligned}(s+2\lambda)\;f_1(s)-\mu f_2(s)&=1,\\ -2\lambda\,f_1(s)+(s+\lambda+\mu)\;f_2(s)&=0.\end{aligned}$$

Hence, we obtain

$$\begin{aligned}f_1(s)&=\frac{(s+\lambda+\mu)}{(s-s_1)\,(s-s_2)},\\ f_2(s)&=\frac{2\lambda}{(s-s_1)\,(s-s_2)},\end{aligned}\qquad(8.7)$$

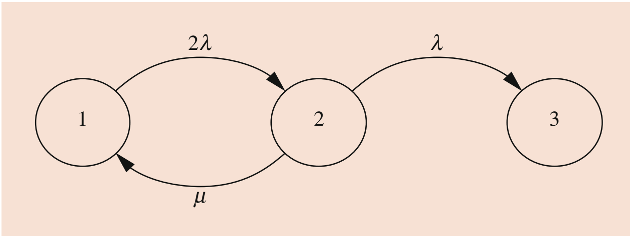

Fig. 8.4 Markov transition rate diagram for a two-component parallel repairable system

where

$$s_1 = \frac{-(3\lambda+\mu)+\sqrt{(3\lambda+\mu)^2-8\lambda^2}}{2},$$
$$s_2 = \frac{-(3\lambda+\mu)-\sqrt{(3\lambda+\mu)^2-8\lambda^2}}{2}. \qquad (8.8)$$

Using the Laplace transform, we obtain

$$P_1(t) = \frac{(s_1+\lambda+\mu)\,e^{-s_1 t}}{(s_1-s_2)} + \frac{(s_2+\lambda+\mu)\,e^{-s_2 t}}{(s_2-s_1)},$$
$$P_2(t) = \frac{2\lambda e^{-s_1 t}}{(s_1-s_2)} + \frac{2\lambda e^{-s_2 t}}{(s_2-s_1)}, \qquad (8.9)$$

where s_1 and s_2 are given in Eq. (8.8).

Thus, the reliability of a two-component parallel repairable system is given by

$$R(t) = P_1(t) + P_2(t) = \frac{(s_1+3\lambda+\mu)\,e^{-s_1 t} - (s_2+3\lambda+\mu)\,e^{-s_2 t}}{(s_1-s_2)} \qquad (8.10)$$

Example 8.2 Consider a three-unit shared load parallel system where

λ_0 is the constant failure rate of a unit when all the three units are operational;
λ_h is the constant failure rate of each of the two surviving units, each of which shares half of the total load; and
λ_f is the constant failure rate of a unit at full load.

For a shared-load parallel system to fail, all the units in the system must fail. We now derive the reliability of a three-unit shared-load parallel system using the Markov model.

In reliability analysis, for the three-unit load-sharing system to work the following events would be considered:

Event 1: All the three units are working until the end of the mission time t where each unit shares one-third of the total load.
Event 2: All the three units are working until time t_1 (each shares one-third of the total load). At time t_1, one of the units (say unit 1) fails, and the other two units (say units 2 and 3) remain to work until the mission time t. Here, once a unit fails at time t_1, the remaining two working units would take half each of the total load and have a constant rate λ_h. As for all identical units, there are three possibilities under this situation.
Event 3: All the three units are working until time t_1 (each shares one-third of the total load) when one (say unit 1) of the three units fails. At time t_2, $(t_2 > t_1)$ one more unit fails (say unit 2) and the remaining unit works until the end of the mission time t. Under this event, there are six possibilities that the probability of two units failing before time t and only one unit remains to work until time t.

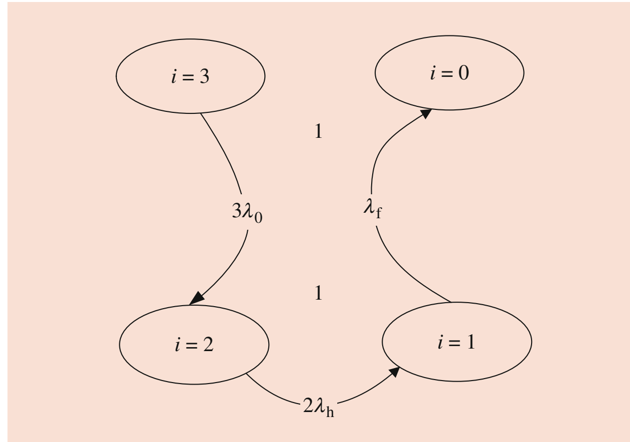

Fig. 8.5 Markov model diagram for a three-unit shared-load parallel system

Define state i represents that i components are working. Let $P_i(t)$ denote the probability that the system is in state i at time t for $i = 0,1,2,3$. Figure 8.5 below shows the Markov diagram of the system.

The Markov modeling system of differential equations based on Fig. 8.1 can be easily derived as follows:

$$\begin{cases} \frac{dP_3(t)}{dt} = -3\lambda_0 P_3(t) \\ \frac{dP_2(t)}{dt} = 3\lambda_0 P_3(t) - 2\lambda_h P_2(t) \\ \frac{dP_1(t)}{dt} = 2\lambda_h P_2(t) - \lambda_f P_1(t) \\ \frac{dP_0(t)}{dt} = \lambda_f P_1(t) \\ P_3(0) = 1 \\ P_j(0) = 0, j \neq 3 \\ P_0(t) + P_1(t) + P_2(t) + P_3(t) = 1 \end{cases} \qquad (8.11)$$

Solving the above differential equations using the Laplace transform method (see Appendix A), we can easily obtain the following results:

$$P_3(t) = e^{-3\lambda_0 t}$$

$$P_2(t) = \frac{3\lambda_0}{3\lambda_h - 3\lambda_0}\left(e^{-3\lambda_0 t} - e^{-2\lambda_h t}\right)$$

$$P_1(t) = \frac{6\lambda_0\lambda_h}{(2\lambda_h - 3\lambda_0)} \left[\frac{e^{-3\lambda_0 t}}{(\lambda_f - 3\lambda_0)} - \frac{e^{-2\lambda_h t}}{(\lambda_f - 2\lambda_h)} + \frac{(2\lambda_h - 3\lambda_0)\, e^{-\lambda_f t}}{(\lambda_f - 3\lambda_0)(\lambda_f - 2\lambda_h)}\right] \tag{8.12}$$

