The Exponential Distribution and the Poisson Process

5.1 Introduction

In making a mathematical model for a real-world phenomenon it is always necessary to make certain simplifying assumptions so as to render the mathematics tractable. On the other hand, however, we cannot make too many simplifying assumptions, for then our conclusions, obtained from the mathematical model, would not be applicable to the real-world situation. Thus, in short, we must make enough simplifying assumptions to enable us to handle the mathematics but not so many that the mathematical model no longer resembles the real-world phenomenon. One simplifying assumption that is often made is to assume that certain random variables are exponentially distributed. The reason for this is that the exponential distribution is both relatively easy to work with and is often a good approximation to the actual distribution.

The property of the exponential distribution that makes it easy to analyze is that it does not deteriorate with time. By this we mean that if the lifetime of an item is exponentially distributed, then an item that has been in use for ten (or any number of) hours is as good as a new item in regards to the amount of time remaining until the item fails. This will be formally defined in Section 5.2 where it will be shown that the exponential is the only distribution that possesses this property.

In Section 5.3 we shall study counting processes with an emphasis on a kind of counting process known as the Poisson process. Among other things we

Introduction to Probability Models, ISBN: 9780123756862

shall discover about this process is its intimate connection with the exponential distribution.

5.2 The Exponential Distribution

5.2.1 Definition

A continuous random variable X is said to have an *exponential distribution* with parameter λ, $\lambda > 0$, if its probability density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0 \\ 0, & x < 0 \end{cases}$$

or, equivalently, if its cdf is given by

$$F(x) = \int_{-\infty}^{x} f(y)\,dy = \begin{cases} 1 - e^{-\lambda x}, & x \geqslant 0 \\ 0, & x < 0 \end{cases}$$

The mean of the exponential distribution, $E[X]$, is given by

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f(x)\,dx \\ &= \int_{0}^{\infty} \lambda x e^{-\lambda x}\,dx \end{aligned}$$

Integrating by parts ($u = x, dv = \lambda e^{-\lambda x} dx$) yields

$$E[X] = -x e^{-\lambda x}\Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x}\,dx = \frac{1}{\lambda}$$

The moment generating function $\phi(t)$ of the exponential distribution is given by

$$\begin{aligned} \phi(t) &= E[e^{tX}] \\ &= \int_0^{\infty} e^{tx} \lambda e^{-\lambda x}\,dx \\ &= \frac{\lambda}{\lambda - t} \qquad \text{for } t < \lambda \end{aligned} \tag{5.1}$$

All the moments of X can now be obtained by differentiating Equation (5.1). For example,

$$\begin{aligned} E[X^2] &= \frac{d^2}{dt^2}\phi(t)\Big|_{t=0} \\ &= \frac{2\lambda}{(\lambda - t)^3}\Big|_{t=0} \\ &= \frac{2}{\lambda^2} \end{aligned}$$

Consequently,

$$\begin{aligned}\mathrm{Var}(X) &= E[X^2] - (E[X])^2 \\ &= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} \\ &= \frac{1}{\lambda^2}\end{aligned}$$

Example 5.1 (Exponential Random Variables and Expected Discounted Returns) Suppose that you are receiving rewards at randomly changing rates continuously throughout time. Let $R(x)$ denote the random rate at which you are receiving rewards at time x. For a value $\alpha \geqslant 0$, called the discount rate, the quantity

$$R = \int_0^\infty e^{-\alpha x} R(x)\, dx$$

represents the total discounted reward. (In certain applications, α is a continuously compounded interest rate, and R is the present value of the infinite flow of rewards.) Whereas

$$E[R] = E\left[\int_0^\infty e^{-\alpha x} R(x)\, dx\right] = \int_0^\infty e^{-\alpha x} E[R(x)]\, dx$$

is the expected total discounted reward, we will show that it is also equal to the expected total reward earned up to an exponentially distributed random time with rate α.

Let T be an exponential random variable with rate α that is independent of all the random variables $R(x)$. We want to argue that

$$\int_0^\infty e^{-\alpha x} E[R(x)]\, dx = E\left[\int_0^T R(x)\, dx\right]$$

To show this define for each $x \geqslant 0$ a random variable $I(x)$ by

$$I(x) = \begin{cases} 1, & \text{if } x \leqslant T \\ 0, & \text{if } x > T \end{cases}$$

and note that

$$\int_0^T R(x)\, dx = \int_0^\infty R(x) I(x)\, dx$$

Thus,

$$\begin{aligned}
E\left[\int_0^T R(x)\,dx\right] &= E\left[\int_0^\infty R(x)I(x)\,dx\right] \\
&= \int_0^\infty E[R(x)I(x)]\,dx \\
&= \int_0^\infty E[R(x)]E[I(x)]\,dx \qquad \text{by independence} \\
&= \int_0^\infty E[R(x)]P\{T \geqslant x\}\,dx \\
&= \int_0^\infty e^{-\alpha x}E[R(x)]\,dx
\end{aligned}$$

Therefore, the expected total discounted reward is equal to the expected total (undiscounted) reward earned by a random time that is exponentially distributed with a rate equal to the discount factor. ■

5.2.2 Properties of the Exponential Distribution

A random variable X is said to be without memory, or *memoryless*, if

$$P\{X > s + t \mid X > t\} = P\{X > s\} \qquad \text{for all } s, t \geqslant 0 \tag{5.2}$$

If we think of X as being the lifetime of some instrument, then Equation (5.2) states that the probability that the instrument lives for at least $s + t$ hours given that it has survived t hours is the same as the initial probability that it lives for at least s hours. In other words, if the instrument is alive at time t, then the distribution of the remaining amount of time that it survives is the same as the original lifetime distribution; that is, the instrument does not remember that it has already been in use for a time t.

The condition in Equation (5.2) is equivalent to

$$\frac{P\{X > s + t, X > t\}}{P\{X > t\}} = P\{X > s\}$$

or

$$P\{X > s + t\} = P\{X > s\}P\{X > t\} \tag{5.3}$$

Since Equation (5.3) is satisfied when X is exponentially distributed (for $e^{-\lambda(s+t)} = e^{-\lambda s}e^{-\lambda t}$), it follows that exponentially distributed random variables are memoryless.

Example 5.2 Suppose that the amount of time one spends in a bank is exponentially distributed with mean ten minutes, that is, $\lambda = \frac{1}{10}$. What is the probability that a customer will spend more than fifteen minutes in the bank? What is the probability that a customer will spend more than fifteen minutes in the bank given that she is still in the bank after ten minutes?

Solution: If X represents the amount of time that the customer spends in the bank, then the first probability is just

$$P\{X > 15\} = e^{-15\lambda} = e^{-3/2} \approx 0.220$$

The second question asks for the probability that a customer who has spent ten minutes in the bank will have to spend at least five more minutes. However, since the exponential distribution does not "remember" that the customer has already spent ten minutes in the bank, this must equal the probability that an entering customer spends at least five minutes in the bank. That is, the desired probability is just

$$P\{X > 5\} = e^{-5\lambda} = e^{-1/2} \approx 0.604$$ ■

Example 5.3 Consider a post office that is run by two clerks. Suppose that when Mr. Smith enters the system he discovers that Mr. Jones is being served by one of the clerks and Mr. Brown by the other. Suppose also that Mr. Smith is told that his service will begin as soon as either Jones or Brown leaves. If the amount of time that a clerk spends with a customer is exponentially distributed with mean $1/\lambda$, what is the probability that, of the three customers, Mr. Smith is the last to leave the post office?

Solution: The answer is obtained by this reasoning: Consider the time at which Mr. Smith first finds a free clerk. At this point either Mr. Jones or Mr. Brown would have just left and the other one would still be in service. However, by the lack of memory of the exponential, it follows that the amount of time that this other man (either Jones or Brown) would still have to spend in the post office is exponentially distributed with mean $1/\lambda$. That is, it is the same as if he were just starting his service at this point. Hence, by symmetry, the probability that he finishes before Smith must equal $\frac{1}{2}$. ■

Example 5.4 The dollar amount of damage involved in an automobile accident is an exponential random variable with mean 1000. Of this, the insurance company only pays that amount exceeding (the deductible amount of) 400. Find the expected value and the standard deviation of the amount the insurance company pays per accident.

Solution: If X is the dollar amount of damage resulting from an accident, then the amount paid by the insurance company is $(X - 400)^+$, (where a^+ is defined to equal a if $a > 0$ and to equal 0 if $a \leqslant 0$). Whereas we could certainly determine the expected value and variance of $(X-400)^+$ from first principles, it is easier to condition on whether X exceeds 400. So, let

$$I = \begin{cases} 1, & \text{if } X > 400 \\ 0, & \text{if } X \leqslant 400 \end{cases}$$

Let $Y=(X-400)^+$ be the amount paid. By the lack of memory property of the exponential, it follows that if a damage amount exceeds 400, then the amount by which it exceeds it is exponential with mean 1000. Therefore,

$$\begin{aligned} E[Y|I=1] &= 1000 \\ E[Y|I=0] &= 0 \\ \mathrm{Var}(Y|I=1) &= (1000)^2 \\ \mathrm{Var}(Y|I=0) &= 0 \end{aligned}$$

which can be conveniently written as

$$E[Y|I] = 10^3 I, \qquad \mathrm{Var}(Y|I) = 10^6 I$$

Because I is a Bernoulli random variable that is equal to 1 with probability $e^{-0.4}$, we obtain

$$E[Y] = E\big[E[Y|I]\big] = 10^3 E[I] = 10^3 e^{-0.4} \approx 670.32$$

and, by the conditional variance formula

$$\begin{aligned} \mathrm{Var}(Y) &= E\big[\mathrm{Var}(Y|I)\big] + \mathrm{Var}\big(E[Y|I]\big) \\ &= 10^6 e^{-0.4} + 10^6 e^{-0.4}(1 - e^{-0.4}) \end{aligned}$$

where the final equality used that the variance of a Bernoulli random variable with parameter p is $p(1-p)$. Consequently,

$$\sqrt{\mathrm{Var}(Y)} \approx 944.09 \qquad \blacksquare$$

It turns out that not only is the exponential distribution "memoryless," but it is the unique distribution possessing this property. To see this, suppose that X is memoryless and let $\bar{F}(x) = P\{X > x\}$. Then by Equation (5.3) it follows that

$$\bar{F}(s+t) = \bar{F}(s)\bar{F}(t)$$

That is, $\bar{F}(x)$ satisfies the functional equation

$$g(s+t) = g(s)g(t)$$

However, it turns out that the only right continuous solution of this functional equation is

$$g(x) = e^{-\lambda x}*$$

and since a distribution function is always right continuous we must have

$$\bar{F}(x) = e^{-\lambda x}$$

or

$$F(x) = P\{X \leqslant x\} = 1 - e^{-\lambda x}$$

which shows that X is exponentially distributed.

Example 5.5 A store must decide how much of a certain commodity to order so as to meet next month's demand, where that demand is assumed to have an exponential distribution with rate λ. If the commodity costs the store c per pound, and can be sold at a price of $s > c$ per pound, how much should be ordered so as to maximize the store's expected profit? Assume that any inventory left over at the end of the month is worthless and that there is no penalty if the store cannot meet all the demand.

Solution: Let X equal the demand. If the store orders the amount t, then the profit, call it P, is given by

$$P = s \min(X, t) - ct$$

Writing

$$\min(X, t) = X - (X - t)^+$$

* This is proven as follows: If $g(s + t) = g(s)g(t)$, then

$$g\left(\frac{2}{n}\right) = g\left(\frac{1}{n} + \frac{1}{n}\right) = g^2\left(\frac{1}{n}\right)$$

and repeating this yields $g(m/n) = g^m(1/n)$. Also,

$$g(1) = g\left(\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}\right) = g^n\left(\frac{1}{n}\right) \qquad \text{or} \qquad g\left(\frac{1}{n}\right) = (g(1))^{1/n}$$

Hence $g(m/n) = (g(1))^{m/n}$, which implies, since g is right continuous, that $g(x) = (g(1))^x$. Since $g(1) = (g(\frac{1}{2}))^2 \geqslant 0$ we obtain $g(x) = e^{-\lambda x}$, where $\lambda = -\log(g(1))$.

we obtain, upon conditioning whether $X > t$ and then using the lack of memory property of the exponential, that

$$\begin{aligned} E[(X-t)^+] &= E[(X-t)^+|X > t]P(X > t) + E[(X-t)^+|X \leqslant t]P(X \leqslant t) \\ &= E[(X-t)^+|X > t]e^{-\lambda t} \\ &= \frac{1}{\lambda}e^{-\lambda t} \end{aligned}$$

where the final equality used the lack of memory property of exponential random variables to conclude that, conditional on X exceeding t, the amount by which it exceeds it is an exponential random variable with rate λ. Hence,

$$E[\min(X,t)] = \frac{1}{\lambda} - \frac{1}{\lambda}e^{-\lambda t}$$

giving that

$$E[P] = \frac{s}{\lambda} - \frac{s}{\lambda}e^{-\lambda t} - ct$$

Differentiation now yields that the maximal profit is attained when $se^{-\lambda t} - c = 0$; that is, when

$$t = \frac{1}{\lambda}\log(s/c)$$

Now, suppose that all unsold inventory can be returned for the amount $r < \min(s,c)$ per pound; and also that there is a penalty cost p per pound of unmet demand. In this case, using our previously derived expression for $E[P]$, we have

$$E[P] = \frac{s}{\lambda} - \frac{s}{\lambda}e^{-\lambda t} - ct + rE[(t-X)^+] - pE[(X-t)^+]$$

Using that

$$\min(X,t) = t - (t-X)^+$$

we see that

$$E[(t-X)^+] = t - E[\min(X,t)] = t - \frac{1}{\lambda} + \frac{1}{\lambda}e^{-\lambda t}$$

Hence,

$$\begin{aligned} E[P] &= \frac{s}{\lambda} - \frac{s}{\lambda}e^{-\lambda t} - ct + rt - \frac{r}{\lambda} + \frac{r}{\lambda}e^{-\lambda t} - \frac{p}{\lambda}e^{-\lambda t} \\ &= \frac{s-r}{\lambda} + \frac{r-s-p}{\lambda}e^{-\lambda t} - (c-r)t \end{aligned}$$

Calculus now yields that the optimal amount to order is

$$t = \frac{1}{\lambda} \log \left(\frac{s + p - r}{c - r} \right)$$

It is worth noting that the optimal amount to order increases in s, p, and r and decreases in λ and c. (Are these monotonicity properties intuitive?) ■

The memoryless property is further illustrated by the failure rate function (also called the hazard rate function) of the exponential distribution.

Consider a continuous positive random variable X having distribution function F and density f. The *failure* (or *hazard*) *rate* function $r(t)$ is defined by

$$r(t) = \frac{f(t)}{1 - F(t)} \tag{5.4}$$

To interpret $r(t)$, suppose that an item, having lifetime X, has survived for t hours, and we desire the probability that it does not survive for an additional time dt. That is, consider $P\{X \in (t, t + dt)|X > t\}$. Now,

$$\begin{aligned} P\{X \in (t, t + dt)|X > t\} &= \frac{P\{X \in (t, t + dt), X > t\}}{P\{X > t\}} \\ &= \frac{P\{X \in (t, t + dt)\}}{P\{X > t\}} \\ &\approx \frac{f(t)\, dt}{1 - F(t)} = r(t)\, dt \end{aligned}$$

That is, $r(t)$ represents the conditional probability density that a t-year-old item will fail.

Suppose now that the lifetime distribution is exponential. Then, by the memoryless property, it follows that the distribution of remaining life for a t-year-old item is the same as for a new item. Hence, $r(t)$ should be constant. This checks out since

$$\begin{aligned} r(t) &= \frac{f(t)}{1 - F(t)} \\ &= \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda \end{aligned}$$

Thus, the failure rate function for the exponential distribution is constant. The parameter λ is often referred to as the *rate* of the distribution. (Note that the rate is the reciprocal of the mean, and vice versa.)

It turns out that the failure rate function $r(t)$ uniquely determines the distribution F. To prove this, we note by Equation (5.4) that

$$r(t) = \frac{\frac{d}{dt} F(t)}{1 - F(t)}$$

Integrating both sides yields

$$\log(1 - F(t)) = -\int_0^t r(t)\,dt + k$$

or

$$1 - F(t) = e^k \exp\left\{-\int_0^t r(t)\,dt\right\}$$

Letting $t = 0$ shows that $k = 0$ and thus

$$F(t) = 1 - \exp\left\{-\int_0^t r(t)\,dt\right\}$$

The preceding identity can also be used to show that exponential random variables are the only ones that are memoryless. Because if X is memoryless, then its failure rate function must be constant. But if $r(t) = c$, then by the preceding equation

$$1 - F(t) = e^{-ct}$$

showing that the random variable is exponential.

Example 5.6 Let $X_1, \ldots, X_n$ be independent exponential random variables with respective rates $\lambda_1, \ldots, \lambda_n$, where $\lambda_i \neq \lambda_j$ when $i \neq j$. Let T be independent of these random variables and suppose that

$$\sum_{j=1}^{n} P_j = 1 \qquad \text{where } P_j = P\{T = j\}$$

The random variable X_T is said to be a *hyperexponential* random variable. To see how such a random variable might originate, imagine that a bin contains n different types of batteries, with a type j battery lasting for an exponential distributed time with rate $\lambda_j, j = 1, \ldots, n$. Suppose further that P_j is the proportion of batteries in the bin that are type j for each $j = 1, \ldots, n$. If a battery is randomly chosen, in the sense that it is equally likely to be any of the batteries in the bin, then the lifetime of the battery selected will have the hyperexponential distribution specified in the preceding.