Hence, the reliability of a three-unit shared load parallel system can be obtained as follows:

$$\begin{aligned} R(t) &= P_3(t) + P_2(t) + P_1(t) \\ &= e^{-3\lambda_0 t} + \frac{3\lambda_0}{2\lambda_h - 3\lambda_0}\left(e^{-3\lambda_0 t} - e^{-2\lambda_h t}\right) \\ &\quad + \frac{6\lambda_0\lambda_h}{(2\lambda_h - 3\lambda_0)}\left[\frac{e^{-3\lambda_0 t}}{(\lambda_f - 3\lambda_0)} - \frac{e^{-2\lambda_h t}}{(\lambda_f - 2\lambda_h)} \right. \\ &\quad \left. + \frac{(2\lambda_h - 3\lambda_0)\, e^{-\lambda_f t}}{(\lambda_f - 3\lambda_0)(\lambda_f - 2\lambda_h)}\right] \end{aligned} \tag{8.13}$$

8.2.1 System Mean Time Between Failures

Another parameter of interest for absorbing Markov processes is the MTBF. Recalling Example 8.1 of a parallel configuration with repair, the differential equations $P_1'(t)$ and $P_2'(t)$ describing the process were

$$\begin{aligned} P_1'(t) &= -2\lambda P_1(t) + \mu P_2(t), \\ P_2'(t) &= 2\lambda P_1(t) - (\lambda + \mu)\, P_2(t). \end{aligned}$$

Integrating both sides of the above equations yields

$$\begin{aligned} \int_0^\infty P_1{}'(t)\mathrm{d}t &= -2\lambda \int_0^\infty P_1(t)\mathrm{d}t + \mu \int_0^\infty P_2(t)\mathrm{d}t, \\ \int_0^\infty P_2{}'(t)\mathrm{d}t &= 2\lambda \int_0^\infty P_1(t)\mathrm{d}t - (\lambda + \mu) \int_0^\infty P_2(t)\mathrm{d}t. \end{aligned}$$

For the repairable system, we have

$$\int_0^\infty R(t)\mathrm{d}t = \text{MTBF}. \tag{8.14}$$

Similarly,

$$\int_0^\infty P_1(t)\mathrm{d}t = \text{mean time spent in state 1, and}$$

$$\int_0^\infty P_2(t)\mathrm{d}t = \text{mean time spent in state 2.}$$

Designating these mean times as T_1 and T_2, respectively, we have

$$\begin{aligned} P_1(t)\mathrm{d}t|_0^\infty &= -2\lambda T_1 + \mu T_2, \\ P_2(t)\mathrm{d}t|_0^\infty &= 2\lambda T_1 - (\lambda + \mu)\, T_2. \end{aligned}$$

But $P_1(t) = 0$ as $t \to \infty$ and $P_1(t) = 1$ for $t = 0$. Likewise, $P_2(t) = 0$ as $t \to \infty$ and $P_2(t) = 0$ for $t = 0$. Thus,

$$\begin{aligned} -1 &= -2\lambda T_1 + \mu T_2, \\ 0 &= 2\lambda T_1 - (\lambda + \mu)\, T_2, \end{aligned}$$

or, equivalently,

$$\begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2\lambda & \mu \\ 2\lambda & -(\lambda + \mu) \end{pmatrix} \begin{pmatrix} T_1 \\ T_2 \end{pmatrix}.$$

Therefore,

$$\begin{aligned} T_1 &= \frac{(\lambda + \mu)}{2\lambda^2}, \quad T_2 = \frac{1}{\lambda}, \\ \text{MTBF} &= T_1 + T_2 = \frac{(\lambda + \mu)}{2\lambda^2} + \frac{1}{\lambda} = \frac{(3\lambda + \mu)}{2\lambda^2}. \end{aligned} \tag{8.15}$$

The MTBF for unmaintained processes is developed in exactly the same way as just shown.

The last case to consider for absorbing processes is that of the availability of a maintained system. The difference between reliability and availability is somewhat subtle for absorbing processes. A good example is that of a communications system where the mission would continue if such a system failed temporarily, but if it failed permanently the mission would be aborted. Consider a cold-standby system consisting of two units: one main unit and one spare unit [1]:

State 1:	Main unit operating and the spare is OK
State 2:	Main unit out and restoration underway
State 3:	Spare unit is installed and operating
State 4:	Permanent failure (no spare available)

The incremental transition matrix is given by

$$[P_{ij}\,(\mathrm{d}t)] = \begin{pmatrix} 1 - \lambda\mathrm{d}t & \lambda\,\mathrm{d}t & 0 & 0 \\ 0 & 1 - \mu\,\mathrm{d}t & \mu\,\mathrm{d}t & 0 \\ 0 & 0 & 1 - \lambda\,\mathrm{d}t & \lambda\,\mathrm{d}t \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We obtain

$$\begin{aligned} P_1'(t) &= -\lambda P_1(t), \\ P_2'(t) &= \lambda P_1(t) - \mu P_2(t), \\ P_3'(t) &= \mu P_2(t) - \lambda P_3(t). \end{aligned}$$

Using the Laplace transform, we obtain the following results.

The probability of full-up performance $P_1(t)$ is given by

$$P_1(t) = e^{-\lambda t}.$$

The probability of a down system that is under repair $P_2(t)$ is

$$P_2(t) = \left(\frac{\lambda}{(\lambda-\mu)}\right)\left(e^{-\mu t} - e^{-\lambda t}\right).$$

Similarly, the probability of a fully up system with no spare available $P_3(t)$ is

$$P_3(t) = \left(\frac{\lambda\mu}{(\lambda-\mu)^2}\right)\left[e^{-\mu t} - e^{-\lambda t} - (\lambda-\mu)\, t\, e^{-\lambda t}\right].$$

Hence, the point availability $A(t)$ is given by

$$A(t) = P_1(t) + P_3(t). \quad (8.16)$$

If average or interval availability is required, this is achieved by

$$\left(\frac{1}{t}\right)\int_0^T A(t)\mathrm{d}t = \left(\frac{1}{t}\right)\int_0^T \left[P_1(t) + P_3(t)\right]\mathrm{d}t,$$

where T is the interval of concern.

Ergodic processes, as opposed to absorbing processes, do not have any absorbing states, and hence movement between states can go on indefinitely. For the latter reason, availability (point, steady-state, or interval) is the only meaningful measure. As an example of an ergodic process, we will use a ground-based power unit configured in parallel.