To obtain the distribution function F of $X = X_T$, condition on T. This yields

$$\begin{aligned} 1 - F(t) &= P\{X > t\} \\ &= \sum_{i=1}^{n} P\{X > t | T = i\} P\{T = i\} \\ &= \sum_{i=1}^{n} P_i e^{-\lambda_i t} \end{aligned}$$

Differentiation of the preceding yields f, the density function of X.

$$f(t) = \sum_{i=1}^{n} \lambda_i P_i e^{-\lambda_i t}$$

Consequently, the failure rate function of a hyperexponential random variable is

$$r(t) = \frac{\sum_{j=1}^{n} P_j \lambda_j e^{-\lambda_j t}}{\sum_{i=1}^{n} P_i e^{-\lambda_i t}}$$

By noting that

$$\begin{aligned} P\{T = j | X > t\} &= \frac{P\{X > t | T = j\} P\{T = j\}}{P\{X > t\}} \\ &= \frac{P_j e^{-\lambda_j t}}{\sum_{i=1}^{n} P_i e^{-\lambda_i t}} \end{aligned}$$

we see that the failure rate function $r(t)$ can also be written as

$$r(t) = \sum_{j=1}^{n} \lambda_j P\{T = j | X > t\}$$

If $\lambda_1 < \lambda_i$, for all $i > 1$, then

$$\begin{aligned} P\{T = 1 | X > t\} &= \frac{P_1 e^{-\lambda_1 t}}{P_1 e^{-\lambda_1 t} + \sum_{i=2}^{n} P_i e^{-\lambda_i t}} \\ &= \frac{P_1}{P_1 + \sum_{i=2}^{n} P_i e^{-(\lambda_i - \lambda_1) t}} \\ &\to 1 \qquad \text{as } t \to \infty \end{aligned}$$

Similarly, $P\{T = i | X > t\} \to 0$ when $i \neq 1$, thus showing that

$$\lim_{t \to \infty} r(t) = \min_i \lambda_i$$

That is, as a randomly chosen battery ages its failure rate converges to the failure rate of the exponential type having the smallest failure rate, which is intuitive since the longer the battery lasts, the more likely it is a battery type with the smallest failure rate. ■

5.2.3 Further Properties of the Exponential Distribution

Let $X_1, \ldots, X_n$ be independent and identically distributed exponential random variables having mean $1/\lambda$. It follows from the results of Example 2.39 that $X_1 + \cdots + X_n$ has a gamma distribution with parameters n and λ. Let us now give a second verification of this result by using mathematical induction. Because there is nothing to prove when $n = 1$, let us start by assuming that $X_1 + \cdots + X_{n-1}$ has density given by

$$f_{X_1+\cdots+X_{n-1}}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-2}}{(n-2)!}$$

Hence,

$$\begin{aligned}
f_{X_1+\cdots+X_{n-1}+X_n}(t) &= \int_0^\infty f_{X_n}(t-s) f_{X_1+\cdots+X_{n-1}}(s)\, ds \\
&= \int_0^t \lambda e^{-\lambda(t-s)} \lambda e^{-\lambda s} \frac{(\lambda s)^{n-2}}{(n-2)!}\, ds \\
&= \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}
\end{aligned}$$

which proves the result.

Another useful calculation is to determine the probability that one exponential random variable is smaller than another. That is, suppose that X_1 and X_2 are independent exponential random variables with respective means $1/\lambda_1$ and $1/\lambda_2$; what is $P\{X_1 < X_2\}$? This probability is easily calculated by conditioning on X_1:

$$\begin{aligned}
P\{X_1 < X_2\} &= \int_0^\infty P\{X_1 < X_2 | X_1 = x\} \lambda_1 e^{-\lambda_1 x}\, dx \\
&= \int_0^\infty P\{x < X_2\} \lambda_1 e^{-\lambda_1 x}\, dx \\
&= \int_0^\infty e^{-\lambda_2 x} \lambda_1 e^{-\lambda_1 x}\, dx \\
&= \int_0^\infty \lambda_1 e^{-(\lambda_1+\lambda_2)x}\, dx \\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2}
\end{aligned} \tag{5.5}$$

Suppose that $X_1, X_2, \ldots, X_n$ are independent exponential random variables, with X_i having rate $\mu_i, i = 1, \ldots, n$. It turns out that the smallest of the X_i is exponential with a rate equal to the sum of the μ_i. This is shown as follows:

$$\begin{aligned}
P\{\text{minimum}(X_1, \ldots, X_n) > x\} &= P\{X_i > x \text{ for each } i = 1, \ldots, n\} \\
&= \prod_{i=1}^n P\{X_i > x\} \qquad \text{(by independence)} \\
&= \prod_{i=1}^n e^{-\mu_i x} \\
&= \exp\left\{-\left(\sum_{i=1}^n \mu_i\right) x\right\}
\end{aligned} \tag{5.6}$$

Example 5.7 (Analyzing Greedy Algorithms for the Assignment Problem) A group of n people is to be assigned to a set of n jobs, with one person assigned to each job. For a given set of n^2 values C_{ij}, $i, j = 1, \ldots, n$, a cost C_{ij} is incurred when person i is assigned to job j. The classical assignment problem is to determine the set of assignments that minimizes the sum of the n costs incurred.

Rather than trying to determine the optimal assignment, let us consider two heuristic algorithms for solving this problem. The first heuristic is as follows. Assign person 1 to the job that results in the least cost. That is, person 1 is assigned to job j_1 where $C(1, j_1) = \text{minimum}_j\, C(1, j)$. Now eliminate that job from consideration and assign person 2 to the job that results in the least cost. That is, person 2 is assigned to job j_2 where $C(2, j_2) = \text{minimum}_{j \neq j_1} C(2, j)$. This procedure is then continued until all n persons are assigned. Since this procedure always selects the best job for the person under consideration, we will call it Greedy Algorithm A.

The second algorithm, which we call Greedy Algorithm B, is a more "global" version of the first greedy algorithm. It considers all n^2 cost values and chooses the pair i_1, j_1 for which $C(i, j)$ is minimal. It then assigns person i_1 to job j_1. It then eliminates all cost values involving either person i_1 or job j_1 (so that $(n-1)^2$ values remain) and continues in the same fashion. That is, at each stage it chooses the person and job that have the smallest cost among all the unassigned people and jobs.

Under the assumption that the C_{ij} constitute a set of n^2 independent exponential random variables each having mean 1, which of the two algorithms results in a smaller expected total cost?

Solution: Suppose first that Greedy Algorithm A is employed. Let C_i denote the cost associated with person $i, i = 1, \ldots, n$. Now C_1 is the minimum of n independent exponentials each having rate 1; so by Equation (5.6) it will be exponential with rate n. Similarly, C_2 is the minimum of $n-1$ independent exponentials with rate 1, and so is exponential with rate $n-1$. Indeed, by the same reasoning C_i will be exponential with rate $n-i+1$, $i = 1, \ldots, n$. Thus, the expected total cost under Greedy Algorithm A is

$$\begin{aligned} E_A[\text{total cost}] &= E[C_1 + \cdots + C_n] \\ &= \sum_{i=1}^{n} 1/i \end{aligned}$$

Let us now analyze Greedy Algorithm B. Let C_i be the cost of the ith person-job pair assigned by this algorithm. Since C_1 is the minimum of all the n^2 values C_{ij}, it follows from Equation (5.6) that C_1 is exponential with rate n^2. Now, it follows from the lack of memory property of the exponential that the amounts by which the other C_{ij} exceed C_1 will be independent exponentials with rates 1. As a result, C_2 is equal to C_1 plus the minimum of $(n-1)^2$ independent exponentials with rate 1. Similarly, C_3 is equal to C_2 plus the minimum of $(n-2)^2$ independent exponentials with rate 1, and so on. Therefore, we see that

$$\begin{aligned}
E[C_1] &= 1/n^2,\\
E[C_2] &= E[C_1] + 1/(n-1)^2,\\
E[C_3] &= E[C_2] + 1/(n-2)^2,\\
&\vdots\\
E[C_j] &= E[C_{j-1}] + 1/(n-j+1)^2,\\
&\vdots\\
E[C_n] &= E[C_{n-1}] + 1
\end{aligned}$$

Therefore,

$$\begin{aligned}
E[C_1] &= 1/n^2,\\
E[C_2] &= 1/n^2 + 1/(n-1)^2,\\
E[C_3] &= 1/n^2 + 1/(n-1)^2 + 1/(n-2)^2,\\
&\vdots\\
E[C_n] &= 1/n^2 + 1/(n-1)^2 + 1/(n-2)^2 + \cdots + 1
\end{aligned}$$

Adding up all the $E[C_i]$ yields

$$\begin{aligned}
E_B[\text{total cost}] &= n/n^2 + (n-1)/(n-1)^2 + (n-2)/(n-2)^2 + \cdots + 1\\
&= \sum_{i=1}^{n} \frac{1}{i}
\end{aligned}$$

The expected cost is thus the same for both greedy algorithms. ■

Let $X_1, \ldots, X_n$ be independent exponential random variables, with respective rates $\lambda_1, \ldots, \lambda_n$. A useful result, generalizing Equation (5.5), is that X_i is the smallest of these with probability $\lambda_i / \sum_j \lambda_j$. This is shown as follows:

$$\begin{aligned}
P\left\{X_i = \min_j X_j\right\} &= P\left\{X_i < \min_{j \neq i} X_j\right\}\\
&= \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j}
\end{aligned}$$

where the final equality uses Equation (5.5) along with the fact that $\min_{j \neq i} X_j$ is exponential with rate $\sum_{j \neq i} \lambda_j$.

Another important fact is that $\min_i X_i$ and the rank ordering of the X_i are independent. To see why this is true, consider the conditional probability that $X_{i_1} < X_{i_2} < \cdots < X_{i_n}$ given that the minimal value is greater than t. Because $\min_i X_i > t$ means that all the X_i are greater than t, it follows from the lack of memory property of exponential random variables that their remaining lives beyond t remain independent exponential random variables with their original rates. Consequently,

$$\begin{aligned} P\left\{X_{i_1} < \cdots < X_{i_n} \middle| \min_i X_i > t\right\} &= P\left\{X_{i_1} - t < \cdots < X_{i_n} - t \middle| \min_i X_i > t\right\} \\ &= P\{X_{i_1} < \cdots < X_{i_n}\} \end{aligned}$$

which proves the result.

Example 5.8 Suppose you arrive at a post office having two clerks at a moment when both are busy but there is no one else waiting in line. You will enter service when either clerk becomes free. If service times for clerk i are exponential with rate $\lambda_i, i = 1, 2$, find $E[T]$, where T is the amount of time that you spend in the post office.

Solution: Let R_i denote the remaining service time of the customer with clerk i, $i = 1, 2$, and note, by the lack of memory property of exponentials, that R_1 and R_2 are independent exponential random variables with respective rates λ_1 and λ_2. Conditioning on which of R_1 or R_2 is the smallest yields

$$\begin{aligned} E[T] &= E[T|R_1 < R_2]P\{R_1 < R_2\} + E[T|R_2 \leqslant R_1]P\{R_2 \leqslant R_1\} \\ &= E[T|R_1 < R_2]\frac{\lambda_1}{\lambda_1 + \lambda_2} + E[T|R_2 \leqslant R_1]\frac{\lambda_2}{\lambda_1 + \lambda_2} \end{aligned}$$

Now, with S denoting your service time

$$\begin{aligned} E[T|R_1 < R_2] &= E[R_1 + S|R_1 < R_2] \\ &= E[R_1|R_1 < R_2] + E[S|R_1 < R_2] \\ &= E[R_1|R_1 < R_2] + \frac{1}{\lambda_1} \\ &= \frac{1}{\lambda_1 + \lambda_2} + \frac{1}{\lambda_1} \end{aligned}$$

The final equation used that conditional on $R_1 < R_2$ the random variable R_1 is the minimum of R_1 and R_2 and is thus exponential with rate $\lambda_1 + \lambda_2$; and also that conditional on $R_1 < R_2$ you are served by server 1.

As we can show in a similar fashion that

$$E[T|R_2 \leqslant R_1] = \frac{1}{\lambda_1 + \lambda_2} + \frac{1}{\lambda_2}$$

we obtain the result

$$E[T] = \frac{3}{\lambda_1 + \lambda_2}$$

Another way to obtain $E[T]$ is to write T as a sum, take expectations, and then condition where needed. This approach yields

$$\begin{aligned}E[T] &= E[\min(R_1, R_2) + S] \\ &= E[\min(R_1, R_2)] + E[S] \\ &= \frac{1}{\lambda_1 + \lambda_2} + E[S]\end{aligned}$$

To compute $E[S]$, we condition on which of R_1 and R_2 is smallest.

$$\begin{aligned}E[S] &= E[S|R_1 < R_2]\frac{\lambda_1}{\lambda_1 + \lambda_2} + E[S|R_2 \leqslant R_1]\frac{\lambda_2}{\lambda_1 + \lambda_2} \\ &= \frac{2}{\lambda_1 + \lambda_2}\end{aligned}$$ ■

Example 5.9 There are n cells in the body, of which cells $1, \ldots, k$ are target cells. Associated with each cell is a weight, with w_i being the weight associated with cell i, $i = 1, \ldots, n$. The cells are destroyed one at a time in a random order, which is such that if S is the current set of surviving cells then, independent of the order in which the cells not in S have been destroyed, the next cell killed is i, $i \in S$, with probability $w_i / \sum_{j \in S} w_j$. In other words, the probability that a given surviving cell is the next one to be killed is the weight of that cell divided by the sum of the weights of all still surviving cells. Let A denote the total number of cells that are still alive at the moment when all the cells $1, 2, \ldots, k$ have been killed, and find $E[A]$.

Solution: Although it would be quite difficult to solve this problem by a direct combinatorial argument, a nice solution can be obtained by relating the order in which cells are killed to a ranking of independent exponential random variables. To do so, let $X_1, \ldots, X_n$ be independent exponential random variables, with X_i having rate w_i, $i = 1, \ldots, n$. Note that X_i will be the smallest of these exponentials with probability $w_i / \sum_j w_j$; further, given that X_i is the smallest, X_r will be the next smallest with probability $w_r / \sum_{j \neq i} w_j$; further, given that X_i and X_r are, respectively, the first and second smallest, X_s, $s \neq i, r$, will be the third smallest with probability $w_s / \sum_{j \neq i,r} w_j$; and so on. Consequently, if we let I_j be the index of the jth smallest of $X_1, \ldots, X_n$—so that $X_{I_1} < X_{I_2} < \cdots < X_{I_n}$—then the order in which the cells are destroyed has the same distribution as $I_1, \ldots, I_n$. So, let us suppose that the order in which the cells are killed is determined by the ordering of $X_1, \ldots, X_n$. (Equivalently, we can suppose that all cells will eventually be killed, with cell i being killed at time X_i, $i = 1, \ldots, n$.)

If we let A_j equal 1 if cell j is still alive at the moment when all the cells $1, \ldots, k$ have been killed, and let it equal 0 otherwise, then

$$A = \sum_{j=k+1}^{n} A_j$$

Because cell j will be alive at the moment when all the cells $1, \ldots, k$ have been killed if X_j is larger than all the values $X_1, \ldots, X_k$, we see that for $j > k$

$$\begin{aligned}
E[A_j] &= P\{A_j = 1\} \\
&= P\{X_j > \max_{i=1,\ldots,k} X_i\} \\
&= \int_0^\infty P\Big\{X_j > \max_{i=1,\ldots,k} X_i | X_j = x\Big\} w_j e^{-w_j x}\, dx \\
&= \int_0^\infty P\{X_i < x \text{ for all } i = 1, \ldots, k\} w_j e^{-w_j x}\, dx \\
&= \int_0^\infty \prod_{i=1}^k (1 - e^{-w_i x})\, w_j e^{-w_j x}\, dx \\
&= \int_0^1 \prod_{i=1}^k (1 - y^{w_i/w_j})\, dy
\end{aligned}$$

where the final equality follows from the substitution $y = e^{-w_j x}$. Thus, we obtain the result

$$E[A] = \sum_{j=k+1}^n \int_0^1 \prod_{i=1}^k (1 - y^{w_i/w_j})\, dy = \int_0^1 \sum_{j=k+1}^n \prod_{i=1}^k (1 - y^{w_i/w_j})\, dy \qquad \blacksquare$$

Example 5.10 Suppose that customers are in line to receive service that is provided sequentially by a server; whenever a service is completed, the next person in line enters the service facility. However, each waiting customer will only wait an exponentially distributed time with rate θ; if its service has not yet begun by this time then it will immediately depart the system. These exponential times, one for each waiting customer, are independent. In addition, the service times are independent exponential random variables with rate μ. Suppose that someone is presently being served and consider the person who is nth in line.

(a) Find P_n, the probability that this customer is eventually served.
(b) Find W_n, the conditional expected amount of time this person spends waiting in line given that she is eventually served.

Solution: Consider the $n + 1$ random variables consisting of the remaining service time of the person in service along with the n additional exponential departure times with rate θ of the first n in line.

(a) Given that the smallest of these $n + 1$ independent exponentials is the departure time of the nth person in line, the conditional probability that this person will be served is 0; on the other hand, given that this person's departure time is not the smallest, the conditional probability that this person will be served is the same as if it were initially in position $n - 1$. Since the probability that a given departure time is the smallest of the $n + 1$ exponentials is

$\theta/(n\theta + \mu)$, we obtain

$$P_n = \frac{(n-1)\theta + \mu}{n\theta + \mu} P_{n-1}$$

Using the preceding with $n - 1$ replacing n gives

$$P_n = \frac{(n-1)\theta + \mu}{n\theta + \mu} \frac{(n-2)\theta + \mu}{(n-1)\theta + \mu} P_{n-2} = \frac{(n-2)\theta + \mu}{n\theta + \mu} P_{n-2}$$

Continuing in this fashion yields the result

$$P_n = \frac{\theta + \mu}{n\theta + \mu} P_1 = \frac{\mu}{n\theta + \mu}$$

(b) To determine an expression for W_n, we use the fact that the minimum of independent exponentials is, independent of their rank ordering, exponential with a rate equal to the sum of the rates. Since the time until the nth person in line enters service is the minimum of these $n + 1$ random variables plus the additional time thereafter, we see, upon using the lack of memory property of exponential random variables, that

$$W_n = \frac{1}{n\theta + \mu} + W_{n-1}$$

Repeating the preceding argument with successively smaller values of n yields the solution

$$W_n = \sum_{i=1}^{n} \frac{1}{i\theta + \mu} \qquad \blacksquare$$

5.2.4 Convolutions of Exponential Random Variables

Let $X_i, i = 1, \ldots, n$, be independent exponential random variables with respective rates $\lambda_i, i = 1, \ldots, n$, and suppose that $\lambda_i \neq \lambda_j$ for $i \neq j$. The random variable $\sum_{i=1}^{n} X_i$ is said to be a *hypoexponential* random variable. To compute its probability density function, let us start with the case $n = 2$. Now,

$$\begin{aligned}
f_{X_1+X_2}(t) &= \int_0^t f_{X_1}(s) f_{X_2}(t-s)\, ds \\
&= \int_0^t \lambda_1 e^{-\lambda_1 s} \lambda_2 e^{-\lambda_2 (t-s)}\, ds \\
&= \lambda_1 \lambda_2 e^{-\lambda_2 t} \int_0^t e^{-(\lambda_1 - \lambda_2)s}\, ds \\
&= \frac{\lambda_1}{\lambda_1 - \lambda_2} \lambda_2 e^{-\lambda_2 t} (1 - e^{-(\lambda_1 - \lambda_2)t}) \\
&= \frac{\lambda_1}{\lambda_1 - \lambda_2} \lambda_2 e^{-\lambda_2 t} + \frac{\lambda_2}{\lambda_2 - \lambda_1} \lambda_1 e^{-\lambda_1 t}
\end{aligned}$$

Using the preceding, a similar computation yields, when $n = 3$,

$$f_{X_1+X_2+X_3}(t) = \sum_{i=1}^{3} \lambda_i e^{-\lambda_i t}\left(\prod_{j\neq i} \frac{\lambda_j}{\lambda_j - \lambda_i}\right)$$

which suggests the general result

$$f_{X_1+\cdots+X_n}(t) = \sum_{i=1}^{n} C_{i,n}\lambda_i e^{-\lambda_i t}$$

where

$$C_{i,n} = \prod_{j\neq i} \frac{\lambda_j}{\lambda_j - \lambda_i}$$

We will now prove the preceding formula by induction on n. Since we have already established it for $n = 2$, assume it for n and consider $n + 1$ arbitrary independent exponentials X_i with distinct rates λ_i, $i = 1, \ldots, n + 1$. If necessary, renumber X_1 and X_{n+1} so that $\lambda_{n+1} < \lambda_1$. Now,

$$\begin{aligned}
f_{X_1+\cdots+X_{n+1}}(t) &= \int_0^t f_{X_1+\cdots+X_n}(s)\lambda_{n+1}e^{-\lambda_{n+1}(t-s)}\,ds \\
&= \sum_{i=1}^{n} C_{i,n}\int_0^t \lambda_i e^{-\lambda_i s}\lambda_{n+1}e^{-\lambda_{n+1}(t-s)}\,ds \\
&= \sum_{i=1}^{n} C_{i,n}\left(\frac{\lambda_i}{\lambda_i - \lambda_{n+1}}\lambda_{n+1}e^{-\lambda_{n+1}t} + \frac{\lambda_{n+1}}{\lambda_{n+1} - \lambda_i}\lambda_i e^{-\lambda_i t}\right) \\
&= K_{n+1}\lambda_{n+1}e^{-\lambda_{n+1}t} + \sum_{i=1}^{n} C_{i,n+1}\lambda_i e^{-\lambda_i t} \qquad (5.7)
\end{aligned}$$

where $K_{n+1} = \sum_{i=1}^{n} C_{i,n}\lambda_i/(\lambda_i - \lambda_{n+1})$ is a constant that does not depend on t. But, we also have that

$$f_{X_1+\cdots+X_{n+1}}(t) = \int_0^t f_{X_2+\cdots+X_{n+1}}(s)\lambda_1 e^{-\lambda_1(t-s)}\,ds$$

which implies, by the same argument that resulted in Equation (5.7), that for a constant K_1

$$f_{X_1+\cdots+X_{n+1}}(t) = K_1\lambda_1 e^{-\lambda_1 t} + \sum_{i=2}^{n+1} C_{i,n+1}\lambda_i e^{-\lambda_i t}$$

Equating these two expressions for $f_{X_1+\cdots+X_{n+1}}(t)$ yields

$$K_{n+1}\lambda_{n+1}e^{-\lambda_{n+1}t} + C_{1,n+1}\lambda_1 e^{-\lambda_1 t} = K_1\lambda_1 e^{-\lambda_1 t} + C_{n+1,n+1}\lambda_{n+1}e^{-\lambda_{n+1}t}$$

Multiplying both sides of the preceding equation by $e^{\lambda_{n+1}t}$ and then letting $t \to \infty$ yields [since $e^{-(\lambda_1-\lambda_{n+1})t} \to 0$ as $t \to \infty$]

$$K_{n+1} = C_{n+1,n+1}$$

and this, using Equation (5.7), completes the induction proof. Thus, we have shown that if $S = \sum_{i=1}^{n} X_i$, then

$$f_S(t) = \sum_{i=1}^{n} C_{i,n}\lambda_i e^{-\lambda_i t} \tag{5.8}$$

where

$$C_{i,n} = \prod_{j\neq i} \frac{\lambda_j}{\lambda_j - \lambda_i}$$

Integrating both sides of the expression for f_S from t to ∞ yields that the tail distribution function of S is given by

$$P\{S > t\} = \sum_{i=1}^{n} C_{i,n}e^{-\lambda_i t} \tag{5.9}$$

Hence, we obtain from Equations (5.8) and (5.9) that $r_S(t)$, the failure rate function of S, is as follows:

$$r_S(t) = \frac{\sum_{i=1}^{n} C_{i,n}\lambda_i e^{-\lambda_i t}}{\sum_{i=1}^{n} C_{i,n}e^{-\lambda_i t}}$$

If we let $\lambda_j = \min(\lambda_1, \ldots, \lambda_n)$, then it follows, upon multiplying the numerator and denominator of $r_S(t)$ by $e^{\lambda_j t}$, that

$$\lim_{t\to\infty} r_S(t) = \lambda_j$$

From the preceding, we can conclude that the remaining lifetime of a hypoexponentially distributed item that has survived to age t is, for t large, approximately that of an exponentially distributed random variable with a rate equal to the minimum of the rates of the random variables whose sums make up the hypoexponential.

Remark Although

$$1 = \int_0^\infty f_S(t)\,dt = \sum_{i=1}^n C_{i,n} = \sum_{i=1}^n \prod_{j\neq i} \frac{\lambda_j}{\lambda_j - \lambda_i}$$

it should not be thought that the $C_{i,n}$, $i = 1, \ldots, n$ are probabilities, because some of them will be negative. Thus, while the form of the hypoexponential density is similar to that of the hyperexponential density (see Example 5.6) these two random variables are very different.

Example 5.11 Let $X_1, \ldots, X_m$ be independent exponential random variables with respective rates $\lambda_1, \ldots, \lambda_m$, where $\lambda_i \neq \lambda_j$ when $i \neq j$. Let N be independent of these random variables and suppose that $\sum_{n=1}^m P_n = 1$, where $P_n = P\{N = n\}$. The random variable

$$Y = \sum_{j=1}^N X_j$$

is said to be a *Coxian* random variable. Conditioning on N gives its density function:

$$\begin{aligned}
f_Y(t) &= \sum_{n=1}^m f_Y(t|N = n)P_n \\
&= \sum_{n=1}^m f_{X_1+\cdots+X_n}(t|N = n)P_n \\
&= \sum_{n=1}^m f_{X_1+\cdots+X_n}(t)P_n \\
&= \sum_{n=1}^m P_n \sum_{i=1}^n C_{i,n}\lambda_i e^{-\lambda_i t}
\end{aligned}$$

Let

$$r(n) = P\{N = n|N \geqslant n\}$$

If we interpret N as a lifetime measured in discrete time periods, then $r(n)$ denotes the probability that an item will die in its nth period of use given that it has survived up to that time. Thus, $r(n)$ is the discrete time analog of the failure rate function $r(t)$, and is correspondingly referred to as the discrete time *failure* (or *hazard*) *rate* function.

Coxian random variables often arise in the following manner. Suppose that an item must go through m stages of treatment to be cured. However, suppose that after each stage there is a probability that the item will quit the program. If we suppose that the amounts of time that it takes the item to pass through the successive stages are independent exponential random variables, and that the probability that an item that has just completed stage n quits the program is (independent of how long it took to go through the n stages) equal to $r(n)$, then the total time that an item spends in the program is a Coxian random variable. ■

5.3 The Poisson Process

5.3.1 Counting Processes

A stochastic process $\{N(t), t \geqslant 0\}$ is said to be a *counting process* if $N(t)$ represents the total number of "events" that occur by time t. Some examples of counting processes are the following:

(a) If we let $N(t)$ equal the number of persons who enter a particular store at or prior to time t, then $\{N(t), t \geqslant 0\}$ is a counting process in which an event corresponds to a person entering the store. Note that if we had let $N(t)$ equal the number of persons in the store at time t, then $\{N(t), t \geqslant 0\}$ would *not* be a counting process (why not?).
(b) If we say that an event occurs whenever a child is born, then $\{N(t), t \geqslant 0\}$ is a counting process when $N(t)$ equals the total number of people who were born by time t. (Does $N(t)$ include persons who have died by time t? Explain why it must.)
(c) If $N(t)$ equals the number of goals that a given soccer player scores by time t, then $\{N(t), t \geqslant 0\}$ is a counting process. An event of this process will occur whenever the soccer player scores a goal.

From its definition we see that for a counting process $N(t)$ must satisfy:

(i) $N(t) \geqslant 0$.
(ii) $N(t)$ is integer valued.
(iii) If $s < t$, then $N(s) \leqslant N(t)$.
(iv) For $s < t$, $N(t) - N(s)$ equals the number of events that occur in the interval $(s, t]$.

A counting process is said to possess *independent increments* if the numbers of events that occur in disjoint time intervals are independent. For example, this means that the number of events that occur by time 10 (that is, $N(10)$) must be independent of the number of events that occur between times 10 and 15 (that is, $N(15) - N(10)$).

The assumption of independent increments might be reasonable for example (a), but it probably would be unreasonable for example (b). The reason for this is that if in example (b) $N(t)$ is very large, then it is probable that there are many people alive at time t; this would lead us to believe that the number of new births between time t and time $t + s$ would also tend to be large (that is, it does not seem reasonable that $N(t)$ is independent of $N(t + s) - N(t)$, and so $\{N(t), t \geqslant 0\}$ would not have independent increments in example (b)). The assumption of independent increments in example (c) would be justified if we believed that the soccer player's chances of scoring a goal today do not depend on "how he's been going." It would not be justified if we believed in "hot streaks" or "slumps."

A counting process is said to possess *stationary increments* if the distribution of the number of events that occur in any interval of time depends only on the length of the time interval. In other words, the process has stationary increments if the number of events in the interval $(s, s+t)$ has the same distribution for all s.

The assumption of stationary increments would only be reasonable in example (a) if there were no times of day at which people were more likely to enter the store. Thus, for instance, if there was a rush hour (say, between 12 P.M. and 1 P.M.) each day, then the stationarity assumption would not be justified. If we believed that the earth's population is basically constant (a belief not held at present by most scientists), then the assumption of stationary increments might be reasonable in example (b). Stationary increments do not seem to be a reasonable assumption in example (c) since, for one thing, most people would agree that the soccer player would probably score more goals while in the age bracket 25–30 than he would while in the age bracket 35–40. It may, however, be reasonable over a smaller time horizon, such as one year.

5.3.2 Definition of the Poisson Process

One of the most important counting processes is the Poisson process, which is defined as follows:

Definition 5.1 The counting process $\{N(t), t \geqslant 0\}$ is said to be a *Poisson process having rate* λ, $\lambda > 0$, if

(i) $N(0) = 0$.
(ii) The process has independent increments.
(iii) The number of events in any interval of length t is Poisson distributed with mean λt. That is, for all s, $t \geqslant 0$

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

Note that it follows from condition (iii) that a Poisson process has stationary increments and also that

$$E[N(t)] = \lambda t$$

which explains why λ is called the rate of the process.

To determine if an arbitrary counting process is actually a Poisson process, we must show that conditions (i), (ii), and (iii) are satisfied. Condition (i), which simply states that the counting of events begins at time $t = 0$, and condition (ii) can usually be directly verified from our knowledge of the process. However, it is not at all clear how we would determine that condition (iii) is satisfied, and for this reason an equivalent definition of a Poisson process would be useful.

As a prelude to giving a second definition of a Poisson process we shall define the concept of a function $f(\cdot)$ being $o(h)$.

Definition 5.2 The function $f(\cdot)$ is said to be $o(h)$ if

$$\lim_{h\to 0}\frac{f(h)}{h}=0$$

Example 5.12

(a) The function $f(x)=x^2$ is $o(h)$ since

$$\lim_{h\to 0}\frac{f(h)}{h}=\lim_{h\to 0}\frac{h^2}{h}=\lim_{h\to 0}h=0$$

(b) The function $f(x)=x$ is not $o(h)$ since

$$\lim_{h\to 0}\frac{f(h)}{h}=\lim_{h\to 0}\frac{h}{h}=\lim_{h\to 0}1=1\neq 0$$

(c) If $f(\cdot)$ is $o(h)$ and $g(\cdot)$ is $o(h)$, then so is $f(\cdot)+g(\cdot)$. This follows since

$$\lim_{h\to 0}\frac{f(h)+g(h)}{h}=\lim_{h\to 0}\frac{f(h)}{h}+\lim_{h\to 0}\frac{g(h)}{h}=0+0=0$$

(d) If $f(\cdot)$ is $o(h)$, then so is $g(\cdot)=cf(\cdot)$. This follows since

$$\lim_{h\to 0}\frac{cf(h)}{h}=c\lim\frac{f(h)}{h}=c\cdot 0=0$$

(e) From (c) and (d) it follows that any finite linear combination of functions, each of which is $o(h)$, is $o(h)$. ■

In order for the function $f(\cdot)$ to be $o(h)$ it is necessary that $f(h)/h$ go to zero as h goes to zero. But if h goes to zero, the only way for $f(h)/h$ to go to zero is for $f(h)$ to go to zero faster than h does. That is, for h small, $f(h)$ must be small compared with h.

We are now in a position to give an alternate definition of a Poisson process.

Definition 5.3 The counting process $\{N(t), t\geqslant 0\}$ is said to be a Poisson process having rate $\lambda, \lambda>0$, if

(i) $N(0)=0$.
(ii) The process has stationary and independent increments.
(iii) $P\{N(h)=1\}=\lambda h+o(h)$.
(iv) $P\{N(h)\geqslant 2\}=o(h)$.

Theorem 5.1 Definitions 5.1 and 5.3 are equivalent.

Proof. We show that Definition 5.3 implies Definition 5.1, and leave it to you to prove the reverse. To start, fix $u \geqslant 0$ and let

$$g(t) = E[\exp\{-uN(t)\}]$$

We derive a differential equation for $g(t)$ as follows:

$$\begin{aligned} g(t+h) &= E[\exp\{-uN(t+h)\}] \\ &= E[\exp\{-uN(t)\} \exp\{-u(N(t+h) - N(t))\}] \\ &= E[\exp\{-uN(t)\}]E[\exp\{-u(N(t+h) - N(t))\}] \\ &\qquad \text{by independent increments} \\ &= g(t)\, E[\exp\{-uN(h)\}] \qquad \text{by stationary increments} \end{aligned} \tag{5.10}$$

Now, assumptions (iii) and (iv) imply that

$$P\{N(h) = 0\} = 1 - \lambda h + o(h)$$

Hence, conditioning on whether $N(h) = 0$ or $N(h) = 1$ or $N(h) \geqslant 2$ yields

$$\begin{aligned} E[\exp\{-uN(h)\}] &= 1 - \lambda h + o(h) + e^{-u}(\lambda h + o(h)) + o(h) \\ &= 1 - \lambda h + e^{-u}\lambda h + o(h) \end{aligned} \tag{5.11}$$

Therefore, from Equations (5.10) and (5.11) we obtain

$$g(t+h) = g(t)(1 - \lambda h + e^{-u}\lambda h) + o(h)$$

implying that

$$\frac{g(t+h) - g(t)}{h} = g(t)\lambda(e^{-u} - 1) + \frac{o(h)}{h}$$

Letting $h \to 0$ gives

$$g'(t) = g(t)\lambda(e^{-u} - 1)$$

or, equivalently,

$$\frac{g'(t)}{g(t)} = \lambda(e^{-u} - 1)$$

Integrating, and using $g(0) = 1$, shows that

$$\log(g(t)) = \lambda t(e^{-u} - 1)$$

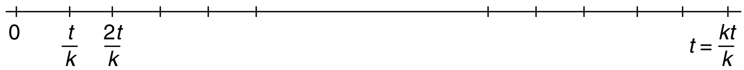

Figure 5.1

or

$$g(t) = \exp\{\lambda t(e^{-u} - 1)\}$$

That is, the Laplace transform of $N(t)$ evaluated at u is $e^{\lambda t(e^{-u}-1)}$. Since that is also the Laplace transform of a Poisson random variable with mean λt, the result follows from the fact that the distribution of a nonnegative random variable is uniquely determined by its Laplace transform. ■

Remarks

(i) The result that $N(t)$ has a Poisson distribution is a consequence of the Poisson approximation to the binomial distribution (see Section 2.2.4). To see this, subdivide the interval $[0, t]$ into k equal parts where k is very large (Figure 5.1). Now it can be shown using axiom (iv) of Definition 5.3 that as k increases to ∞ the probability of having two or more events in any of the k subintervals goes to 0. Hence, $N(t)$ will (with a probability going to 1) just equal the number of subintervals in which an event occurs. However, by stationary and independent increments this number will have a binomial distribution with parameters k and $p = \lambda t/k + o(t/k)$. Hence, by the Poisson approximation to the binomial we see by letting k approach ∞ that $N(t)$ will have a Poisson distribution with mean equal to

$$\begin{aligned}\lim_{k\to\infty} k\left[\lambda\frac{t}{k} + o\left(\frac{t}{k}\right)\right] &= \lambda t + \lim_{k\to\infty}\frac{to(t/k)}{t/k}\\ &= \lambda t\end{aligned}$$

by using the definition of $o(h)$ and the fact that $t/k \to 0$ as $k \to \infty$.

(ii) The explicit assumption that the process has stationary increments can be eliminated from Definition 5.3 provided that we change assumptions (iii) and (iv) to require that for any t the probability of one event in the interval $(t, t+h)$ is $\lambda h + o(h)$ and the probability of two or more events in that interval is $o(h)$. That is, assumptions (ii), (iii), and (iv) of Definition 5.3 can be replaced by

The process has independent increments.

$P\{N(t+h) - N(t) = 1\} = \lambda h + o(h)$.

$P\{N(t+h) - N(t) \geqslant 2\} = o(h)$.

5.3.3 Interarrival and Waiting Time Distributions

Consider a Poisson process, and let us denote the time of the first event by T_1. Further, for $n > 1$, let T_n denote the elapsed time between the $(n-1)$st and the

nth event. The sequence $\{T_n, n = 1, 2, \ldots\}$ is called the *sequence of interarrival times*. For instance, if $T_1 = 5$ and $T_2 = 10$, then the first event of the Poisson process would have occurred at time 5 and the second at time 15.

We shall now determine the distribution of the T_n. To do so, we first note that the event $\{T_1 > t\}$ takes place if and only if no events of the Poisson process occur in the interval $[0, t]$ and thus,

$$P\{T_1 > t\} = P\{N(t) = 0\} = e^{-\lambda t}$$

Hence, T_1 has an exponential distribution with mean $1/\lambda$. Now,

$$P\{T_2 > t\} = E[P\{T_2 > t | T_1\}]$$

However,

$$\begin{aligned} P\{T_2 > t \mid T_1 = s\} &= P\{0 \text{ events in } (s, s+t] \mid T_1 = s\} \\ &= P\{0 \text{ events in } (s, s+t]\} \\ &= e^{-\lambda t} \end{aligned} \tag{5.12}$$

where the last two equations followed from independent and stationary increments. Therefore, from Equation (5.12) we conclude that T_2 is also an exponential random variable with mean $1/\lambda$ and, furthermore, that T_2 is independent of T_1. Repeating the same argument yields the following.

Proposition 5.1 $T_n, n = 1, 2, \ldots$, are independent identically distributed exponential random variables having mean $1/\lambda$.

Remark The proposition should not surprise us. The assumption of stationary and independent increments is basically equivalent to asserting that, at any point in time, the process *probabilistically* restarts itself. That is, the process from any point on is independent of all that has previously occurred (by independent increments), and also has the same distribution as the original process (by stationary increments). In other words, the process has no *memory*, and hence exponential interarrival times are to be expected.