The parallel units are identical, each with exponential failure and repair times with means $1/\lambda$ and $1/\mu$, respectively. Assume a two-repairmen capability if required (both units down), then

State 1:	Fully up (both units operating)
State 2:	One unit down and under repair (other unit up)
State 3:	Both units down and under repair

It should be noted that, as in the case of failure events, two or more repairs cannot be made in the dt interval.

$$[P_{ij}(\mathrm{d}t)] = \begin{pmatrix} 1-2\lambda\,\mathrm{d}t & 2\lambda\,\mathrm{d}t & 0 \\ \mu\,\mathrm{d}t & 1-(\lambda+\mu)\,\mathrm{d}t & \lambda\mathrm{d}t \\ 0 & 2\mu\,\mathrm{d}t & 1-2\mu\,\mathrm{d}t \end{pmatrix}.$$

Case I: *Point Availability – Ergodic Process.* For an ergodic process, as $t \to \infty$ the availability settles down to a constant level. Point availability allows us to study the process before this "settling down," and it reflects the initial conditions in the process. We can obtain a solution for the point availability in a similar way to that for absorbing processes, except that the last row and column of the transition matrix must be retained and entered into the system of equations. For example, the system of differential equations becomes

$$\begin{pmatrix} P_1'(t) \\ P_2'(t) \\ P_3'(t) \end{pmatrix} = \begin{pmatrix} -2\lambda & \mu & 0 \\ 2\lambda & -(\lambda+\mu) & 2\mu \\ 0 & \lambda & -2\mu \end{pmatrix}\begin{pmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \end{pmatrix}.$$

Similar to the absorbing case, the Laplace transform can be used to solve for $P_1(t)$, $P_2(t)$ and $P_3(t)$; the point availability $A(t)$ is given by

$$A(t) = P_1(t) + P_2(t).$$

Case II: *Interval Availability – Ergodic Process.* This is the same as the absorbing case, with integration over the time period T of interest. The interval availability, $A(T)$, is

$$A(T) = \frac{1}{T}\int_0^T A(t)\mathrm{d}t. \quad (8.17)$$

Case III: *Steady State Availability – Ergodic Process.* Here, the process is examined as $t \to \infty$, with complete "washout" of the initial conditions. By letting $t \to \infty$, the system of differential equations can be transformed into linear algebraic equations. Thus,

$$\lim_{t\to\infty}\begin{pmatrix} P_1'(t) \\ P_2'(t) \\ P_3'(t) \end{pmatrix} = \lim_{t\to\infty}\begin{pmatrix} -2\lambda & \mu & 0 \\ 2\lambda & -(\lambda+\mu) & 2\mu \\ 0 & \lambda & -2\mu \end{pmatrix}\begin{pmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \end{pmatrix}.$$

As $t \to \infty$, $P_i(t) \to$ constant and $P_1'(t) \to 0$. This leads to an unsolvable system, namely,

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2\lambda & \mu & 0 \\ 2\lambda & -(\lambda+\mu) & 2\mu \\ 0 & \lambda & -2\mu \end{pmatrix}\begin{pmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \end{pmatrix}.$$

To avoid the above difficulty, an additional equation is introduced:

$$\sum_{i=1}^{3} P_i(t) = 1.$$

With the introduction of the new equation, one of the original equations is deleted and a new system is formed:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ -2\lambda & \mu & 0 \\ 2\lambda & -(\lambda+\mu) & 2\mu \end{pmatrix}\begin{pmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \end{pmatrix}$$

or, equivalently,

$$\begin{pmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ -2\lambda & \mu & 0 \\ 2\lambda & -(\lambda+\mu) & 2\mu \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

We now obtain the following results:

$$P_1(t) = \frac{\mu^2}{(\mu+\lambda)^2},$$
$$P_2(t) = \frac{2\lambda\mu}{(\mu+\lambda)^2},$$

and

$$P_3(t) = 1 - P_1(t) - P_2(t) = \frac{\lambda^2}{(\mu+\lambda)^2}.$$

Therefore, the steady state availability $A(\infty)$ is given by

$$A_3(\infty) = P_1(t) + P_2(t) = \frac{\mu(\mu+2\lambda)}{(\mu+\lambda)^2}. \tag{8.18}$$

Note that Markov methods can also be employed when failure or repair times are not exponential but can be represented as the sum of exponential times with identical means (an Erlang distribution or gamma distribution with integer-valued shape parameters). Basically, the method involves introducing "dummy" states which, although being of no particular interest in themselves, change the hazard function from constant to increasing.

Example 8.3 A system is composed of eight identical active power supplies, at least seven of the eight are required for the system to function. In other words, when two of the eight power supplies fail, the system fails. When all eight power supplies are operating, each has a constant failure rate λ_a per hour. If one power supply fails, each remaining power supply has a failure rate λ_b per hour where $\lambda_a \leq \lambda_b$ We assume that a failed power supply can be repaired with a constant rate μ per hour. The system reliability function, $R(t)$, is defined as the probability that the system continues to function throughout the interval $(0, t)$. Here we wish to determine the system mean time to failure (MTTF).

Define

State 0: All 8 units are working
State 1: 7 units are working
State 2: More than one unit failed and system does not work
The initial condition: $P_0(0) = 1$, $P_1(0) = P_2(0) = 0$

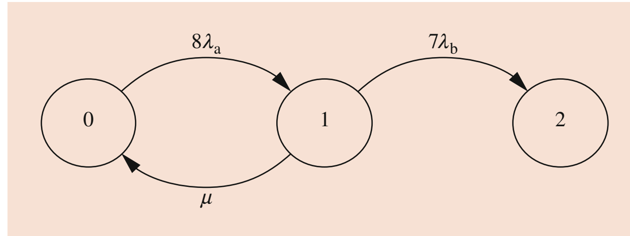

Fig. 8.6 Markov transition rate diagram for a 7-out-8 dependent system

The Markov modeling of differential equations (see Fig. 8.6) can be written as follows:

$$P_0'(t) = -8\lambda_a P_0(t) + \mu P_1(t)$$
$$P_1'(t) = 8\lambda_a P_0(t) - (7\lambda_b + \mu) P_1(t)$$
$$P_2'(t) = 7\lambda_b P_1(t)$$

Using the Laplace transform, we obtain

$$\begin{cases} sF_0(s) - P_0(0) = -8\lambda_a F_0(s) + \mu F_1(s) \\ sF_1(s) - P_1(0) = 8\lambda_a F_0(s) - (7\lambda_b + \mu) F_1(s) \\ sF_2(s) - P_2(0) = 7\lambda_b F_1(s) \end{cases} \tag{8.19}$$