Another quantity of interest is S_n, the arrival time of the nth event, also called the *waiting time* until the nth event. It is easily seen that

$$S_n = \sum_{i=1}^{n} T_i, \qquad n \geqslant 1$$

and hence from Proposition 5.1 and the results of Section 2.2 it follows that S_n has a gamma distribution with parameters n and λ. That is, the probability density of S_n is given by

$$f_{S_n}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}, \qquad t \geqslant 0 \tag{5.13}$$

Equation (5.13) may also be derived by noting that the nth event will occur prior to or at time t if and only if the number of events occurring by time t is at least n. That is,

$$N(t) \geqslant n \quad \Leftrightarrow \quad S_n \leqslant t$$

Hence,

$$F_{S_n}(t) = P\{S_n \leqslant t\} = P\{N(t) \geqslant n\} = \sum_{j=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!}$$

which, upon differentiation, yields

$$\begin{aligned} f_{S_n}(t) &= -\sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^j}{j!} + \sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{j-1}}{(j-1)!} \\ &= \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} + \sum_{j=n+1}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{j-1}}{(j-1)!} - \sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^j}{j!} \\ &= \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \end{aligned}$$

Example 5.13 Suppose that people immigrate into a territory at a Poisson rate $\lambda = 1$ per day.

(a) What is the expected time until the tenth immigrant arrives?
(b) What is the probability that the elapsed time between the tenth and the eleventh arrival exceeds two days?

Solution:
(a) $E[S_{10}] = 10/\lambda = 10$ days.
(b) $P\{T_{11} > 2\} = e^{-2\lambda} = e^{-2} \approx 0.133$. ■

Proposition 5.1 also gives us another way of defining a Poisson process. Suppose we start with a sequence $\{T_n, n \geqslant 1\}$ of independent identically distributed exponential random variables each having mean $1/\lambda$. Now let us define a counting process by saying that the nth event of this process occurs at time

$$S_n \equiv T_1 + T_2 + \cdots + T_n$$

The resultant counting process $\{N(t), t \geqslant 0\}$* will be Poisson with rate λ.

* A formal definition of $N(t)$ is given by $N(t) \equiv \max\{n\colon S_n \leqslant t\}$ where $S_0 \equiv 0$.

Remark Another way of obtaining the density function of S_n is to note that because S_n is the time of the nth event,

$$\begin{aligned}
P\{t < S_n < t+h\} &= P\{N(t) = n-1, \text{ one event in } (t, t+h)\} + o(h) \\
&= P\{N(t) = n-1\}P\{\text{one event in } (t, t+h)\} + o(h) \\
&= e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}[\lambda h + o(h)] + o(h) \\
&= \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}h + o(h)
\end{aligned}$$

where the first equality uses the fact that the probability of 2 or more events in $(t, t+h)$ is $o(h)$. If we now divide both sides of the preceding equation by h and then let $h \to 0$, we obtain

$$f_{S_n}(t) = \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}$$

5.3.4 Further Properties of Poisson Processes

Consider a Poisson process $\{N(t), t \geqslant 0\}$ having rate λ, and suppose that each time an event occurs it is classified as either a type I or a type II event. Suppose further that each event is classified as a type I event with probability p or a type II event with probability $1-p$, independently of all other events. For example, suppose that customers arrive at a store in accordance with a Poisson process having rate λ; and suppose that each arrival is male with probability $\frac{1}{2}$ and female with probability $\frac{1}{2}$. Then a type I event would correspond to a male arrival and a type II event to a female arrival.

Let $N_1(t)$ and $N_2(t)$ denote respectively the number of type I and type II events occurring in $[0, t]$. Note that $N(t) = N_1(t) + N_2(t)$.

Proposition 5.2 $\{N_1(t), t \geqslant 0\}$ and $\{N_2(t), t \geqslant 0\}$ are both Poisson processes having respective rates λp and $\lambda(1-p)$. Furthermore, the two processes are independent.

Proof. It is easy to verify that $\{N_1(t), t \geqslant 0\}$ is a Poisson process with rate λp by verifying that it satisfies Definition 5.3.

- $N_1(0) = 0$ follows from the fact that $N(0) = 0$.
- It is easy to see that $\{N_1(t), t \geqslant 0\}$ inherits the stationary and independent increment properties of the process $\{N(t), t \geqslant 0\}$. This is true because the distribution of the number of type I events in an interval can be obtained by conditioning on the number of events in that interval, and the distribution of this latter quantity depends only on the length of the interval and is independent of what has occurred in any nonoverlapping interval.

- $$\begin{aligned} P\{N_1(h) = 1\} &= P\{N_1(h) = 1 \mid N(h) = 1\}P\{N(h) = 1\} \\ &\quad + P\{N_1(h) = 1 \mid N(h) \geqslant 2\}P\{N(h) \geqslant 2\} \\ &= p(\lambda h + o(h)) + o(h) \\ &= \lambda ph + o(h) \end{aligned}$$
- $P\{N_1(h) \geqslant 2\} \leqslant P\{N(h) \geqslant 2\} = o(h)$

Thus we see that $\{N_1(t),\ t \geqslant 0\}$ is a Poisson process with rate λp and, by a similar argument, that $\{N_2(t), t \geqslant 0\}$ is a Poisson process with rate $\lambda(1 - p)$. Because the probability of a type I event in the interval from t to $t + h$ is independent of all that occurs in intervals that do not overlap $(t, t + h)$, it is independent of knowledge of when type II events occur, showing that the two Poisson processes are independent. (For another way of proving independence, see Example 3.23.) ■

Example 5.14 If immigrants to area A arrive at a Poisson rate of ten per week, and if each immigrant is of English descent with probability $\frac{1}{12}$, then what is the probability that no people of English descent will emigrate to area A during the month of February?

Solution: By the previous proposition it follows that the number of Englishmen emigrating to area A during the month of February is Poisson distributed with mean $4 \cdot 10 \cdot \frac{1}{12} = \frac{10}{3}$. Hence, the desired probability is $e^{-10/3}$. ■

Example 5.15 Suppose nonnegative offers to buy an item that you want to sell arrive according to a Poisson process with rate λ. Assume that each offer is the value of a continuous random variable having density function $f(x)$. Once the offer is presented to you, you must either accept it or reject it and wait for the next offer. We suppose that you incur costs at a rate c per unit time until the item is sold, and that your objective is to maximize your expected total return, where the total return is equal to the amount received minus the total cost incurred. Suppose you employ the policy of accepting the first offer that is greater than some specified value y. (Such a type of policy, which we call a y-policy, can be shown to be optimal.) What is the best value of y?

Solution: Let us compute the expected total return when you use the y-policy, and then choose the value of y that maximizes this quantity. Let X denote the value of a random offer, and let $\bar{F}(x) = P\{X > x\} = \int_x^\infty f(u)\,du$ be its tail distribution function. Because each offer will be greater than y with probability $\bar{F}(y)$, it follows that such offers occur according to a Poisson process with rate $\lambda\bar{F}(y)$. Hence, the time until an offer is accepted is an exponential random variable with rate $\lambda\bar{F}(y)$. Letting $R(y)$ denote the total return from the policy that accepts the first offer that is greater than y, we have

$$\begin{aligned} E[R(y)] &= E[\text{accepted offer}] - cE[\text{time to accept}] \\ &= E[X|X > y] - \frac{c}{\lambda\bar{F}(y)} \end{aligned}$$

$$= \int_0^\infty x f_{X|X>y}(x)\,dx - \frac{c}{\lambda \bar{F}(y)}$$
$$= \int_y^\infty x \frac{f(x)}{\bar{F}(y)}\,dx - \frac{c}{\lambda \bar{F}(y)}$$
$$= \frac{\int_y^\infty x f(x)\,dx - c/\lambda}{\bar{F}(y)}$$

Differentiation yields

$$\frac{d}{dy}E[R(y)] = 0 \Leftrightarrow -\bar{F}(y) y f(y) + \left(\int_y^\infty x f(x)\,dx - \frac{c}{\lambda}\right) f(y) = 0$$

Therefore, the optimal value of y satisfies

$$y\bar{F}(y) = \int_y^\infty x f(x)\,dx - \frac{c}{\lambda}$$

or

$$y \int_y^\infty f(x)\,dx = \int_y^\infty x f(x)\,dx - \frac{c}{\lambda}$$

or

$$\int_y^\infty (x - y) f(x)\,dx = \frac{c}{\lambda} \tag{5.14}$$

We now argue that the left-hand side of the preceding is a nonincreasing function of y. To do so, note that, with a^+ defined to equal a if $a > 0$ or to equal 0 otherwise, we have

$$\int_y^\infty (x - y) f(x)\,dx = E[(X - y)^+]$$

Because $(X - y)^+$ is a nonincreasing function of y, so is its expectation, thus showing that the left hand side of Equation (5.14) is a nonincreasing function of y. Consequently, if $E[X] < c/\lambda$—in which case there is no solution of Equation (5.14)—then it is optimal to accept any offer; otherwise, the optimal value y is the unique solution of Equation (5.14). ■

It follows from Proposition 5.2 that if each of a Poisson number of individuals is independently classified into one of two possible groups with respective probabilities p and $1 - p$, then the number of individuals in each of the two groups will be independent Poisson random variables. Because this result easily generalizes to the case where the classification is into any one of r possible groups, we have the following application to a model of employees moving about in an organization.

Example 5.16 Consider a system in which individuals at any time are classified as being in one of r possible states, and assume that an individual changes states in accordance with a Markov chain having transition probabilities $P_{ij}, i, j = 1, \ldots, r$. That is, if an individual is in state i during a time period then, independently of its previous states, it will be in state j during the next time period with probability P_{ij}. The individuals are assumed to move through the system independently of each other. Suppose that the numbers of people initially in states $1, 2, \ldots, r$ are independent Poisson random variables with respective means $\lambda_1, \lambda_2, \ldots, \lambda_r$. We are interested in determining the joint distribution of the numbers of individuals in states $1, 2, \ldots, r$ at some time n.

Solution: For fixed i, let $N_j(i), j = 1, \ldots, r$ denote the number of those individuals, initially in state i, that are in state j at time n. Now each of the (Poisson distributed) number of people initially in state i will, independently of each other, be in state j at time n with probability P_{ij}^n, where P_{ij}^n is the n-stage transition probability for the Markov chain having transition probabilities P_{ij}. Hence, the $N_j(i), j = 1, \ldots, r$ will be independent Poisson random variables with respective means $\lambda_i P_{ij}^n, j = 1, \ldots, r$. Because the sum of independent Poisson random variables is itself a Poisson random variable, it follows that the number of individuals in state j at time n—namely $\sum_{i=1}^r N_j(i)$—will be independent Poisson random variables with respective means $\sum_i \lambda_i P_{ij}^n$, for $j = 1, \ldots, r$. ■

Example 5.17 (The Coupon Collecting Problem) There are m different types of coupons. Each time a person collects a coupon it is, independently of ones previously obtained, a type j coupon with probability p_j, $\sum_{j=1}^m p_j = 1$. Let N denote the number of coupons one needs to collect in order to have a complete collection of at least one of each type. Find $E[N]$.

Solution: If we let N_j denote the number one must collect to obtain a type j coupon, then we can express N as

$$N = \max_{1 \leqslant j \leqslant m} N_j$$

However, even though each N_j is geometric with parameter p_j, the foregoing representation of N is not that useful, because the random variables N_j are not independent.

We can, however, transform the problem into one of determining the expected value of the maximum of *independent* random variables. To do so, suppose that coupons are collected at times chosen according to a Poisson process with rate $\lambda = 1$. Say that an event of this Poisson process is of type j, $1 \leqslant j \leqslant m$, if the coupon obtained at that time is a type j coupon. If we now let $N_j(t)$ denote the number of type j coupons collected by time t, then it follows from Proposition 5.2 that $\{N_j(t), t \geqslant 0\}, j = 1, \ldots, m$ are independent Poisson processes with respective rates $\lambda p_j = p_j$. Let X_j denote the time of the first event of the jth process, and let

$$X = \max_{1 \leqslant j \leqslant m} X_j$$

denote the time at which a complete collection is amassed. Since the X_j are independent exponential random variables with respective rates p_j, it follows that

$$\begin{aligned} P\{X < t\} &= P\{\max_{1 \leqslant j \leqslant m} X_j < t\} \\ &= P\{X_j < t, \text{ for } j = 1, \ldots, m\} \\ &= \prod_{j=1}^{m} (1 - e^{-p_j t}) \end{aligned}$$

Therefore,

$$\begin{aligned} E[X] &= \int_0^\infty P\{X > t\}\, dt \\ &= \int_0^\infty \left\{1 - \prod_{j=1}^{m} (1 - e^{-p_j t})\right\} dt \end{aligned} \tag{5.15}$$

It remains to relate $E[X]$, the expected time until one has a complete set, to $E[N]$, the expected number of coupons it takes. This can be done by letting T_i denote the ith interarrival time of the Poisson process that counts the number of coupons obtained. Then it is easy to see that

$$X = \sum_{i=1}^{N} T_i$$

Since the T_i are independent exponentials with rate 1, and N is independent of the T_i, we see that

$$E[X|N] = NE[T_i] = N$$

Therefore,

$$E[X] = E[N]$$

and so $E[N]$ is as given in Equation (5.15).

Let us now compute the expected number of types that appear only once in the complete collection. Letting I_i equal 1 if there is only a single type i coupon in the final set, and letting it equal 0 otherwise, we thus want

$$\begin{aligned} E\left[\sum_{i=1}^{m} I_i\right] &= \sum_{i=1}^{m} E[I_i] \\ &= \sum_{i=1}^{m} P\{I_i = 1\} \end{aligned}$$

Now there will be a single type i coupon in the final set if a coupon of each type has appeared before the second coupon of type i is obtained. Thus, letting S_i denote the time at which the second type i coupon is obtained, we have

$$P\{I_i = 1\} = P\{X_j < S_i, \text{ for all } j \neq i\}$$

Using that S_i has a gamma distribution with parameters $(2, p_i)$, this yields

$$\begin{aligned} P\{I_i = 1\} &= \int_0^\infty P\{X_j < S_i \text{ for all } j \neq i | S_i = x\} p_i e^{-p_i x} p_i x \, dx \\ &= \int_0^\infty P\{X_j < x, \text{ for all } j \neq i\} p_i^2 \, x \, e^{-p_i x} \, dx \\ &= \int_0^\infty \prod_{j \neq i} (1 - e^{-p_j x}) p_i^2 x e^{-p_i x} \, dx \end{aligned}$$

Therefore, we have the result

$$\begin{aligned} E\left[\sum_{i=1}^{m} I_i\right] &= \int_0^\infty \sum_{i=1}^{m} \prod_{j \neq i} (1 - e^{-p_j x}) p_i^2 x e^{-p_i x} \, dx \\ &= \int_0^\infty x \prod_{j=1}^{m} (1 - e^{-p_j x}) \sum_{i=1}^{m} p_i^2 \frac{e^{-p_i x}}{1 - e^{-p_i x}} \, dx \end{aligned}$$ ■

The next probability calculation related to Poisson processes that we shall determine is the probability that n events occur in one Poisson process before m events have occurred in a second and independent Poisson process. More formally let $\{N_1(t), t \geqslant 0\}$ and $\{N_2(t), t \geqslant 0\}$ be two independent Poisson processes having respective rates λ_1 and λ_2. Also, let S_n^1 denote the time of the nth event of the first process, and S_m^2 the time of the mth event of the second process. We seek

$$P\{S_n^1 < S_m^2\}$$

Before attempting to calculate this for general n and m, let us consider the special case $n = m = 1$. Since S_1^1, the time of the first event of the $N_1(t)$ process, and S_1^2, the time of the first event of the $N_2(t)$ process, are both exponentially distributed random variables (by Proposition 5.1) with respective means $1/\lambda_1$ and $1/\lambda_2$, it follows from Section 5.2.3 that

$$P\{S_1^1 < S_1^2\} = \frac{\lambda_1}{\lambda_1 + \lambda_2} \tag{5.16}$$

Let us now consider the probability that two events occur in the $N_1(t)$ process before a single event has occurred in the $N_2(t)$ process. That is, $P\{S_2^1 < S_1^2\}$. To calculate this we reason as follows: In order for the $N_1(t)$ process to have two events before a single event occurs in the $N_2(t)$ process, it is first necessary for the initial event that occurs to be an event of the $N_1(t)$ process (and this occurs, by Equation (5.16), with probability $\lambda_1/(\lambda_1 + \lambda_2)$). Now, given that the initial event is from the $N_1(t)$ process, the next thing that must occur for S_2^1 to be less than S_1^2 is for the second event also to be an event of the $N_1(t)$ process. However, when the first event occurs both processes start all over again (by the memoryless property of Poisson processes) and hence this conditional probability is also $\lambda_1/(\lambda_1 + \lambda_2)$; thus, the desired probability is given by

$$P\{S_2^1 < S_1^2\} = \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^2$$

In fact, this reasoning shows that *each event that occurs is going to be an event of the $N_1(t)$ process with probability $\lambda_1/(\lambda_1 + \lambda_2)$ or an event of the $N_2(t)$ process with probability $\lambda_2/(\lambda_1 + \lambda_2)$, independent of all that has previously occurred*. In other words, the probability that the $N_1(t)$ process reaches n before the $N_2(t)$ process reaches m is just the probability that n heads will appear before m tails if one flips a coin having probability $p = \lambda_1/(\lambda_1 + \lambda_2)$ of a head appearing. But by noting that this event will occur if and only if the first $n + m - 1$ tosses result in n or more heads, we see that our desired probability is given by

$$P\{S_n^1 < S_m^2\} = \sum_{k=n}^{n+m-1} \binom{n+m-1}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n+m-1-k}$$

5.3.5 Conditional Distribution of the Arrival Times

Suppose we are told that exactly one event of a Poisson process has taken place by time t, and we are asked to determine the distribution of the time at which the event occurred. Now, since a Poisson process possesses stationary and independent increments it seems reasonable that each interval in $[0, t]$ of equal length should have the same probability of containing the event. In other words, the time of the event should be uniformly distributed over $[0, t]$. This is easily checked since, for $s \leqslant t$,

$$
\begin{aligned}
P\{T_1 < s | N(t) = 1\} &= \frac{P\{T_1 < s, N(t) = 1\}}{P\{N(t) = 1\}} \\
&= \frac{P\{1 \text{ event in } [0, s), 0 \text{ events in } [s, t]\}}{P\{N(t) = 1\}} \\
&= \frac{P\{1 \text{ event in } [0, s)\} P\{0 \text{ events in } [s, t]\}}{P\{N(t) = 1\}} \\
&= \frac{\lambda s e^{-\lambda s} e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}} \\
&= \frac{s}{t}
\end{aligned}
$$

This result may be generalized, but before doing so we need to introduce the concept of order statistics.

Let $Y_1, Y_2, \dots, Y_n$ be n random variables. We say that $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$ are the *order statistics* corresponding to $Y_1, Y_2, \dots, Y_n$ if $Y_{(k)}$ is the kth smallest value among $Y_1, \dots, Y_n$, $k = 1, 2, \dots, n$. For instance, if $n = 3$ and $Y_1 = 4$, $Y_2 = 5$, $Y_3 = 1$ then $Y_{(1)} = 1$, $Y_{(2)} = 4$, $Y_{(3)} = 5$. If the Y_i, $i = 1, \dots, n$, are independent identically distributed continuous random variables with probability density f, then the joint density of the order statistics $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$ is given by

$$
f(y_1, y_2, \dots, y_n) = n! \prod_{i=1}^{n} f(y_i), \qquad y_1 < y_2 < \cdots < y_n
$$

The preceding follows since

(i) $(Y_{(1)}, Y_{(2)}, \dots, Y_{(n)})$ will equal $(y_1, y_2, \dots, y_n)$ if $(Y_1, Y_2, \dots, Y_n)$ is equal to any of the $n!$ permutations of $(y_1, y_2, \dots, y_n)$;

and

(ii) the probability density that $(Y_1, Y_2, \dots, Y_n)$ is equal to $y_{i_1}, \dots, y_{i_n}$ is $\prod_{j=1}^{n} f(y_{i_j}) = \prod_{j=1}^{n} f(y_j)$ when $i_1, \dots, i_n$ is a permutation of $1, 2, \dots, n$.

If the $Y_i, i = 1, \dots, n$, are uniformly distributed over $(0, t)$, then we obtain from the preceding that the joint density function of the order statistics $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$ is

$$
f(y_1, y_2, \dots, y_n) = \frac{n!}{t^n}, \qquad 0 < y_1 < y_2 < \cdots < y_n < t
$$

We are now ready for the following useful theorem.

Theorem 5.2 Given that $N(t) = n$, the n arrival times $S_1, \dots, S_n$ have the same distribution as the order statistics corresponding to n independent random variables uniformly distributed on the interval $(0, t)$.

Proof. To obtain the conditional density of $S_1, \dots, S_n$ given that $N(t) = n$ note that for $0 < s_i < \cdots < s_n < t$ the event that $S_1 = s_1, S_2 = s_2, \dots, S_n = s_n, N(t) = n$ is equivalent to the event that the first $n + 1$ interarrival times satisfy $T_1 = s_1$, $T_2 = s_2 - s_1, \dots, T_n = s_n - s_{n-1}, T_{n+1} > t - s_n$. Hence, using Proposition 5.1, we have that the conditional joint density of $S_1, \dots, S_n$ given that $N(t) = n$ is as follows:

$$\begin{aligned} f(s_1, \dots, s_n \mid n) &= \frac{f(s_1, \dots, s_n, n)}{P\{N(t) = n\}} \\ &= \frac{\lambda e^{-\lambda s_1} \lambda e^{-\lambda(s_2 - s_1)} \cdots \lambda e^{-\lambda(s_n - s_{n-1})} e^{-\lambda(t - s_n)}}{e^{-\lambda t}(\lambda t)^n / n!} \\ &= \frac{n!}{t^n}, \qquad 0 < s_1 < \cdots < s_n < t \end{aligned}$$

which proves the result. ■

Remark The preceding result is usually paraphrased as stating that, under the condition that n events have occurred in $(0, t)$, the times $S_1, \dots, S_n$ at which events occur, considered as unordered random variables, are distributed independently and uniformly in the interval $(0, t)$.

Application of Theorem 5.2 (Sampling a Poisson Process) In Proposition 5.2 we showed that if each event of a Poisson process is independently classified as a type I event with probability p and as a type II event with probability $1 - p$ then the counting processes of type I and type II events are independent Poisson processes with respective rates λp and $\lambda(1-p)$. Suppose now, however, that there are k possible types of events and that the probability that an event is classified as a type i event, $i = 1, \dots, k$, depends on the time the event occurs. Specifically, suppose that if an event occurs at time y then it will be classified as a type i event, independently of anything that has previously occurred, with probability $P_i(y), i = 1, \dots, k$ where $\sum_{i=1}^k P_i(y) = 1$. Upon using Theorem 5.2 we can prove the following useful proposition.

Proposition 5.3 If $N_i(t), i = 1, \dots, k$, represents the number of type i events occurring by time t then $N_i(t), i = 1, \dots, k$, are independent Poisson random variables having means

$$E[N_i(t)] = \lambda \int_0^t P_i(s)\, ds$$

Before proving this proposition, let us first illustrate its use.

Example 5.18 (An Infinite Server Queue) Suppose that customers arrive at a service station in accordance with a Poisson process with rate λ. Upon arrival the customer is immediately served by one of an infinite number of possible servers, and the service times are assumed to be independent with a common distribution G. What is the distribution of $X(t)$, the number of customers that have completed service by time t? What is the distribution of $Y(t)$, the number of customers that are being served at time t?

To answer the preceding questions let us agree to call an entering customer a type I customer if he completes his service by time t and a type II customer if he does not complete his service by time t. Now, if the customer enters at time s, $s \leqslant t$, then he will be a type I customer if his service time is less than $t - s$. Since the service time distribution is G, the probability of this will be $G(t - s)$. Similarly, a customer entering at time s, $s \leqslant t$, will be a type II customer with probability $\bar{G}(t-s) = 1 - G(t-s)$. Hence, from Proposition 5.3 we obtain that the distribution of $X(t)$, the number of customers that have completed service by time t, is Poisson distributed with mean

$$E[X(t)] = \lambda \int_0^t G(t-s)\, ds = \lambda \int_0^t G(y)\, dy \tag{5.17}$$

Similarly, the distribution of $Y(t)$, the number of customers being served at time t is Poisson with mean

$$E[Y(t)] = \lambda \int_0^t \bar{G}(t-s)\, ds = \lambda \int_0^t \bar{G}(y)\, dy \tag{5.18}$$

Furthermore, $X(t)$ and $Y(t)$ are independent.

Suppose now that we are interested in computing the joint distribution of $Y(t)$ and $Y(t + s)$—that is, the joint distribution of the number in the system at time t and at time $t + s$. To accomplish this, say that an arrival is

type 1: if he arrives before time t and completes service between t and $t + s$,
type 2: if he arrives before t and completes service after $t + s$,
type 3: if he arrives between t and $t + s$ and completes service after $t + s$,
type 4: otherwise.

Hence, an arrival at time y will be type i with probability $P_i(y)$ given by

$$P_1(y) = \begin{cases} G(t+s-y) - G(t-y), & \text{if } y < t \\ 0, & \text{otherwise} \end{cases}$$

$$P_2(y) = \begin{cases} \bar{G}(t+s-y), & \text{if } y < t \\ 0, & \text{otherwise} \end{cases}$$

$$P_3(y) = \begin{cases} \bar{G}(t+s-y), & \text{if } t < y < t+s \\ 0, & \text{otherwise} \end{cases}$$

$$P_4(y) = 1 - P_1(y) - P_2(y) - P_3(y)$$

Thus, if $N_i = N_i(s + t)$, $i = 1, 2, 3$, denotes the number of type i events that occur, then from Proposition 5.3, $N_i, i = 1, 2, 3$, are independent Poisson random variables with respective means

$$E[N_i] = \lambda \int_0^{t+s} P_i(y)\, dy, \qquad i = 1, 2, 3$$

Because

$$\begin{aligned} Y(t) &= N_1 + N_2, \\ Y(t+s) &= N_2 + N_3 \end{aligned}$$

it is now an easy matter to compute the joint distribution of $Y(t)$ and $Y(t + s)$. For instance,

$$\begin{aligned} &\text{Cov}[Y(t), Y(t+s)] \\ &\quad = \text{Cov}(N_1 + N_2, N_2 + N_3) \\ &\quad = \text{Cov}(N_2, N_2) \qquad \text{by independence of } N_1, N_2, N_3 \\ &\quad = \text{Var}(N_2) \\ &\quad = \lambda \int_0^t \bar{G}(t + s - y)\, dy = \lambda \int_0^t \bar{G}(u + s)\, du \end{aligned}$$

where the last equality follows since the variance of a Poisson random variable equals its mean, and from the substitution $u = t - y$. Also, the joint distribution of $Y(t)$ and $Y(t + s)$ is as follows:

$$\begin{aligned} P\{Y(t) = i, Y(t+s) = j\} &= P\{N_1 + N_2 = i, N_2 + N_3 = j\} \\ &= \sum_{l=0}^{\min(i,j)} P\{N_2 = l, N_1 = i - l, N_3 = j - l\} \\ &= \sum_{l=0}^{\min(i,j)} P\{N_2 = l\} P\{N_1 = i - l\} P\{N_3 = j - l\} \end{aligned}$$

■

Example 5.19 (Minimizing the Number of Encounters) Suppose that cars enter a one-way highway in accordance with a Poisson process with rate λ. The cars enter at point a and depart at point b (see Figure 5.2). Each car travels at a constant speed that is randomly determined, independently from car to car, from the distribution G. When a faster car encounters a slower one, it passes it with no time being lost. If your car enters the highway at time s and you are able to choose your speed, what speed minimizes the expected number of encounters you

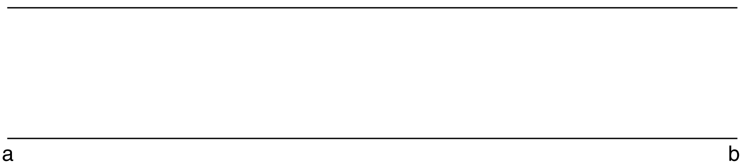

Figure 5.2 Cars enter at point a and depart at b.

will have with other cars, where we say that an encounter occurs each time your car either passes or is passed by another car?

Solution: We will show that for large s the speed that minimizes the expected number of encounters is the median of the speed distribution G. To see this, suppose that the speed x is chosen. Let $d = b - a$ denote the length of the road. Upon choosing the speed x, it follows that your car will enter the road at time s and will depart at time $s + t_0$, where $t_0 = d/x$ is the travel time.

Now, the other cars enter the road according to a Poisson process with rate λ. Each of them chooses a speed X according to the distribution G, and this results in a travel time $T = d/X$. Let F denote the distribution of travel time T. That is,

$$F(t) = P\{T \leqslant t\} = P\{d/X \leqslant t\} = P\{X \geqslant d/t\} = \bar{G}(d/t)$$

Let us say that an event occurs at time t if a car enters the highway at time t. Also, say that the event is a type 1 event if it results in an encounter with your car. Now, your car will enter the road at time s and will exit at time $s + t_0$. Hence, a car will encounter your car if it enters before s and exits after $s + t_0$ (in which case your car will pass it on the road) or if it enters after s but exits before $s + t_0$ (in which case it will pass your car). As a result, a car that enters the road at time t will encounter your car if its travel time T is such that

$$\begin{aligned} &t + T > s + t_0, && \text{if } t < s \\ &t + T < s + t_0, && \text{if } s < t < s + t_0 \end{aligned}$$

From the preceding, we see that an event at time t will, independently of other events, be a type 1 event with probability $p(t)$ given by

$$p(t) = \begin{cases} P\{t + T > s + t_0\} = \bar{F}(s + t_0 - t), & \text{if } t < s \\ P\{t + T < s + t_0\} = F(s + t_0 - t), & \text{if } s < t < s + t_0 \\ 0, & \text{if } t > s + t_0 \end{cases}$$

Since events (that is, cars entering the road) are occurring according to a Poisson process it follows, upon applying Proposition 5.3, that the total number of type 1 events that ever occurs is Poisson with mean

$$\begin{aligned}\lambda \int_0^\infty p(t)\,dt &= \lambda \int_0^s \bar{F}(s+t_0-t)\,dt + \lambda \int_s^{s+t_0} F(s+t_0-t)\,dt \\ &= \lambda \int_{t_0}^{s+t_0} \bar{F}(y)\,dy + \lambda \int_0^{t_0} F(y)\,dy\end{aligned}$$

To choose the value of t_0 that minimizes the preceding quantity, we differentiate. This gives

$$\frac{d}{dt_0}\left\{\lambda \int_0^\infty p(t)\,dt\right\} = \lambda\{\bar{F}(s+t_0) - \bar{F}(t_0) + F(t_0)\}$$

Setting this equal to 0, and using that $\bar{F}(s+t_0) \approx 0$ when s is large, we see that the optimal travel time t_0 is such that

$$F(t_0) - \bar{F}(t_0) = 0$$

or

$$F(t_0) - [1 - F(t_0)] = 0$$

or

$$F(t_0) = \tfrac{1}{2}$$

That is, the optimal travel time is the median of the travel time distribution. Since the speed X is equal to the distance d divided by the travel time T, it follows that the optimal speed $x_0 = d/t_0$ is such that

$$F(d/x_0) = \tfrac{1}{2}$$

Since

$$F(d/x_0) = \bar{G}(x_0)$$

we see that $\bar{G}(x_0) = \frac{1}{2}$, and so the optimal speed is the median of the distribution of speeds.

Summing up, we have shown that for any speed x the number of encounters with other cars will be a Poisson random variable, and the mean of this Poisson will be smallest when the speed x is taken to be the median of the distribution G. ■

Example 5.20 (Tracking the Number of HIV Infections) There is a relatively long incubation period from the time when an individual becomes infected with the HIV virus, which causes AIDS, until the symptoms of the disease appear. As a result, it is difficult for public health officials to be certain of the number of members of the population that are infected at any given time. We will now present a first approximation model for this phenomenon, which can be used to obtain a rough estimate of the number of infected individuals.

Let us suppose that individuals contract the HIV virus in accordance with a Poisson process whose rate λ is unknown. Suppose that the time from when an individual becomes infected until symptoms of the disease appear is a random variable having a known distribution G. Suppose also that the incubation times of different infected individuals are independent.

Let $N_1(t)$ denote the number of individuals who have shown symptoms of the disease by time t. Also, let $N_2(t)$ denote the number who are HIV positive but have not yet shown any symptoms by time t. Now, since an individual who contracts the virus at time s will have symptoms by time t with probability $G(t-s)$ and will not with probability $\bar{G}(t-s)$, it follows from Proposition 5.3 that $N_1(t)$ and $N_2(t)$ are independent Poisson random variables with respective means

$$E[N_1(t)] = \lambda \int_0^t G(t-s)\,ds = \lambda \int_0^t G(y)\,dy$$

and

$$E[N_2(t)] = \lambda \int_0^t \bar{G}(t-s)\,ds = \lambda \int_0^t \bar{G}(y)\,dy$$

Now, if we knew λ, then we could use it to estimate $N_2(t)$, the number of individuals infected but without any outward symptoms at time t, by its mean value $E[N_2(t)]$. However, since λ is unknown, we must first estimate it. Now, we will presumably know the value of $N_1(t)$, and so we can use its known value as an estimate of its mean $E[N_1(t)]$. That is, if the number of individuals who have exhibited symptoms by time t is n_1, then we can estimate that

$$n_1 \approx E[N_1(t)] = \lambda \int_0^t G(y)\,dy$$

Therefore, we can estimate λ by the quantity $\hat{\lambda}$ given by

$$\hat{\lambda} = n_1 \Big/ \int_0^t G(y)\,dy$$

Using this estimate of λ, we can estimate the number of infected but symptomless individuals at time t by

$$\begin{aligned}\text{estimate of } N_2(t) &= \hat{\lambda} \int_0^t \bar{G}(y)\,dy \\ &= \frac{n_1 \int_0^t \bar{G}(y)\,dy}{\int_0^t G(y)\,dy}\end{aligned}$$

For example, suppose that G is exponential with mean μ. Then $\bar{G}(y) = e^{-y/\mu}$, and a simple integration gives that

$$\text{estimate of } N_2(t) = \frac{n_1 \mu(1 - e^{-t/\mu})}{t - \mu(1 - e^{-t/\mu})}$$

If we suppose that $t = 16$ years, $\mu = 10$ years, and $n_1 = 220$ thousand, then the estimate of the number of infected but symptomless individuals at time 16 is

$$\text{estimate} = \frac{2{,}200(1 - e^{-1.6})}{16 - 10(1 - e^{-1.6})} = 218.96$$

That is, if we suppose that the foregoing model is approximately correct (and we should be aware that the assumption of a constant infection rate λ that is unchanging over time is almost certainly a weak point of the model), then if the incubation period is exponential with mean 10 years and if the total number of individuals who have exhibited AIDS symptoms during the first 16 years of the epidemic is 220 thousand, then we can expect that approximately 219 thousand individuals are HIV positive though symptomless at time 16. ■

Proof of Proposition 5.3. Let us compute the joint probability $P\{N_i(t) = n_i, i = 1, \ldots, k\}$. To do so note first that in order for there to have been n_i type i events for $i = 1, \ldots, k$ there must have been a total of $\sum_{i=1}^{k} n_i$ events. Hence, conditioning on $N(t)$ yields

$$\begin{aligned} P\{N_1(t) &= n_1, \ldots, N_k(t) = n_k\} \\ &= P\left\{N_1(t) = n_1, \ldots, N_k(t) = n_k \;\middle|\; N(t) = \sum_{i=1}^{k} n_i\right\} \\ &\quad \times P\left\{N(t) = \sum_{i=1}^{k} n_i\right\} \end{aligned}$$

Now consider an arbitrary event that occurred in the interval $[0, t]$. If it had occurred at time s, then the probability that it would be a type i event would be $P_i(s)$. Hence, since by Theorem 5.2 this event will have occurred at some time uniformly distributed on $[0, t]$, it follows that the probability that this event will be a type i event is

$$P_i = \frac{1}{t} \int_0^t P_i(s)\, ds$$

independently of the other events. Hence,

$$P\left\{N_i(t) = n_i, i = 1, \ldots, k \,\middle|\, N(t) = \sum_{i=1}^{k} n_i\right\}$$

will just equal the multinomial probability of n_i type i outcomes for $i = 1, \ldots, k$ when each of $\sum_{i=1}^{k} n_i$ independent trials results in outcome i with probability P_i, $i = 1, \ldots, k$. That is,

$$P\left\{N_1(t) = n_1, \ldots, N_k(t) = n_k \,\middle|\, N(t) = \sum_{i=1}^{k} n_i\right\} = \frac{(\sum_{i=1}^{k} n_i)!}{n_1! \cdots n_k!} P_1^{n_1} \cdots P_k^{n_k}$$

Consequently,

$$\begin{aligned} P\{N_1(t) = n_1, \ldots, N_k(t) = n_k\} &= \frac{(\sum_i n_i)!}{n_1! \cdots n_k!} P_1^{n_1} \cdots P_k^{n_k} e^{-\lambda t} \frac{(\lambda t)^{\sum_i n_i}}{(\sum_i n_i)!} \\ &= \prod_{i=1}^{k} e^{-\lambda t P_i} (\lambda t P_i)^{n_i} / n_i! \end{aligned}$$

and the proof is complete. ■

We now present some additional examples of the usefulness of Theorem 5.2.