When $s = 0$:

$$F_i(0) = \int_0^\infty P_i(t)\mathrm{d}t.$$

Thus, the system reliability function and system MTTF, respectively, are

$$R(t) = P_0(t) + P_1(t). \tag{8.20}$$

and

$$\mathrm{MTTF} = \int_0^\infty R(t)\mathrm{d}t = \int_0^\infty [P_0(t) + P_1(t)]\,\mathrm{d}t = \sum_{i=1}^{2} F_i(0). \tag{8.21}$$

From Eq. (8.19), when $s = 0$, we have

$$\begin{cases} -1 = -8\lambda_a F_0(0) + \mu F_1(0) \\ 0 = 8\lambda_a F_0(0) - (7\lambda_b + \mu) F_1(0). \end{cases} \tag{8.22}$$

From Eq. (8.22), after some arrangements, we can obtain

$$7\lambda_b F_1(0) = 1 \Rightarrow F_1(0) = \frac{1}{7\lambda_b}$$

and

$$F_0(0) = \frac{7\lambda_b + \mu}{8\lambda_a} F_1(0)$$
$$= \frac{7\lambda_b + \mu}{8\lambda_a} \frac{1}{7\lambda_b} = \frac{7\lambda_b + \mu}{56\lambda_a\lambda_b}$$

From Eq. (8.21), the system MTTF can be obtained

$$\text{MTTF} = \int_0^\infty R(t)\mathrm{d}t = \int_0^\infty [P_0(t) + P_1(t)]\,\mathrm{d}t$$
$$= F_0(0) + F_1(0)$$
$$= \frac{7\lambda_b + \mu}{56\lambda_a\lambda_b} + \frac{1}{7\lambda_b} = \frac{\mu + 8\lambda_a + 7\lambda_b}{56\lambda_a\lambda_b}.$$

Given $\lambda_a = 3 \times 10^{-3} = 0.003$, $\lambda_b = 5 \times 10^{-2} = 0.05$, and $\mu = 0.8$, then the system mean time to failure is given by:

$$\text{MTTF} = \frac{\mu + 8\lambda_a + 7\lambda_b}{56\lambda_a\lambda_b}$$
$$= \frac{0.8 + 8(0.003) + 7(0.05)}{56(0.003)(0.05)} = \frac{1.174}{0.0084} = 139.762 \text{ hours.}$$

8.3 Counting Processes

Among various discrete stochastic processes, counting processes are widely used in engineering statistics to describe the appearance of events in time, such as failures, the number of perfect repairs, etc. The simplest counting process is a Poisson process. The Poisson process plays a special role in many applications related to reliability [1]. A classic example of such an application is the decay of uranium. Here, radioactive particles from nuclear material strike a certain target in accordance with a Poisson process of some fixed intensity. One well-known counting process is the so-called renewal process. This process is described as a sequence of events where the intervals between the events are independent and identically distributed random variables. In reliability theory, this type of mathematical model is used to describe the number of occurrences of an event or the number of renewals (i.e., replacements of objects) over a time interval. A light bulb is shining in your living room and it blows ups suddenly. You replace it by a new bulb. It lasts for a few months, then burns out again. You then replace it again, and so on. One would be interested to know about the total number of bulbs is needed to be replaced in 2 years.

In this subsection, we discuss the concepts and properties of the Poisson process, renewal process, quasi-renewal process, and nonhomogeneous Poisson process.

A non-negative, integer-valued stochastic process $N(t)$ is called a counting process if $N(t)$ represents the total number of occurrences of an event in the time interval $[0, t]$ and satisfies these two properties:

1. If $t_1 < t_2$, then $N(t_1) \le N(t_2)$
2. If $t_1 < t_2$, then $N(t_2) - N(t_1)$ is the number of occurrences of the event in the interval $[t_1, t_2]$.

For example, if $N(t)$ equals the number of persons who have entered a restaurant at or prior to time t, then $N(t)$ is a counting process in which an event occurs whenever a person enters the restaurant.

8.3.1 Poisson Processes

One of the most important counting processes is the Poisson process.

Definition 8.3 A counting process $N(t)$ is said to be a Poisson process with intensity λ if

1. The failure process $N(t)$ has stationary independent increments
2. The number of failures in any time interval of length s has a Poisson distribution with a mean of λs; in other words

$$P\{N(t+s) - N(t) = n\} = \frac{e^{-\lambda s}(\lambda s)^n}{n!} \qquad n = 0, 1, 2, \ldots; \tag{8.23}$$

3. The initial condition is $N(0) = 0$

This model is also called a homogeneous Poisson process, indicating that the failure rate λ does not depend on time t. In other words, the number of failures that occur during the time interval $(t, t+s]$ does not depend on the current time t, only the length of the time interval s. A counting process is said to possess independent increments if the number of events in disjoint time intervals are independent.

For a stochastic process with independent increments, the autocovariance function is

$$\text{Cov}[X(t_1), X(t_2)] = \begin{cases} \text{Var}[N(t_1+s) - N(t_2)] & \text{for } 0 < t_2 - t_1 < s \\ 0 & \text{otherwise} \end{cases},$$

where

$$X(t) = N(t+s) - N(t).$$

If $X(t)$ is Poisson-distributed, then the variance of the Poisson distribution is

$$\text{Cov}[X(t_1), X(t_2)] = \begin{cases} \lambda[s - (t_2 - t_1)] & \text{for } 0 < t_2 - t_1 < s \\ 0 & \text{otherwise} \end{cases}.$$

8

This result shows that the Poisson increment process is covariance stationary. We now present several properties of the Poisson process.

Property 8.1 The sum of independent Poisson processes $N_1(t)$, $N_2(t)$, $\ldots$, $N_k(t)$ with mean values $\lambda_1 t$, $\lambda_2 t, \ldots, \lambda_k t$, respectively, is also a Poisson process with mean $\left(\sum_{i=1}^{k} \lambda_i\right) t$. In other words, the sum of the independent Poisson processes is also a Poisson process with a mean that is equal to the sum of the means of the individual Poisson processes.