Example 5.21 Insurance claims are made at times distributed according to a Poisson process with rate λ; the successive claim amounts are independent random variables having distribution G with mean μ, and are independent of the claim arrival times. Let S_i and C_i denote, respectively, the time and the amount of the ith claim. The total discounted cost of all claims made up to time t, call it $D(t)$, is defined by

$$D(t) = \sum_{i=1}^{N(t)} e^{-\alpha S_i} C_i$$

where α is the discount rate and $N(t)$ is the number of claims made by time t. To determine the expected value of $D(t)$, we condition on $N(t)$ to obtain

$$E[D(t)] = \sum_{n=0}^{\infty} E[D(t)|N(t) = n] e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

Now, conditional on $N(t) = n$, the claim arrival times $S_1, \ldots, S_n$ are distributed as the ordered values $U_{(1)}, \ldots, U_{(n)}$ of n independent uniform $(0, t)$ random variables $U_1, \ldots, U_n$. Therefore,

$$\begin{aligned}
E[D(t)|N(t) = n] &= E\left[\sum_{i=1}^{n} C_i e^{-\alpha U_{(i)}}\right] \\
&= \sum_{i=1}^{n} E[C_i e^{-\alpha U_{(i)}}] \\
&= \sum_{i=1}^{n} E[C_i] E[e^{-\alpha U_{(i)}}]
\end{aligned}$$

where the final equality used the independence of the claim amounts and their arrival times. Because $E[C_i] = \mu$, continuing the preceding gives

$$\begin{aligned}
E[D(t)|N(t) = n] &= \mu \sum_{i=1}^{n} E[e^{-\alpha U_{(i)}}] \\
&= \mu E\left[\sum_{i=1}^{n} e^{-\alpha U_{(i)}}\right] \\
&= \mu E\left[\sum_{i=1}^{n} e^{-\alpha U_i}\right]
\end{aligned}$$

The last equality follows because $U_{(1)}, \ldots, U_{(n)}$ are the values $U_1, \ldots, U_n$ in increasing order, and so $\sum_{i=1}^{n} e^{-\alpha U_{(i)}} = \sum_{i=1}^{n} e^{-\alpha U_i}$. Continuing the string of equalities yields

$$\begin{aligned}
E[D(t)|N(t) = n] &= n\mu E[e^{-\alpha U}] \\
&= n\frac{\mu}{t} \int_0^t e^{-\alpha x}\, dx \\
&= n\frac{\mu}{\alpha t}(1 - e^{-\alpha t})
\end{aligned}$$

Therefore,

$$E[D(t)|N(t)] = N(t)\frac{\mu}{\alpha t}(1 - e^{-\alpha t})$$

Taking expectations yields the result

$$E[D(t)] = \frac{\lambda\mu}{\alpha}(1 - e^{-\alpha t}) \qquad \blacksquare$$

Example 5.22 (An Optimization Example) Suppose that items arrive at a processing plant in accordance with a Poisson process with rate λ. At a fixed time T, all items are dispatched from the system. The problem is to choose an intermediate time, $t \in (0, T)$, at which all items in the system are dispatched, so as to minimize the total expected wait of all items.

If we dispatch at time t, $0 < t < T$, then the expected total wait of all items will be

$$\frac{\lambda t^2}{2} + \frac{\lambda (T-t)^2}{2}$$

To see why this is true, we reason as follows: The expected number of arrivals in $(0, t)$ is λt, and each arrival is uniformly distributed on $(0, t)$, and hence has expected wait $t/2$. Thus, the expected total wait of items arriving in $(0, t)$ is $\lambda t^2/2$. Similar reasoning holds for arrivals in (t, T), and the preceding follows. To minimize this quantity, we differentiate with respect to t to obtain

$$\frac{d}{dt}\left[\lambda \frac{t^2}{2} + \lambda \frac{(T-t)^2}{2}\right] = \lambda t - \lambda (T-t)$$

and equating to 0 shows that the dispatch time that minimizes the expected total wait is $t = T/2$. ■

We end this section with a result, quite similar in spirit to Theorem 5.2, which states that given S_n, the time of the nth event, then the first $n-1$ event times are distributed as the ordered values of a set of $n-1$ random variables uniformly distributed on $(0, S_n)$.

Proposition 5.4 Given that $S_n = t$, the set $S_1, \ldots, S_{n-1}$ has the distribution of a set of $n-1$ independent uniform $(0, t)$ random variables.

Proof. We can prove the preceding in the same manner as we did Theorem 5.2, or we can argue more loosely as follows:

$$\begin{aligned} S_1, \ldots, S_{n-1} \mid S_n = t &\sim S_1, \ldots, S_{n-1} \mid S_n = t, N(t^-) = n-1 \\ &\sim S_1, \ldots, S_{n-1} \mid N(t^-) = n-1 \end{aligned}$$

where $\sim$ means "has the same distribution as" and t^- is infinitesimally smaller than t. The result now follows from Theorem 5.2. ■

5.3.6 Estimating Software Reliability

When a new computer software package is developed, a testing procedure is often put into effect to eliminate the faults, or bugs, in the package. One common procedure is to try the package on a set of well-known problems to see if any errors result. This goes on for some fixed time, with all resulting errors being noted. Then the testing stops and the package is carefully checked to determine the specific bugs that were responsible for the observed errors. The package is then altered to remove these bugs. Because we cannot be certain that all the bugs in the package have been eliminated, however, a problem of great importance is the estimation of the error rate of the revised software package.

To model the preceding, let us suppose that initially the package contains an unknown number, m, of bugs, which we will refer to as bug 1, bug 2, ..., bug m. Suppose also that bug i will cause errors to occur in accordance with a Poisson process having an unknown rate $\lambda_i, i = 1, \ldots, m$. Then, for instance, the number of errors due to bug i that occurs in any s units of operating time is Poisson distributed with mean $\lambda_i s$. Also suppose that these Poisson processes caused by bugs $i, i = 1, \ldots, m$ are independent. In addition, suppose that the package is to be run for t time units with all resulting errors being noted. At the end of this time a careful check of the package is made to determine the specific bugs that caused the errors (that is, a *debugging*, takes place). These bugs are removed, and the problem is then to determine the error rate for the revised package.

If we let

$$\psi_i(t) = \begin{cases} 1, & \text{if bug } i \text{ has not caused an error by } t \\ 0, & \text{otherwise} \end{cases}$$

then the quantity we wish to estimate is

$$\Lambda(t) = \sum_i \lambda_i \psi_i(t)$$

the error rate of the final package. To start, note that

$$\begin{aligned} E[\Lambda(t)] &= \sum_i \lambda_i E[\psi_i(t)] \\ &= \sum_i \lambda_i e^{-\lambda_i t} \end{aligned} \tag{5.19}$$

Now each of the bugs that is discovered would have been responsible for a certain number of errors. Let us denote by $M_j(t)$ the number of bugs that were responsible for j errors, $j \geqslant 1$. That is, $M_1(t)$ is the number of bugs that caused exactly one error, $M_2(t)$ is the number that caused two errors, and so on, with $\sum_j jM_j(t)$ equaling the total number of errors that resulted. To compute $E[M_1(t)]$, let us define the indicator variables, $I_i(t), i \geqslant 1$, by

$$I_i(t) = \begin{cases} 1, & \text{bug } i \text{ causes exactly 1 error} \\ 0, & \text{otherwise} \end{cases}$$

Then,

$$M_1(t) = \sum_i I_i(t)$$

and so

$$E[M_1(t)] = \sum_i E[I_i(t)] = \sum_i \lambda_i t e^{-\lambda_i t} \tag{5.20}$$

Thus, from (5.19) and (5.20) we obtain the intriguing result that

$$E\left[\Lambda(t) - \frac{M_1(t)}{t}\right] = 0 \tag{5.21}$$

Thus suggests the possible use of $M_1(t)/t$ as an estimate of $\Lambda(t)$. To determine whether or not $M_1(t)/t$ constitutes a "good" estimate of $\Lambda(t)$ we shall look at how far apart these two quantities tend to be. That is, we will compute

$$\begin{aligned} E\left[\left(\Lambda(t) - \frac{M_1(t)}{t}\right)^2\right] &= \mathrm{Var}\left(\Lambda(t) - \frac{M_1(t)}{t}\right) \qquad \text{from (5.21)} \\ &= \mathrm{Var}(\Lambda(t)) - \frac{2}{t}\mathrm{Cov}(\Lambda(t), M_1(t)) + \frac{1}{t^2}\mathrm{Var}(M_1(t)) \end{aligned}$$

Now,

$$\begin{aligned} \mathrm{Var}(\Lambda(t)) &= \sum_i \lambda_i^2 \,\mathrm{Var}(\psi_i(t)) = \sum_i \lambda_i^2 e^{-\lambda_i t}(1 - e^{-\lambda_i t}), \\ \mathrm{Var}(M_1(t)) &= \sum_i \mathrm{Var}(I_i(t)) = \sum_i \lambda_i t e^{-\lambda_i t}(1 - \lambda_i t e^{-\lambda_i t}), \\ \mathrm{Cov}(\Lambda(t), M_1(t)) &= \mathrm{Cov}\left(\sum_i \lambda_i \psi_i(t),\ \sum_j I_j(t)\right) \\ &= \sum_i \sum_j \mathrm{Cov}(\lambda_i \psi_i(t), I_j(t)) \\ &= \sum_i \lambda_i \mathrm{Cov}(\psi_i(t), I_i(t)) \\ &= -\sum_i \lambda_i e^{-\lambda_i t} \lambda_i t e^{-\lambda_i t} \end{aligned}$$

where the last two equalities follow since $\psi_i(t)$ and $I_j(t)$ are independent when $i \neq j$ because they refer to different Poisson processes and $\psi_i(t)I_i(t) = 0$. Hence, we obtain

$$E\left[\left(\Lambda(t)-\frac{M_1(t)}{t}\right)^2\right]=\sum_i \lambda_i^2 e^{-\lambda_i t}+\frac{1}{t}\sum_i \lambda_i e^{-\lambda_i t}$$
$$=\frac{E[M_1(t)+2M_2(t)]}{t^2}$$

where the last equality follows from (5.20) and the identity (which we leave as an exercise)

$$E[M_2(t)]=\frac{1}{2}\sum_i (\lambda_i t)^2 e^{-\lambda_i t} \tag{5.22}$$

Thus, we can estimate the average square of the difference between $\Lambda(t)$ and $M_1(t)/t$ by the observed value of $M_1(t)+2M_2(t)$ divided by t^2.

Example 5.23 Suppose that in 100 units of operating time 20 bugs are discovered of which two resulted in exactly one, and three resulted in exactly two, errors. Then we would estimate that $\Lambda(100)$ is something akin to the value of a random variable whose mean is equal to $1/50$ and whose variance is equal to $8/10{,}000$. ■

5.4 Generalizations of the Poisson Process

5.4.1 Nonhomogeneous Poisson Process

In this section we consider two generalizations of the Poisson process. The first of these is the nonhomogeneous, also called the nonstationary, Poisson process, which is obtained by allowing the arrival rate at time t to be a function of t.

Definition 5.4 The counting process $\{N(t),\ t \geqslant 0\}$ is said to be a *nonhomogeneous Poisson process with intensity function* $\lambda(t),\ t \geqslant 0$, if

(i) $N(0)=0$.
(ii) $\{N(t), t \geqslant 0\}$ has independent increments.
(iii) $P\{N(t+h)-N(t) \geqslant 2\}=o(h)$.
(iv) $P\{N(t+h)-N(t)=1\}=\lambda(t)h+o(h)$.

Time sampling an ordinary Poisson process generates a nonhomogeneous Poisson process. That is, let $\{N(t), t \geqslant 0\}$ be a Poisson process with rate λ, and suppose that an event occurring at time t is, independently of what has occurred prior to t, counted with probability $p(t)$. With $N_c(t)$ denoting the number of counted events by time t, the counting process $\{N_c(t), t \geqslant 0\}$ is a nonhomogeneous Poisson process with intensity function $\lambda(t)=\lambda p(t)$. This is verified by noting that $\{N_c(t), t \geqslant 0\}$ satisfies the nonhomogeneous Poisson process axioms.

1. $N_c(0) = 0$.
2. The number of counted events in $(s, s + t)$ depends solely on the number of events of the Poisson process in $(s, s + t)$, which is independent of what has occurred prior to time s. Consequently, the number of counted events in $(s, s + t)$ is independent of the process of counted events prior to s, thus establishing the independent increment property.
3. Let $N_c(t, t + h) = N_c(t + h) - N_c(t)$, with a similar definition of $N(t, t + h)$.

$$P\{N_c(t, t + h) \geqslant 2\} \leqslant P\{N(t, t + h) \geqslant 2\} = o(h)$$

4. To compute $P\{N_c(t, t + h) = 1\}$, condition on $N(t, t + h)$.

$$\begin{aligned}
&P\{N_c(t, t + h) = 1\} \\
&\quad = P\{N_c(t, t + h) = 1|N(t, t + h) = 1\}P\{N(t, t + h) = 1\} \\
&\qquad + P\{N_c(t, t + h) = 1|N(t, t + h) \geqslant 2\}P\{N(t, t + h) \geqslant 2\} \\
&\quad = P\{N_c(t, t + h) = 1|N(t, t + h) = 1\}\lambda h + o(h) \\
&\quad = p(t)\lambda h + o(h)
\end{aligned}$$

Not only does time sampling a Poisson process result in a nonhomogeneous Poisson process, but it also works the other way: every nonhomogeneous Poisson process with a bounded intensity function can be thought of as being a time sampling of a Poisson process. To show this, we start by showing that the superposition of two independent nonhomogeneous Poisson processes remains a nonhomogeneous Poisson process.

Proposition 5.4 Let $\{N(t), t \geqslant 0\}$, and $\{M(t), t \geqslant 0\}$, be independent nonhomogeneous Poisson processes, with respective intensity functions $\lambda(t)$ and $\mu(t)$, and let $N^*(t) = N(t) + M(t)$. Then, the following are true.

(a) $\{N^*(t), t \geqslant 0\}$ is a nonhomogeneous Poisson process with intensity function $\lambda(t) + \mu(t)$.
(b) Given that an event of the $\{N^*(t)\}$ process occurs at time t then, independent of what occurred prior to t, the event at t was from the $\{N(t)\}$ process with probability $\frac{\lambda(t)}{\lambda(t)+\mu(t)}$.

Proof. To verify that $\{N^*(t), t \geqslant 0\}$ is a nonhomogeneous Poisson process with intensity function $\lambda(t) + \mu(t)$, we will argue that it satisfies the nonhomogeneous Poisson process axioms.

1. $N^*(0) = N(0) + M(0) = 0$.
2. To verify independent increments, let $I_1, \ldots, I_n$ be nonoverlapping intervals. Let $N(I)$ and $M(I)$ denote, respectively, the number of events from the $\{N(t)\}$ process and from the $\{M(t)\}$ process that are in the interval I. Because each counting process has independent increments, and the two processes are independent of each other, it follows that $N(I_1), \ldots, N(I_n), M(I_1), \ldots, M(I_n)$ are all independent, and thus so are $N(I_1) + M(I_1), \ldots, N(I_n) + M(I_n)$, which shows that $\{N^*(t), t \geqslant 0\}$ also possesses independent increments.
3. In order for there to be exactly one event of the N^* process between t and $t + h$, either there must be one event of the N process and 0 events of the M process or the

reverse. The first of these mutually exclusive possibilities occurs with probability

$$\begin{aligned} &P\{N(t,t+h)=1, M(t,t+h)=0\} \\ &\quad = P\{N(t,t+h)=1\}\, P\{M(t,t+h)=0\} \\ &\quad = (\lambda(t)h+o(h))\,(1-\mu(t)h+o(h)) \\ &\quad = \lambda(t)h+o(h) \end{aligned}$$

Similarly, the second possibility occurs with probability

$$P\{N(t,t+h)=0, M(t,t+h)=1\} = \mu(t)h + o(h)$$

yielding

$$P\{N^*(t+h)-N^*(t)=1\} = (\lambda(t)+\mu(t))h + o(h)$$

4. In order for there to be at least two events of the N^* process between t and $t+h$, one of the following three possibilities must occur: there is at least two events of the N process between t and $t+h$; there is at least two events of the M process between t and $t+h$; or both processes have exactly one event between t and $t+h$. Each of the first two of these possibilities occurs with probability $o(h)$, while the third occurs with probability $(\lambda(t)h+o(h))(\mu(t)h+o(h)) = o(h)$. Thus,

$$P\{N^*(t+h)-N^*(t)\geqslant 2\} \leqslant o(h)$$

Thus, part (a) is proven.

To prove (b), note first that it follows from independent increments that which process caused the event at time t is independent of what occurred prior to t. To find the conditional probability that the event at time t is from the N process, we use that

$$\begin{aligned} P\{N(t,t+h)=1|N^*(t,t+h)=1\} &= \frac{P\{N(t,t+h)=1, M(t,t+h)=0\}}{P\{N^*(t,t+h)=1\}} \\ &= \frac{\lambda(t)h+o(h)}{(\lambda(t)+\mu(t))h+o(h)} \\ &= \frac{\lambda(t)+\frac{o(h)}{h}}{\lambda(t)+\mu(t)+\frac{o(h)}{h}} \end{aligned}$$

Letting $h\to 0$ in the preceding proves (b). ■

Now, suppose that $\{N(t), t\geqslant 0\}$ is a nonhomogeneous Poisson process with a bounded intensity function $\lambda(t)$, and suppose that λ is such that $\lambda(t)\leqslant\lambda$, for all $t\geqslant 0$. Letting $\{M(t),\ t\geqslant 0\}$ be a nonhomogeneous Poisson process with intensity function $\mu(t)=\lambda-\lambda(t), t\geqslant 0$, that is independent of $\{N(t), t\geqslant 0\}$, it follows from Proposition 5.4 that $\{N(t), t\geqslant 0\}$ can be regarded as being the process of time-sampled events of the Poisson process $\{N(t)+M(t), t\geqslant 0\}$, where an event of the Poisson process that occurs at time t is counted with probability $p(t)=\lambda(t)/\lambda$.

With this interpretation of a nonhomogeneous Poisson process as being a time-sampled Poisson process, the number of events of the nonhomogeneous Poisson process by time t, namely, $N(t)$, is equal to the number of counted events of the Poisson process by time t. Consequently, from Proposition 5.3 it follows that $N(t)$ is a Poisson random variable with mean

$$E[N(t)] = \lambda \int_0^t \frac{\lambda(y)}{\lambda}\, dy = \int_0^t \lambda(y)\, dy$$

Moreover, by regarding the nonhomogeneous Poisson process as starting at time s, the preceding yields that $N(t+s) - N(t)$, the number of events in its first t time units, is a Poisson random variable with mean $\int_0^t \lambda(s+y)\, dy = \int_s^{s+t} \lambda(y)\, dy$.

The function $m(t)$ defined by

$$m(t) = \int_0^t \lambda(y)\, dy$$

is called the *mean value function* of the nonhomogeneous Poisson process.

Remark That $N(s+t) - N(s)$ has a Poisson distribution with mean $\int_s^{s+t} \lambda(y)\, dy$ is a consequence of the Poisson limit of the sum of independent Bernoulli random variables (see Example 2.47). To see why, subdivide the interval $[s, s+t]$ into n subintervals of length $\frac{t}{n}$, where subinterval i goes from $s + (i-1)\frac{t}{n}$ to $s + i\frac{t}{n}$, $i = 1, \dots, n$. Let $N_i = N(s + i\frac{t}{n}) - N(s + (i-1)\frac{t}{n})$ be the number of events that occur in subinterval i, and note that

$$\begin{aligned} P\{\geqslant 2 \text{ events in some subinterval}\} &= P\left(\bigcup_{i=1}^{n} \{N_i \geqslant 2\}\right) \\ &\leqslant \sum_{i=1}^{n} P\{N_i \geqslant 2\} \\ &= n o(t/n) \qquad \text{by Axiom } (iii) \end{aligned}$$

Because

$$\lim_{n\to\infty} n o(t/n) = \lim_{n\to\infty} t \frac{o(t/n)}{t/n} = 0$$

it follows that, as n increases to ∞, the probability of having two or more events in any of the n subintervals goes to 0. Consequently, with a probability going to 1, $N(t)$ will equal the number of subintervals in which an event occurs. Because the probability of an event in subinterval i is $\lambda(s + i\frac{t}{n})\frac{t}{n} + o(\frac{t}{n})$, it follows, because the number of events in different subintervals are independent, that when n is large the number of subintervals that contain an event is approximately a Poisson random variable with mean

$$\sum_{i=1}^{n} \lambda\left(s + i\frac{t}{n}\right)\frac{t}{n} + no(t/n)$$

But,

$$\lim_{n\to\infty} \sum_{i=1}^{n} \lambda\left(s + i\frac{t}{n}\right)\frac{t}{n} + no(t/n) = \int_{s}^{s+t} \lambda(y)\,dy$$

and the result follows. ■

The importance of the nonhomogeneous Poisson process resides in the fact that we no longer require the condition of stationary increments. Thus we now allow for the possibility that events may be more likely to occur at certain times than during other times.

Example 5.24 Siegbert runs a hot dog stand that opens at 8 A.M. From 8 until 11 A.M. customers seem to arrive, on the average, at a steadily increasing rate that starts with an initial rate of 5 customers per hour at 8 A.M. and reaches a maximum of 20 customers per hour at 11 A.M. From 11 A.M. until 1 P.M. the (average) rate seems to remain constant at 20 customers per hour. However, the (average) arrival rate then drops steadily from 1 P.M. until closing time at 5 P.M. at which time it has the value of 12 customers per hour. If we assume that the numbers of customers arriving at Siegbert's stand during disjoint time periods are independent, then what is a good probability model for the preceding? What is the probability that no customers arrive between 8:30 A.M. and 9:30 A.M. on Monday morning? What is the expected number of arrivals in this period?

Solution: A good model for the preceding would be to assume that arrivals constitute a nonhomogeneous Poisson process with intensity function $\lambda(t)$ given by

$$\lambda(t) = \begin{cases} 5 + 5t, & 0 \leqslant t \leqslant 3 \\ 20, & 3 \leqslant t \leqslant 5 \\ 20 - 2(t-5), & 5 \leqslant t \leqslant 9 \end{cases}$$

and

$$\lambda(t) = \lambda(t-9) \qquad \text{for } t > 9$$

Note that $N(t)$ represents the number of arrivals during the first t hours that the store is open. That is, we do not count the hours between 5 P.M. and 8 A.M. If for some reason we wanted $N(t)$ to represent the number of arrivals during the first t hours regardless of whether the store was open or not, then, assuming that the process begins at midnight we would let

$$\lambda(t) = \begin{cases} 0, & 0 \leqslant t \leqslant 8 \\ 5 + 5(t-8), & 8 \leqslant t \leqslant 11 \\ 20, & 11 \leqslant t \leqslant 13 \\ 20 - 2(t-13), & 13 \leqslant t \leqslant 17 \\ 0, & 17 < t \leqslant 24 \end{cases}$$

and

$$\lambda(t) = \lambda(t-24) \qquad \text{for } t > 24$$

As the number of arrivals between 8:30 A.M. and 9:30 A.M. will be Poisson with mean $m(\frac{3}{2}) - m(\frac{1}{2})$ in the first representation (and $m(\frac{19}{2}) - m(\frac{17}{2})$ in the second representation), we have that the probability that this number is zero is

$$\exp\left\{-\int_{1/2}^{3/2} (5+5t)\,dt\right\} = e^{-10}$$

and the mean number of arrivals is

$$\int_{1/2}^{3/2} (5+5t)\,dt = 10$$ ■

Suppose that events occur according to a Poisson process with rate λ, and suppose that, independent of what has previously occurred, an event at time s is a type 1 event with probability $P_1(s)$ or a type 2 event with probability $P_2(s) = 1 - P_1(s)$. If $N_i(t), t \geqslant 0$, denotes the number of type i events by time t, then it easily follows from Definition 5.4 that $\{N_1(t), t \geqslant 0\}$ and $\{N_2(t), t \geqslant 0\}$ are independent nonhomogeneous Poisson processes with respective intensity functions $\lambda_i(t) = \lambda P_i(t), i = 1, 2$. (The proof mimics that of Proposition 5.2.) This result gives us another way of understanding (or of proving) the time sampling Poisson process result of Proposition 5.3, which states that $N_1(t)$ and $N_2(t)$ are independent Poisson random variables with means $E[N_i(t)] = \lambda \int_0^t P_i(s)\,ds$, $i = 1, 2$.

Example 5.25 (The Output Process of an Infinite Server Poisson Queue) It turns out that the output process of the $M/G/\infty$ queue—that is, of the infinite server queue having Poisson arrivals and general service distribution G—is a nonhomogeneous Poisson process having intensity function $\lambda(t) = \lambda G(t)$. To verify this claim, let us first argue that the departure process has independent increments. Towards this end, consider nonoverlapping intervals $O_1, \ldots, O_k$; now say that an arrival is type $i, i = 1, \ldots, k$, if that arrival departs in the interval O_i.

By Proposition 5.3, it follows that the numbers of departures in these intervals are independent, thus establishing independent increments. Now, suppose that an arrival is "counted" if that arrival departs between t and $t + h$. Because an arrival at time $s, s < t + h$, will be counted with probability $G(t-s + h)-G(t-s)$, it follows from Proposition 5.3 that the number of departures in $(t, t + h)$ is a Poisson random variable with mean

$$\begin{aligned}
\lambda \int_0^{t+h} [G(t - s + h) - G(t - s)]\, ds &= \lambda \int_0^{t+h} [G'(t - s + h)h + o(h)]\, ds \\
&= \lambda h \int_0^{t+h} G'(y)\, dy + o(h) \\
&= \lambda G(t)h + o(h)
\end{aligned}$$

Therefore,

$$P\{1 \text{ departure in } (t, t + h)\} = \lambda G(t)h\, e^{-\lambda G(t)h} + o(h) = \lambda G(t)h + o(h)$$

and

$$P\{\geqslant 2 \text{ departures in } (t, t + h)\} = o(h)$$

which completes the verification. ■

If we let S_n denote the time of the nth event of the nonhomogeneous Poisson process, then we can obtain its density as follows:

$$\begin{aligned}
P\{t < S_n < t + h\} &= P\{N(t) = n - 1,\ \text{one event in } (t, t + h)\} + o(h) \\
&= P\{N(t) = n - 1\}P\{\text{one event in } (t, t + h)\} + o(h) \\
&= e^{-m(t)} \frac{[m(t)]^{n-1}}{(n - 1)!} [\lambda(t)h + o(h)] + o(h) \\
&= \lambda(t)e^{-m(t)} \frac{[m(t)]^{n-1}}{(n - 1)!} h + o(h)
\end{aligned}$$

which implies that

$$f_{S_n}(t) = \lambda(t)e^{-m(t)} \frac{[m(t)]^{n-1}}{(n - 1)!}$$

where

$$m(t) = \int_0^t \lambda(s)\, ds$$

5.4.2 Compound Poisson Process

A stochastic process $\{X(t), t \geqslant 0\}$ is said to be a *compound Poisson process* if it can be represented as

$$X(t) = \sum_{i=1}^{N(t)} Y_i, \qquad t \geqslant 0 \tag{5.23}$$

where $\{N(t), t \geqslant 0\}$ is a Poisson process, and $\{Y_i, i \geqslant 1\}$ is a family of independent and identically distributed random variables that is also independent of $\{N(t), t \geqslant 0\}$. As noted in Chapter 3, the random variable $X(t)$ is said to be a compound Poisson random variable.

Examples of Compound Poisson Processes

(i) If $Y_i \equiv 1$, then $X(t) = N(t)$, and so we have the usual Poisson process.
(ii) Suppose that buses arrive at a sporting event in accordance with a Poisson process, and suppose that the numbers of fans in each bus are assumed to be independent and identically distributed. Then $\{X(t), t \geqslant 0\}$ is a compound Poisson process where $X(t)$ denotes the number of fans who have arrived by t. In Equation (5.23) Y_i represents the number of fans in the ith bus.
(iii) Suppose customers leave a supermarket in accordance with a Poisson process. If the Y_i, the amount spent by the ith customer, $i = 1, 2, \ldots$, are independent and identically distributed, then $\{X(t), t \geqslant 0\}$ is a compound Poisson process when $X(t)$ denotes the total amount of money spent by time t. ■

Because $X(t)$ is a compound Poisson random variable with Poisson parameter λt, we have from Examples 3.10 and 3.17 that

$$E[X(t)] = \lambda t E[Y_1] \tag{5.24}$$

and

$$\mathrm{Var}(X(t)) = \lambda t E\left[Y_1^2\right] \tag{5.25}$$

Example 5.26 Suppose that families migrate to an area at a Poisson rate $\lambda = 2$ per week. If the number of people in each family is independent and takes on the values 1, 2, 3, 4 with respective probabilities $\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}$, then what is the expected value and variance of the number of individuals migrating to this area during a fixed five-week period?

Solution: Letting Y_i denote the number of people in the ith family, we have

$$E[Y_i] = 1 \cdot \tfrac{1}{6} + 2 \cdot \tfrac{1}{3} + 3 \cdot \tfrac{1}{3} + 4 \cdot \tfrac{1}{6} = \tfrac{5}{2},$$

$$E\left[Y_i^2\right] = 1^2 \cdot \tfrac{1}{6} + 2^2 \cdot \tfrac{1}{3} + 3^2 \cdot \tfrac{1}{3} + 4^2 \cdot \tfrac{1}{6} = \tfrac{43}{6}$$

Hence, letting $X(5)$ denote the number of immigrants during a five-week period, we obtain from Equations (5.24) and (5.25) that

$$E[X(5)] = 2 \cdot 5 \cdot \tfrac{5}{2} = 25$$

and

$$\text{Var}[X(5)] = 2 \cdot 5 \cdot \tfrac{43}{6} = \tfrac{215}{3}$$ ■

Example 5.27 (Busy Periods in Single-Server Poisson Arrival Queues) Consider a single-server service station in which customers arrive according to a Poisson process having rate λ. An arriving customer is immediately served if the server is free; if not, the customer waits in line (that is, he or she joins the queue). The successive service times are independent with a common distribution.

Such a system will alternate between idle periods when there are no customers in the system, so the server is idle, and busy periods when there are customers in the system, so the server is busy. A busy period will begin when an arrival finds the system empty, and because of the memoryless property of the Poisson arrivals it follows that the distribution of the length of a busy period will be the same for each such period. Let B denote the length of a busy period. We will compute its mean and variance.

To begin, let S denote the service time of the first customer in the busy period and let $N(S)$ denote the number of arrivals during that time. Now, if $N(S) = 0$ then the busy period will end when the initial customer completes his service, and so B will equal S in this case. Now, suppose that one customer arrives during the service time of the initial customer. Then, at time S there will be a single customer in the system who is just about to enter service. Because the arrival stream from time S on will still be a Poisson process with rate λ, it thus follows that the additional time from S until the system becomes empty will have the same distribution as a busy period. That is, if $N(S) = 1$ then

$$B = S + B_1$$

where B_1 is independent of S and has the same distribution as B.

Now, consider the general case where $N(S) = n$, so there will be n customers waiting when the server finishes his initial service. To determine the distribution of remaining time in the busy period note that the order in which customers are served will not affect the remaining time. Hence, let us suppose that the n arrivals, call them $C_1, \ldots, C_n$, during the initial service period are served as follows. Customer C_1 is served first, but C_2 is not served until the only customers in the system are $C_2, \ldots, C_n$. For instance, any customers arriving during C_1's service time will be served before C_2. Similarly, C_3 is not served until the system is free of all customers but $C_3, \ldots, C_n$, and so on. A little thought reveals that the times between the beginnings of service of customers C_i and C_{i+1}, $i = 1, \ldots, n-1$, and the time from the beginning of service of C_n until there are no customers in the system, are independent random variables, each distributed as a busy period.

It follows from the preceding that if we let $B_1, B_2, \ldots$ be a sequence of independent random variables, each distributed as a busy period, then we can express B as

$$B = S + \sum_{i=1}^{N(S)} B_i$$

Hence,

$$E[B|S] = S + E\left[\sum_{i=1}^{N(S)} B_i|S\right]$$

and

$$\mathrm{Var}(B|S) = \mathrm{Var}\left(\sum_{i=1}^{N(S)} B_i|S\right)$$

However, given $S, \sum_{i=1}^{N(S)} B_i$ is a compound Poisson random variable, and thus from Equations (5.24) and (5.25) we obtain

$$E[B|S] = S + \lambda S E[B] = (1 + \lambda E[B])S$$
$$\mathrm{Var}(B|S) = \lambda S E[B^2]$$

Hence,

$$E[B] = E[E[B|S]] = (1 + \lambda E[B])E[S]$$

implying, provided that $\lambda E[S] < 1$, that

$$E[B] = \frac{E[S]}{1 - \lambda E[S]}$$

Also, by the conditional variance formula

$$\begin{aligned}\mathrm{Var}(B) &= \mathrm{Var}(E[B|S]) + E[\mathrm{Var}(B|S)]\\ &= (1 + \lambda E[B])^2 \mathrm{Var}(S) + \lambda E[S]E[B^2]\\ &= (1 + \lambda E[B])^2 \mathrm{Var}(S) + \lambda E[S](\mathrm{Var}(B) + E^2[B])\end{aligned}$$

yielding

$$\mathrm{Var}(B) = \frac{\mathrm{Var}(S)(1 + \lambda E[B])^2 + \lambda E[S]E^2[B]}{1 - \lambda E[S]}$$

Using $E[B] = E[S]/(1 - \lambda E[S])$, we obtain

$$\mathrm{Var}(B) = \frac{\mathrm{Var}(S) + \lambda E^3[S]}{(1 - \lambda E[S])^3} \qquad \blacksquare$$

There is a very nice representation of the compound Poisson process when the set of possible values of the Y_i is finite or countably infinite. So let us suppose that there are numbers $\alpha_j, j \geqslant 1$, such that

$$P\{Y_i = \alpha_j\} = p_j, \quad \sum_j p_j = 1$$

Now, a compound Poisson process arises when events occur according to a Poisson process and each event results in a random amount Y being added to the cumulative sum. Let us say that the event is a type j event whenever it results in adding the amount $\alpha_j, j \geqslant 1$. That is, the ith event of the Poisson process is a type j event if $Y_i = \alpha_j$. If we let $N_j(t)$ denote the number of type j events by time t, then it follows from Proposition 5.2 that the random variables $N_j(t)$, $j \geqslant 1$, are independent Poisson random variables with respective means

$$E[N_j(t)] = \lambda p_j t$$

Since, for each j, the amount α_j is added to the cumulative sum a total of $N_j(t)$ times by time t, it follows that the cumulative sum at time t can be expressed as

$$X(t) = \sum_j \alpha_j N_j(t) \tag{5.26}$$

As a check of Equation (5.26), let us use it to compute the mean and variance of $X(t)$. This yields

$$\begin{aligned} E[X(t)] &= E\left[\sum_j \alpha_j N_j(t)\right] \\ &= \sum_j \alpha_j E[N_j(t)] \\ &= \sum_j \alpha_j \lambda p_j t \\ &= \lambda t\, E[Y_1] \end{aligned}$$

Also,

$$\begin{aligned}
\mathrm{Var}[X(t)] &= \mathrm{Var}\left[\sum_j \alpha_j N_j(t)\right] \\
&= \sum_j \alpha_j^2 \,\mathrm{Var}[N_j(t)] \quad \text{by the independence of the } N_j(t),\ j \geqslant 1 \\
&= \sum_j \alpha_j^2 \,\lambda p_j t \\
&= \lambda t E[Y_1^2]
\end{aligned}$$

where the next to last equality follows since the variance of the Poisson random variable $N_j(t)$ is equal to its mean.

Thus, we see that the representation (5.26) results in the same expressions for the mean and variance of $X(t)$ as were previously derived.

One of the uses of the representation (5.26) is that it enables us to conclude that as t grows large, the distribution of $X(t)$ converges to the normal distribution. To see why, note first that it follows by the central limit theorem that the distribution of a Poisson random variable converges to a normal distribution as its mean increases. (Why is this?) Therefore, each of the random variables $N_j(t)$ converges to a normal random variable as t increases. Because they are independent, and because the sum of independent normal random variables is also normal, it follows that $X(t)$ also approaches a normal distribution as t increases.

Example 5.28 In Example 5.26, find the approximate probability that at least 240 people migrate to the area within the next 50 weeks.

Solution: Since $\lambda = 2, E[Y_i] = 5/2, E[Y_i^2] = 43/6$, we see that

$$E[X(50)] = 250, \qquad \mathrm{Var}[X(50)] = 4300/6$$

Now, the desired probability is

$$\begin{aligned}
P\{X(50) \geqslant 240\} &= P\{X(50) \geqslant 239.5\} \\
&= P\left\{\frac{X(50) - 250}{\sqrt{4300/6}} \geqslant \frac{239.5 - 250}{\sqrt{4300/6}}\right\} \\
&= 1 - \phi(-0.3922) \\
&= \phi(0.3922) \\
&= 0.6525
\end{aligned}$$

where Table 2.3 was used to determine $\phi(0.3922)$, the probability that a standard normal is less than 0.3922. ■

Another useful result is that if $\{X(t), t \geqslant 0\}$ and $\{Y(t), t \geqslant 0\}$ are independent compound Poisson processes with respective Poisson parameters and distributions λ_1, F_1 and λ_2, F_2, then $\{X(t) + Y(t), t \geqslant 0\}$ is also a compound Poisson process. This is true because in this combined process events will occur according to a Poisson process with rate $\lambda_1 + \lambda_2$, and each event independently will be from the first compound Poisson process with probability $\lambda_1/(\lambda_1 + \lambda_2)$. Consequently, the combined process will be a compound Poisson process with Poisson parameter $\lambda_1 + \lambda_2$, and with distribution function F given by

$$F(x) = \frac{\lambda_1}{\lambda_1 + \lambda_2} F_1(x) + \frac{\lambda_2}{\lambda_1 + \lambda_2} F_2(x)$$

5.4.3 Conditional or Mixed Poisson Processes

Let $\{N(t), t \geqslant 0\}$ be a counting process whose probabilities are defined as follows. There is a positive random variable L such that, conditional on $L = \lambda$, the counting process is a Poisson process with rate λ. Such a counting process is called a *conditional* or a *mixed* Poisson process.

Suppose that L is continuous with density function g. Because

$$\begin{aligned} P\{N(t+s) - N(s) = n\} &= \int_0^\infty P\{N(t+s) - N(s) = n \mid L = \lambda\} g(\lambda)\, d\lambda \\ &= \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} g(\lambda)\, d\lambda \end{aligned} \tag{5.27}$$

we see that a conditional Poisson process has stationary increments. However, because knowing how many events occur in an interval gives information about the possible value of L, which affects the distribution of the number of events in any other interval, it follows that a conditional Poisson process does not generally have independent increments. Consequently, a conditional Poisson process is not generally a Poisson process.