Property 8.2 The difference between two independent Poisson processes, $N_1(t)$ and $N_2(t)$, with mean $\lambda_1 t$ and $\lambda_2 t$, respectively, is not a Poisson process. Instead, it has a probability mass function of

$$\begin{aligned} &P[N_1(t) - N_2(t) = k] \\ &= \mathrm{e}^{-(\lambda_1+\lambda_2)t}\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{k}{2}} I_k\left(2\sqrt{\lambda_1\lambda_2 t}\right), \end{aligned}$$

where $I_k(.)$ is a modified Bessel function of order k.

Property 8.3 If the Poisson process $N(t)$ with mean λt is filtered such that not every occurrence of the event is counted, then the process has a constant probability p of being counted. The result of this process is a Poisson process with mean λpt.

Property 8.4 Let $N(t)$ be a Poisson process and Y_i a family of independent and identically distributed random variables which are also independent of $N(t)$. A stochastic process $X(t)$ is said to be a compound Poisson process if it can be represented as

$$X(t) = \sum_{i=1}^{N(t)} Y_i.$$

8.3.2 Renewal Processes

A renewal process is a more general case of the Poisson process in which the inter-arrival times of the process or the times between failures do not necessarily follow the exponential distribution. For convenience, we will call the occurrence of an event a renewal, the inter-arrival time the renewal period, and the waiting time the renewal time.

Definition 8.3 A counting process $N(t)$ that represents the total number of occurrences of an event in the time interval $(0, t]$ is called a renewal process if the times between the failures are independent and identically distributed random variables.

The probability that exactly n failures occur by time t can be written as

$$P[N(t) = n] = P[N(t) \geq n] - P[N(t) > n]. \tag{8.24}$$

Note that the times between the failures are $T_1, T_2, \ldots, T_n$, so the failures occurring at time W_k are

$$W_k = \sum_{i=1}^{k} T_i$$

and

$$T_k = W_k - W_{k-1}.$$

Thus,

$$\begin{aligned} P[N(t) = n] &= P[N(t) \geq n] - P[N(t) > n] \\ &= P[W_n \leq t] - P[W_{n+1} \leq t] \\ &= F_n(t) - F_{n+1}(t), \end{aligned} \tag{8.25}$$

where $F_n(t)$ is the cumulative distribution function for the time of the nth failure and $n = 0, 1, 2, \ldots$.

Example 8.4 Consider a software testing model for which the time at which an error is found during the testing phase has an exponential distribution with a failure rate of X. It can be shown that the time of the nth failure follows the gamma distribution with parameters k and n. From Eq. (8.24), we obtain

$$\begin{aligned} P[N(t) = n] &= P[N(t) \leq n] - P[N(t) \leq n-1] \\ &= \sum_{k=0}^{n} \frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t} - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t} \\ &= \frac{(\lambda t)^n}{n!}\mathrm{e}^{-\lambda t} \quad \text{for} \quad n = 0, 1, 2, \ldots. \end{aligned} \tag{8.26}$$

Several important properties of the renewal function are given below.

Property 8.5 The mean value function of the renewal process, denoted by $m(t)$, is equal to the sum of the distribution functions for all renewal times, that is,

$$\begin{aligned} m(t) &= E[N(t)] \\ &= \sum_{n=1}^{\infty} F_n(t). \end{aligned} \tag{8.27}$$

Property 8.6 The renewal function $m(t)$ satisfies the following equation:

$$m(t) = F_{\mathrm{a}}(t) + \int_0^t m(t-s)\,\mathrm{d}F_{\mathrm{a}}(s), \tag{8.28}$$

where $F_a(t)$ is the distribution function of the inter-arrival time or the renewal period.

In general, let $y(t)$ be an unknown function to be evaluated and $x(t)$ be any non-negative and integrable function associated with the renewal process. Assume that $F_a(t)$ is the distribution function of the renewal period. We can then obtain the following result.

Property 8.7 Let the renewal equation be

$$y(t) = x(t) + \int_0^t y(t-s)\,\mathrm{d}F_a(s). \tag{8.29}$$

Then its solution is given by

$$y(t) = x(t) + \int_0^t x(t-s)\,\mathrm{d}m(s)$$

where $m(t)$ is the mean value function of the renewal process.

The proof of the above property can be easily derived using the Laplace transform. Let $x(t) = a$. Thus, in Property 8.7, the solution $y(t)$ is given by

$$\begin{aligned} y(t) &= x(t) + \int_0^t x(t-s)\,\mathrm{d}m(s). \\ &= a + \int_0^t a\,\mathrm{d}m(s) \\ &= a\left\{1 + E[N(t)]\right\}. \end{aligned}$$

8.3.3 Quasi-Renewal Processes

In this section we discuss a general renewal process: the quasi-renewal process. Let $\{N(t), t > 0\}$ be a counting process and let X_n be the time between the $(n-1)$th and the nth event of this process, $n \geq 1$.

Definition 8.4 [2] If the sequence of non-negative random variables $\{X_1, X_2, \ldots\}$ is independent and

$$X_i = \alpha X_{i-1} \tag{8.30}$$

for $i \geq 2$ where $\alpha > 0$ is a constant, then the counting process $\{N(t),\ t \geq 0\}$ is said to be a quasi-renewal process with parameter α and the first inter-arrival time X_1.

When $\alpha = 1$, this process becomes the ordinary renewal process. This quasi-renewal process can be used to model reliability growth processes in software testing phases and hardware burn-in stages for $\alpha > 1$, and in hardware maintenance processes when $\alpha \leq 1$.

Assume that the probability density function (pdf), cumulative distribution function (cdf), survival function, and failure rate of random variable X_1 are $f_1(x)$, $F_1(x)$, $s_1(x)$ and $r_1(x)$, respectively. Then the pdf, cdf, survival function, and failure rate of X_n for $n = 1, 2, 3, \ldots$ are, respectively, given below [2]:

$$\begin{aligned} f_n(x) &= \frac{1}{\alpha^{n-1}} f_1\left(\frac{1}{\alpha^{n-1}}x\right), \\ F_n(x) &= F_1\left(\frac{1}{\alpha^{n-1}}x\right), \\ s_n(x) &= s_1\left(\frac{1}{\alpha^{n-1}}x\right), \\ f_n(x) &= \frac{1}{\alpha^{n-1}} r_1\left(\frac{1}{\alpha^{n-1}}x\right). \end{aligned} \tag{8.31}$$

Similarly, the mean and variance of X_n is given as

$$\begin{aligned} E(X_n) &= \alpha^{n-1} E(X_1), \\ \mathrm{Var}(X_n) &= \alpha^{2n-2}\mathrm{Var}(X_1). \end{aligned} \tag{8.32}$$

Because of the non-negativity of X_1, and the fact that X_1 is not identically 0, we obtain

$$E(X_1) = \mu_1 \neq 0.$$

It is worth noting that the shape parameters for X_n are the same for $n = 1, 2, 3, \ldots$ for a quasi-renewal process if X_1 follows the gamma, Weibull, or log normal distribution.

This means that the shape parameters of the inter-arrival time will not change after "renewal". In software reliability, the assumption that the software debugging process does not change the error-free distribution seems reasonable. Thus, if a quasi-renewal process model is used, the error-free times that occur during software debugging will have the same shape parameters. In this sense, a quasi-renewal process is suitable for modeling the increase in software reliability. It is worth noting that

$$\begin{aligned} \lim_{n\to\infty} \frac{E(X_1 + X_2 + \cdots + X_n)}{n} &= \lim_{n\to\infty} \frac{\mu_1(1-\alpha^n)}{(1-\alpha)n} \\ &= \begin{cases} 0 & \text{if } \alpha < 1, \\ \infty & \text{if } \alpha > 1. \end{cases} \end{aligned} \tag{8.33}$$

Therefore, if the inter-arrival time represents the error-free time of a software system, then the average error-free time approaches infinity when its debugging process has been operating for a long debugging time.

Distribution of *N(t)*

Consider a quasi-renewal process with parameter α and a first inter-arrival time X_1. Clearly, the total number of renewals $N(t)$ that occur up to time t has the following relationship to the arrival time of the nth renewal SS_n:

$$N(t) \geq n \quad \text{if and only if} \quad \mathrm{SS}_n \leq t.$$

8

In other words, $N(t)$ is at least n if and only if the nth renewal occurs prior to time t. It is easily seen that

$$\mathrm{SS}_n = \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} \alpha^{i-1} X_1 \quad \text{for } n \geq 1. \tag{8.34}$$

Here, $\mathrm{SS}_0 = 0$. Thus, we have

$$\begin{aligned} P\{N(t) = n\} &= P\{N(t) \geq n\} - P\{N(t) \geq n+1\} \\ &= P\{\mathrm{SS}_n \leq t\} - P\{\mathrm{SS}_{n+1} \leq t\} \\ &= G_n(t) - G_{n+1}(t), \end{aligned}$$

where $G_n(t)$ is the convolution of the inter-arrival times F_1, $F_2, F_3, \ldots, F_n$. In other words,

$$G_n(t) = P\{F_1 + F_2 + \cdots + F_n \leq t\}.$$

If the mean value of $N(t)$ is defined as the renewal function $m(t)$, then

$$\begin{aligned} m(t) &= E[N(t)] \\ &= \sum_{n=1}^{\infty} P\{N(t) \geq n\} \\ &= \sum_{n=1}^{\infty} P\{\mathrm{SS}_n \leq t\} \\ &= \sum_{n=1}^{\infty} G_n(t). \end{aligned} \tag{8.35}$$

The derivative of $m(t)$ is known as the renewal density

$$\lambda(t) = m'(t).$$

In renewal theory, random variables representing inter-arrival distributions assume only non-negative values, and the Laplace transform of its distribution $F_1(t)$ is defined by

$$\mathcal{L}\{F_1(s)\} = \int_0^{\infty} \mathrm{e}^{-sx} \mathrm{d}F_1(x).$$

Therefore,

$$\mathcal{L}F_n(s) = \int_0^{\infty} \mathrm{e}^{-a^{n-1}st} \mathrm{d}F_1(t) = \mathcal{L}F_1\left(\alpha^{n-1}s\right)$$

and

$$\begin{aligned} \mathcal{L}m_n(s) &= \sum_{n-1}^{\infty} \mathcal{L}G_n(s) \\ = \sum_{n-1}^{\infty} & \mathcal{L}F_1(s)\mathcal{L}F_1(\alpha s) \cdots \mathcal{L}F_1\left(\alpha^{n-1}s\right). \end{aligned} \tag{8.36}$$

Since there is a one-to-one correspondence between distribution functions and its Laplace transform, it follows that the first inter-arrival distribution of a quasi-renewal process uniquely determines its renewal function.

If the inter-arrival time represents the error-free time (time to first failure), a quasi-renewal process can be used to model reliability growth in both software and hardware.

Suppose that all software faults have the same chance of being detected. If the inter-arrival time of a quasi-renewal process represents the error-free time of a software system, then the expected number of software faults in the time interval $[0, t]$ can be defined by the renewal function, $m(t)$, with parameter $\alpha > 1$. Denoted by $m_\mathrm{r}(t)$, the number of remaining software faults at time t, it follows that

$$m_\mathrm{r}(t) = m(T_\mathrm{c}) - m(t),$$

where $m(T_\mathrm{c})$ is the number of faults that will eventually be detected through a software lifecycle T_c.

8.3.4 Nonhomogeneous Poisson Processes

The nonhomogeneous Poisson process model (NHPP), which represents the number of failures experienced up to time t, is a nonhomogeneous Poisson process $\{N(t)$ with $t \geq 0\}$. The main issue with the NHPP model is to determine an appropriate mean value function to denote the expected number of failures experienced up to a certain time.

Different assumptions mean that the model will end up with different functional forms of the mean value function. Note that the exponential assumption for the inter-arrival time between failures is relaxed in a renewal process, and the stationary assumption is relaxed in the NHPP.

The NHPP model is based on the following assumptions:

- The failure process has an independent increment; in other words, the number of failures during the time interval $(t, t+s)$ depends on the current time t and the length of the time interval s, and does not depend on the past history of the process.
- The failure rate of the process is given by

$$\begin{aligned} &P\{\text{exactly one failure in } (t, t+\Delta t)\} \\ &\quad = P\{N(t+\Delta t) - N(t) = 1\} \\ &\quad = \lambda(t)\Delta t + o(\Delta t), \end{aligned}$$

  where $\lambda(t)$ is the intensity function.
- During a small interval Δt, the probability of more than one failure is negligible; that is,

$$P\{\text{two or more failures in } (t, t+\Delta t)\} = o(\Delta t),$$

- The initial condition is $N(0) = 0$.

Based on these assumptions, the probability that exactly n failures occur during the time interval $(0, t)$ for the NHPP is given by

$$\Pr\{N(t) = n\} = \frac{[m(t)]^n}{n!} e^{-m(t)} \quad n = 0, 1, 2, \ldots, \tag{8.37}$$

where $m(t) = E[N(t)] = \int_0^t \lambda(s)\ \mathrm{d}s$ and $\lambda(t)$ is the intensity function. It is easily shown that the mean value function $m(t)$ is nondecreasing.

The reliability $R(t)$, defined as the probability that there are no failures in the time interval $(0, t)$, is given by

$$\begin{aligned} R(t) &= P\{N(t) = 0\} \\ &= e^{-m(t)}. \end{aligned} \tag{8.38}$$

In general, the reliability $R(x \mid t)$ – the probability that there are no failures in the interval $(t, t + x)$ – is given by

$$\begin{aligned} R(x \mid t) &= P\{N(t+x) - N(t) = 0\} \\ &= e^{-[m(t+x) - m(t)]} \end{aligned}$$

and its density is given by

$$f(x) = \lambda(t+x)\, e^{-[m(t+x) - m(t)]},$$

where

$$\lambda(x) = \frac{\partial}{\partial x}[m(x)].$$

The variance of the NHPP can be obtained as follows:

$$\operatorname{Var}[N(t)] = \int_0^t \lambda(s) \mathrm{d}s$$

and the autocorrelation function is given by

$$\begin{aligned} \operatorname{Cor}[s] &= E[N(t)]\, E[N(t+s) - N(t)] + E\left[N^2(t)\right] \\ &= \int_0^t \lambda(s)\mathrm{d}s \int_0^{t+s} \lambda(s)\mathrm{d}s + \int_0^t \lambda(s)\mathrm{d}s \\ &= \int_0^t \lambda(s)\mathrm{d}s \left[1 + \int_0^{t+s} \lambda(s)\mathrm{d}s\right]. \end{aligned} \tag{8.39}$$

Example 8.5 Assume that the intensity λ is a random variable with pdf $f(\lambda)$. Then the probability that exactly n failures occur during the time interval $(0, t)$ is given by

$$P\{N(t) = n\} = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} f(\lambda)\, \mathrm{d}\lambda. \tag{8.40}$$

If the pdf $f(\lambda)$ is given as the following gamma density function with parameters k and m:

$$f(\lambda) = \frac{1}{\Gamma(m)} k^m \lambda^{m-1} e^{-k\lambda} \quad \text{for } \lambda \geq 0 \tag{8.41}$$

then it can be shown that

$$P\{N(t) = n\} = \binom{n+m-1}{n} p^m q^n \qquad n = 0, 1, 2, \ldots \tag{8.42}$$

(this is also called a negative binomial density function), where

$$p = \frac{k}{t+k} \quad \text{and} \quad q = \frac{t}{t+k} = 1 - p. \tag{8.43}$$

8.4 Further Reading

The reader interested in a deeper understanding of advanced probability theory and stochastic processes should note the following citations, which refer to highly recommended books: Feller [3]; Pinksy and Karlin [4], Parzen [5], Melsa and Sage [6].

Appendix A: Laplace Transformation Functions

Let X be a nonnegative life time having probability density function f. The Laplace transform of a function $f(x)$, denote f^*, is defined as

$$\ell\{f(x)\} = f^*(s) = \int_0^\infty e^{-sx} f(x)\mathrm{d}x \ \text{ for } \ s \geq 0. \tag{8.44}$$

The function f^* is called the Laplace transform of the function f. The symbol ℓ in Eq. (8.44) is called the Laplace transform operator. Note that $f^*(0) = 1$. By taking a differential derivative of $f^*(s)$, we obtain

$$\frac{\partial f^*(s)}{\partial s} = -\int_0^\infty x e^{-sx} f(x)\mathrm{d}x.$$

Substitute $s = 0$ into the above equation, the first derivative of f^*, it yields the negative of the expected value of X or the first moment of X:

$$\left.\frac{\partial f^*(s)}{\partial s}\right|_{s=0} = -E(X).$$

Similarly, the second derivative yields the second moment of X when $s = 0$, that is,

$$\left.\frac{\partial f^*(s)}{\partial s}\right|_{s=0} = \left.\int_0^\infty x^2 e^{-sx} f(x)\mathrm{d}x\right|_{s=0} = E\left(X^2\right).$$

In general, it can be shown that

$$\left.\frac{\partial^n f^*(s)}{\partial^n s}\right|_{s=0} = \left.\int_0^\infty (-x)^n e^{-sx} f(x)\mathrm{d}x\right|_{s=0} = (-1)^n E\left(X^n\right)$$

Note that

$$e^{-sx} = \sum_{n=0}^{\infty} \frac{(-sx)^n}{n!}$$

then $f^*(s)$ can be rewritten as

$$f^*(s) = \sum_{n=0}^{\infty} \frac{(-s)^n}{n!}\mu_n,$$

where

$$\mu_n = (-1)^n \left.\frac{\partial^n f^*(s)}{\partial^n s}\right|_{s=0} = \left.\int_0^\infty x^n e^{-sx} f(x)\mathrm{d}x\right|_{s=0} = E\left(X^n\right).$$

We can easily show that ℓ is a linear operator, that is

$$\ell\left\{c_1 f_1(x) + c_2 f_2(x)\right\} = c_1 \ell\left\{f_1(x)\right\} + c_2 \ell\left\{f_2(x)\right\}.$$

If $\ell\{f(t)\} = f^*(s)$, then we call $f(t)$ the inverse Laplace transform of $f^*(s)$ and write $\ell^{-1}\{f^*(s)\} = f(t)$ A summary of some common Laplace transform functions is listed in Table 8.1.

Example 8.6 Use the Laplace transforms to solve the following

$$\frac{\partial f(t)}{\partial t} + 3f(t) = e^{-t}, \tag{8.45}$$

with an initial condition: $f(0) = 0$. Obtain the solution $f(t)$. Here the Laplace transforms of $\frac{\partial f(t)}{\partial t}$, $f(t)$, and e^{-t} are

$$sf^*(s) - f(0), f^*(s), \text{ and } \frac{1}{s+1}$$

respectively. Thus, the Laplace transform of Eq. (8.45) is given by

$$sf^*(s) - f(0) + 3f^*(s) = \frac{1}{s+1}.$$

Since $f(0) = 0$ we have

$$(s+3)f^*(s) = \frac{1}{s+1} \text{ or } f^*(s) = \frac{1}{(s+1)(s+3)}.$$

Table 8.1 List of common Laplace transforms

$f(t)$	$\ell\{f(t)\} = f^*(s)$
$f(t)$	$f^*(s) = \int_0^\infty e^{-st} f(t)\mathrm{d}t$
$\frac{\partial f(t)}{\partial t}$	$s f^*(s) - f(0)$
$\frac{\partial^2}{\partial t^2}[f(t)]$	$s^2 f^*(s) - s\,f(0) - \frac{\partial}{\partial t}f(0)$
$\frac{\partial^n}{\partial t^n}[f(t)]$	$s^n f^*(s) - s^{n-1} f(0) - \cdots - \frac{\partial^{n-1}}{\partial t^{n-1}}f(0)$
$f(at)$	$\frac{1}{a}f^*\left(\frac{s}{a}\right)$
1	$\frac{1}{s}$
t	$\frac{1}{s^2}$
a	$\frac{a}{s}$
e^{-at}	$\frac{1}{s+a}$
te^{at}	$\frac{1}{(s-a)^2}$
$(1+at)e^{at}$	$\frac{s}{(s-a)^2}$
$\frac{1}{a}e^{-\frac{t}{a}}$	$\frac{1}{(1+sa)}$
t^p for $p > -1$	$\frac{\Gamma(p+1)}{s^{p+1}}$ for $s > 0$
t^n $n = 1, 2, 3, \ldots$	$\frac{n!}{s^{n+1}}$ $s > 0$
$\frac{1}{a}\left(1 - e^{-at}\right)$	$\frac{1}{s(s+a)}$
$\frac{1}{a}\left(e^{at} - 1\right)$	$\frac{1}{s(s-a)}$
$\frac{1}{a^2}\left(e^{-at} + at - 1\right)$	$\frac{1}{s^2(s+a)}$
$\frac{1}{b-a}\left(e^{-at} - e^{-bt}\right)$	$\frac{1}{(s+a)(s+b)}$ $a \neq b$
$\frac{\left(ae^{at} - be^{bt}\right)}{a-b}$	$\frac{s}{(s-a)(s-b)}$ $a \neq b$
$\frac{\alpha^k t^{k-1} e^{-\alpha t}}{\Gamma(k)}$	$\left(\frac{\alpha}{\alpha+s}\right)^k$

From Table 8.1, the inverse transform is

$$f(t) = \frac{1}{3-1}\left(e^{-t} - e^{-3t}\right) = \frac{1}{2}\left(e^{-t} - e^{-3t}\right). \tag{8.46}$$

Example 8.7 Let X be an exponential random variable with constant failure rate λ, that is, $f(x) = \lambda e^{-\lambda x}$ then we have

$$f^*(s) = \int_0^\infty \lambda e^{-sx} e^{-\lambda x}\mathrm{d}x = \frac{\lambda}{s+\lambda}. \tag{8.47}$$

If X and Y are two independent random variables that represent life times with densities f_1 and f_2, respectively, then the total life time's Z of those two X and Y, says $Z = X + Y$, has a pdf g that can be obtained as follows

$$g(z) = \int_0^z f_1(x) f_2(z-x)\,\mathrm{d}x.$$

The Laplace transform of g in terms of f_1 and f_2 can be written as

$$\begin{aligned} g^*(s) &= \int_0^\infty \mathrm{e}^{-sz} g(z)\mathrm{d}z = \int_0^\infty \int_0^z \mathrm{e}^{-sz} f_1(x) f_2(z-x)\, \mathrm{d}x\mathrm{d}z \\ &= \int_0^\infty \mathrm{e}^{-sx} f_1(x)\mathrm{d}x \int_x^\infty \mathrm{e}^{-s(z-x)} f_2(z-x)\, \mathrm{d}z \\ &= f_1^*(s) f_2^*(s). \end{aligned} \tag{8.48}$$

Example 8.8 If X and Y are both independent having the following pdfs: $f_1(x) = \lambda \mathrm{e}^{-\lambda x}$ and $f_2(y) = \lambda \mathrm{e}^{-\lambda y}$ for $x, y \geq 0$ and $\lambda \geq 0$ then we have

$$g^*(s) = f_1^*(s) f_2^*(s) = \left(\frac{\lambda}{s+\lambda}\right)^2. \tag{8.49}$$

From the Laplace transform Table 8.1, the inverse transform to solve for $g(z)$ is

$$g(z) = \frac{\lambda^2 t\, \mathrm{e}^{-\lambda t}}{\Gamma(2)}$$

which is a special case of gamma pdf.

From Eq. (8.48), one can easily show the Laplace transform of the density function g_n of the total life time S_n of n independent life time's X_i with their pdf f_i for $i = 1,2,\ldots,n$, that

$$g_n^*(s) = f_1^*(s) f_2^*(s) \ldots f_n^*(s) = \prod_{i=1}^{n} f_1^*(s) \tag{8.50}$$

If the pdf of n life time X_1, X_2,...,X_n are independent and identically distributed (i.i.d.) having a constant failure rate λ, then

$$g_n^*(s) = \left(f^*(s)\right)^n = \left(\frac{\lambda}{s+\lambda}\right)^n.$$

From the Laplace transform table, we obtain the inverse transform for the solution function g as follows

$$g_n(z) = \frac{\lambda^n t^{n-1} \mathrm{e}^{-\lambda t}}{\Gamma(n)}. \tag{8.51}$$

References

1. Pham, H.: Software Reliability. Springer, Berlin, Heidelberg (2000)
2. Wang, H., Pham, H.: A quasi renewal process and its applications in imperfect maintenance. Int. J. Syst. Sci. **27**(10), 1055–1062 (1996)
3. Feller, W.: An Introduction to Probability Theory and Its Applications, 3rd edn. Wiley, New York (1994)
4. Pinsky, M., Karlin, S.: Introduction to Stochastic Modeling, 4th edn. Academic Press (2010)
5. Parzen, E.: Stochastic Processes. SIAM (1987)
6. Melsa, J.L., Sage, A.P.: An Introduction to Probability and Stochastic Processes. Dover, Mineola; New York (2013)

Hoang Pham is a Distinguished Professor and former Chairman of the Department of Industrial & Systems Engineering at Rutgers University. He is the author or coauthor of 7 books and has published over 200 journal articles, 100 conference papers, and edited 20 books. His numerous awards include the 2009 IEEE Reliability Society *Engineer of the Year Award.* He is a Fellow of the IEEE and IISE.