Example 5.29 If g is the gamma density with parameters m and θ,

$$g(\lambda) = \theta e^{-\theta\lambda} \frac{(\theta\lambda)^{m-1}}{(m-1)!}, \qquad \lambda > 0$$

then

$$\begin{aligned} P\{N(t) = n\} &= \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} \theta e^{-\theta\lambda} \frac{(\theta\lambda)^{m-1}}{(m-1)!}\, d\lambda \\ &= \frac{t^n \theta^m}{n!(m-1)!} \int_0^\infty e^{-(t+\theta)\lambda} \lambda^{n+m-1}\, d\lambda \end{aligned}$$

Multiplying and dividing by $\frac{(n+m-1)!}{(t+\theta)^{n+m}}$ gives

$$P\{N(t)=n\} = \frac{t^n \theta^m (n+m-1)!}{n!(m-1)!(t+\theta)^{n+m}} \int_0^\infty (t+\theta)e^{-(t+\theta)\lambda} \frac{((t+\theta)\lambda)^{n+m-1}}{(n+m-1)!}\,d\lambda$$

Because $(t+\theta)e^{-(t+\theta)\lambda}((t+\theta)\lambda)^{n+m-1}/(n+m-1)!$ is the density function of a gamma $(n+m, t+\theta)$ random variable, its integral is 1, giving the result

$$P\{N(t)=n\} = \binom{n+m-1}{n}\left(\frac{\theta}{t+\theta}\right)^m \left(\frac{t}{t+\theta}\right)^n$$

Therefore, the number of events in an interval of length t has the same distribution of the number of failures that occur before a total of m successes are amassed, when each trial is a success with probability $\frac{\theta}{t+\theta}$. ■

To compute the mean and variance of $N(t)$, condition on L. Because, conditional on L, $N(t)$ is Poisson with mean Lt, we obtain

$$\begin{aligned} E[N(t)|L] &= Lt \\ \mathrm{Var}(N(t)|L) &= Lt \end{aligned}$$

where the final equality used that the variance of a Poisson random variable is equal to its mean. Consequently, the conditional variance formula yields

$$\begin{aligned} \mathrm{Var}(N(t)) &= E[Lt] + \mathrm{Var}(Lt) \\ &= tE[L] + t^2\mathrm{Var}(L) \end{aligned}$$

We can compute the conditional distribution function of L, given that $N(t)=n$, as follows.

$$\begin{aligned} P\{L \leqslant x | N(t)=n\} &= \frac{P\{L \leqslant x, N(t)=n\}}{P\{N(t)=n\}} \\ &= \frac{\int_0^\infty P\{L \leqslant x, N(t)=n | L=\lambda\} g(\lambda)\,d\lambda}{P\{N(t)=n\}} \\ &= \frac{\int_0^x P\{N(t)=n | L=\lambda\} g(\lambda)\,d\lambda}{P\{N(t)=n\}} \\ &= \frac{\int_0^x e^{-\lambda t}(\lambda t)^n g(\lambda)\,d\lambda}{\int_0^\infty e^{-\lambda t}(\lambda t)^n g(\lambda)\,d\lambda} \end{aligned}$$

where the final equality used Equation (5.27). In other words, the conditional density function of L given that $N(t) = n$ is

$$f_{L|N(t)}(\lambda \mid n) = \frac{e^{-\lambda t}\lambda^n g(\lambda)}{\int_0^\infty e^{-\lambda t}\lambda^n g(\lambda)\,d\lambda}, \quad \lambda \geqslant 0 \tag{5.28}$$

Example 5.30 An insurance company feels that each of its policyholders has a rating value and that a policyholder having rating value λ will make claims at times distributed according to a Poisson process with rate λ, when time is measured in years. The firm also believes that rating values vary from policyholder to policyholder, with the probability distribution of the value of a new policyholder being uniformly distributed over $(0, 1)$. Given that a policyholder has made n claims in his or her first t years, what is the conditional distribution of the time until the policyholder's next claim?

Solution: If T is the time until the next claim, then we want to compute $P\{T > x \mid N(t) = n\}$. Conditioning on the policyholder's rating value gives, upon using Equation (5.28),

$$\begin{aligned} P\{T > x \mid N(t) = n\} &= \int_0^\infty P\{T > x \mid L = \lambda, N(t) = n\} f_{L|N(t)}(\lambda \mid n)\,d\lambda \\ &= \frac{\int_0^1 e^{-\lambda x} e^{-\lambda t}\lambda^n\,d\lambda}{\int_0^1 e^{-\lambda t}\lambda^n\,d\lambda} \end{aligned}$$

■

There is a nice formula for the probability that more than n events occur in an interval of length t. In deriving it we will use the identity

$$\sum_{j=n+1}^{\infty} e^{-\lambda t}\frac{(\lambda t)^j}{j!} = \int_0^t \lambda e^{-\lambda x}\frac{(\lambda x)^n}{n!}\,dx \tag{5.29}$$

which follows by noting that it equates the probability that the number of events by time t of a Poisson process with rate λ is greater than n with the probability that the time of the $(n + 1)$st event of this process (which has a gamma $(n + 1, \lambda)$ distribution) is less than t. Interchanging λ and t in Equation (5.29) yields the equivalent identity

$$\sum_{j=n+1}^{\infty} e^{-\lambda t}\frac{(\lambda t)^j}{j!} = \int_0^\lambda t e^{-tx}\frac{(tx)^n}{n!}\,dx \tag{5.30}$$

Using Equation (5.27) we now have

$$P\{N(t) > n\} = \sum_{j=n+1}^{\infty} \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!} g(\lambda)\, d\lambda$$

$$= \int_0^{\infty} \sum_{j=n+1}^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!} g(\lambda)\, d\lambda \qquad \text{(by interchanging)}$$

$$= \int_0^{\infty} \int_0^{\lambda} t e^{-tx} \frac{(tx)^n}{n!}\, dx g(\lambda)\, d\lambda \qquad \text{(using (5.30))}$$

$$= \int_0^{\infty} \int_x^{\infty} g(\lambda)\, d\lambda t e^{-tx} \frac{(tx)^n}{n!}\, dx \qquad \text{(by interchanging)}$$

$$= \int_0^{\infty} \bar{G}(x) t e^{-tx} \frac{(tx)^n}{n!}\, dx$$

Exercises

1. The time T required to repair a machine is an exponentially distributed random variable with mean $\frac{1}{2}$ (hours).
   (a) What is the probability that a repair time exceeds $\frac{1}{2}$ hour?
   (b) What is the probability that a repair takes at least $12\frac{1}{2}$ hours given that its duration exceeds 12 hours?
2. Suppose that you arrive at a single-teller bank to find five other customers in the bank, one being served and the other four waiting in line. You join the end of the line. If the service times are all exponential with rate μ, what is the expected amount of time you will spend in the bank?
3. Let X be an exponential random variable. Without any computations, tell which one of the following is correct. Explain your answer.
   (a) $E[X^2|X > 1] = E[(X + 1)^2]$
   (b) $E[X^2|X > 1] = E[X^2] + 1$
   (c) $E[X^2|X > 1] = (1 + E[X])^2$
4. Consider a post office with two clerks. Three people, A, B, and C, enter simultaneously. A and B go directly to the clerks, and C waits until either A or B leaves before he begins service. What is the probability that A is still in the post office after the other two have left when
   (a) the service time for each clerk is exactly (nonrandom) ten minutes?
   (b) the service times are i with probability $\frac{1}{3}$, $i = 1, 2, 3$?
   (c) the service times are exponential with mean $1/\mu$?
5. The lifetime of a radio is exponentially distributed with a mean of ten years. If Jones buys a ten-year-old radio, what is the probability that it will be working after an additional ten years?

6. In Example 5.3 if server i serves at an exponential rate λ_i, $i = 1, 2$, show that

$$P\{\text{Smith is not last}\} = \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^2 + \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^2$$

*7. If X_1 and X_2 are independent nonnegative continuous random variables, show that

$$P\{X_1 < X_2 | \min(X_1, X_2) = t\} = \frac{r_1(t)}{r_1(t) + r_2(t)}$$

where $r_i(t)$ is the failure rate function of X_i.

8. Let $X_i, i = 1, \ldots, n$ be independent continuous random variables, with X_i having failure rate function $r_i(t)$. Let T be independent of this sequence, and suppose that $\sum_{i=1}^{n} P\{T = i\} = 1$. Show that the failure rate function $r(t)$ of X_T is given by

$$r(t) = \sum_{i=1}^{n} r_i(t) P\{T = i | X > t\}$$

9. Machine 1 is currently working. Machine 2 will be put in use at a time t from now. If the lifetime of machine i is exponential with rate λ_i, $i = 1, 2$, what is the probability that machine 1 is the first machine to fail?

*10. Let X and Y be independent exponential random variables with respective rates λ and μ. Let $M = \min(X, Y)$. Find
   (a) $E[MX|M = X]$,
   (b) $E[MX|M = Y]$,
   (c) $\text{Cov}(X, M)$.

11. Let X, $Y_1, \ldots, Y_n$ be independent exponential random variables; X having rate λ, and Y_i having rate μ. Let A_j be the event that the jth smallest of these $n + 1$ random variables is one of the Y_i. Find $p = P\{X > \max_i Y_i\}$, by using the identity

$$p = P(A_1 \cdots A_n) = P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 \cdots A_{n-1})$$

Verify your answer when $n = 2$ by conditioning on X to obtain p.

12. If X_i, $i = 1, 2, 3$, are independent exponential random variables with rates λ_i, $i = 1, 2, 3$, find
   (a) $P\{X_1 < X_2 < X_3\}$,
   (b) $P\{X_1 < X_2 | \max(X_1, X_2, X_3) = X_3\}$,
   (c) $E[\max X_i | X_1 < X_2 < X_3]$,
   (d) $E[\max X_i]$.

13. Find, in Example 5.10, the expected time until the nth person on line leaves the line (either by entering service or departing without service).

14. Let X be an exponential random variable with rate λ.
   (a) Use the definition of conditional expectation to determine $E[X|X < c]$.
   (b) Now determine $E[X|X < c]$ by using the following identity:

$$E[X] = E[X|X < c]P\{X < c\} + E[X|X > c]P\{X > c\}$$

15. One hundred items are simultaneously put on a life test. Suppose the lifetimes of the individual items are independent exponential random variables with mean 200 hours. The test will end when there have been a total of 5 failures. If T is the time at which the test ends, find $E[T]$ and $\text{Var}(T)$.

16. There are three jobs that need to be processed, with the processing time of job i being exponential with rate μ_i. There are two processors available, so processing on two of the jobs can immediately start, with processing on the final job to start when one of the initial ones is finished.
    (a) Let T_i denote the time at which the processing of job i is completed. If the objective is to minimize $E[T_1 + T_2 + T_3]$, which jobs should be initially processed if $\mu_1 < \mu_2 < \mu_3$?
    (b) Let M, called the *makespan*, be the time until all three jobs have been processed. With S equal to the time that there is only a single processor working, show that

$$2E[M] = E[S] + \sum_{i=1}^{3} 1/\mu_i$$

    For the rest of this problem, suppose that $\mu_1 = \mu_2 = \mu,\quad \mu_3 = \lambda$. Also, let $P(\mu)$ be the probability that the last job to finish is either job 1 or job 2, and let $P(\lambda) = 1 - P(\mu)$ be the probability that the last job to finish is job 3.
    (c) Express $E[S]$ in terms of $P(\mu)$ and $P(\lambda)$.
    Let $P_{i,j}(\mu)$ be the value of $P(\mu)$ when i and j are the jobs that are initially started.
    (d) Show that $P_{1,2}(\mu) \leqslant P_{1,3}(\mu)$.
    (e) If $\mu > \lambda$ show that $E[M]$ is minimized when job 3 is one of the jobs that is initially started.
    (f) If $\mu < \lambda$ show that $E[M]$ is minimized when processing is initially started on jobs 1 and 2.

17. A set of n cities is to be connected via communication links. The cost to construct a link between cities i and j is C_{ij}, $i \neq j$. Enough links should be constructed so that for each pair of cities there is a path of links that connects them. As a result, only $n - 1$ links need be constructed. A minimal cost algorithm for solving this problem (known as the minimal spanning tree problem) first constructs the cheapest of all the $\binom{n}{2}$ links. Then, at each additional stage it chooses the cheapest link that connects a city without any links to one with links. That is, if the first link is between cities 1 and 2, then the second link will either be between 1 and one of the links $3, \ldots, n$ or between 2 and one of the links $3, \ldots, n$. Suppose that all of the $\binom{n}{2}$ costs C_{ij} are independent exponential random variables with mean 1. Find the expected cost of the preceding algorithm if
    (a) $n = 3$,
    (b) $n = 4$.

*18. Let X_1 and X_2 be independent exponential random variables, each having rate μ. Let

$$X_{(1)} = \text{minimum}(X_1, X_2) \quad \text{and} \quad X_{(2)} = \text{maximum}(X_1, X_2)$$

Find
(a) $E[X_{(1)}]$,
(b) $\text{Var}[X_{(1)}]$,
(c) $E[X_{(2)}]$,
(d) $\text{Var}[X_{(2)}]$.

19. Repeat Exercise 18, but this time suppose that the X_i are independent exponentials with respective rates μ_i, $i = 1, 2$.

20. Consider a two-server system in which a customer is served first by server 1, then by server 2, and then departs. The service times at server i are exponential random variables with rates μ_i, $i = 1, 2$. When you arrive, you find server 1 free and two customers at server 2—customer A in service and customer B waiting in line.
(a) Find P_A, the probability that A is still in service when you move over to server 2.
(b) Find P_B, the probability that B is still in the system when you move over to server 2.
(c) Find $E[T]$, where T is the time that you spend in the system.

Hint: Write

$$T = S_1 + S_2 + W_A + W_B$$

where S_i is your service time at server i, W_A is the amount of time you wait in queue while A is being served, and W_B is the amount of time you wait in queue while B is being served.

21. In a certain system, a customer must first be served by server 1 and then by server 2. The service times at server i are exponential with rate μ_i, $i = 1, 2$. An arrival finding server 1 busy waits in line for that server. Upon completion of service at server 1, a customer either enters service with server 2 if that server is free or else remains with server 1 (blocking any other customer from entering service) until server 2 is free. Customers depart the system after being served by server 2. Suppose that when you arrive there is one customer in the system and that customer is being served by server 1. What is the expected total time you spend in the system?

22. Suppose in Exercise 21 you arrive to find two others in the system, one being served by server 1 and one by server 2. What is the expected time you spend in the system? Recall that if server 1 finishes before server 2, then server 1's customer will remain with him (thus blocking your entrance) until server 2 becomes free.

*23. A flashlight needs two batteries to be operational. Consider such a flashlight along with a set of n functional batteries—battery 1, battery 2, …, battery n. Initially, battery 1 and 2 are installed. Whenever a battery fails, it is immediately replaced by the lowest numbered functional battery that has not yet been put in use. Suppose that the lifetimes of the different batteries are independent exponential random variables each having rate μ. At a random time, call it T, a battery will fail and our stockpile will be empty. At that moment exactly one of the batteries—which we call battery X—will not yet have failed.
(a) What is $P\{X = n\}$?
(b) What is $P\{X = 1\}$?
(c) What is $P\{X = i\}$?

(d) Find $E[T]$.
(e) What is the distribution of T?

24. There are two servers available to process n jobs. Initially, each server begins work on a job. Whenever a server completes work on a job, that job leaves the system and the server begins processing a new job (provided there are still jobs waiting to be processed). Let T denote the time until all jobs have been processed. If the time that it takes server i to process a job is exponentially distributed with rate μ_i, $i = 1, 2$, find $E[T]$ and $\mathrm{Var}(T)$.

25. Customers can be served by any of three servers, where the service times of server i are exponentially distributed with rate μ_i, $i = 1, 2, 3$. Whenever a server becomes free, the customer who has been waiting the longest begins service with that server.
(a) If you arrive to find all three servers busy and no one waiting, find the expected time until you depart the system.
(b) If you arrive to find all three servers busy and one person waiting, find the expected time until you depart the system.

26. Each entering customer must be served first by server 1, then by server 2, and finally by server 3. The amount of time it takes to be served by server i is an exponential random variable with rate μ_i, $i = 1, 2, 3$. Suppose you enter the system when it contains a single customer who is being served by server 3.
(a) Find the probability that server 3 will still be busy when you move over to server 2.
(b) Find the probability that server 3 will still be busy when you move over to server 3.
(c) Find the expected amount of time that you spend in the system. (Whenever you encounter a busy server, you must wait for the service in progress to end before you can enter service.)
(d) Suppose that you enter the system when it contains a single customer who is being served by server 2. Find the expected amount of time that you spend in the system.

27. Show, in Example 5.7, that the distributions of the total cost are the same for the two algorithms.

28. Consider n components with independent lifetimes, which are such that component i functions for an exponential time with rate λ_i. Suppose that all components are initially in use and remain so until they fail.
(a) Find the probability that component 1 is the second component to fail.
(b) Find the expected time of the second failure.

Hint: Do not make use of part (a).

29. Let X and Y be independent exponential random variables with respective rates λ and μ, where $\lambda > \mu$. Let $c > 0$.
(a) Show that the conditional density function of X, given that $X + Y = c$, is

$$f_{X|X+Y}(x|c) = \frac{(\lambda - \mu)e^{-(\lambda-\mu)x}}{1 - e^{-(\lambda-\mu)c}}, \qquad 0 < x < c$$

(b) Use part (a) to find $E[X|X + Y = c]$.
(c) Find $E[Y|X + Y = c]$.

30. The lifetimes of A's dog and cat are independent exponential random variables with respective rates λ_d and λ_c. One of them has just died. Find the expected additional lifetime of the other pet.

31. A doctor has scheduled two appointments, one at 1 P.M. and the other at 1:30 P.M. The amounts of time that appointments last are independent exponential random variables with mean 30 minutes. Assuming that both patients are on time, find the expected amount of time that the 1:30 appointment spends at the doctor's office.

32. Let X be a uniform random variable on (0, 1), and consider a counting process where events occur at times $X + i$, for $i = 0, 1, 2, \ldots$.
    (a) Does this counting process have independent increments?
    (b) Does this counting process have stationary increments?

33. Let X and Y be independent exponential random variables with respective rates λ and μ.
    (a) Argue that, conditional on $X > Y$, the random variables $\min(X, Y)$ and $X - Y$ are independent.
    (b) Use part (a) to conclude that for any positive constant c

$$\begin{aligned}E[\min(X, Y)|X > Y + c] &= E[\min(X, Y)|X > Y]\\ &= E[\min(X, Y)] = \frac{1}{\lambda + \mu}\end{aligned}$$

    (c) Give a verbal explanation of why $\min(X, Y)$ and $X - Y$ are (unconditionally) independent.

34. Two individuals, A and B, both require kidney transplants. If she does not receive a new kidney, then A will die after an exponential time with rate μ_A, and B after an exponential time with rate μ_B. New kidneys arrive in accordance with a Poisson process having rate λ. It has been decided that the first kidney will go to A (or to B if B is alive and A is not at that time) and the next one to B (if still living).
    (a) What is the probability that A obtains a new kidney?
    (b) What is the probability that B obtains a new kidney?

35. Show that Definition 5.1 of a Poisson process implies Definition 5.3.

*36. Let $S(t)$ denote the price of a security at time t. A popular model for the process $\{S(t), t \geqslant 0\}$ supposes that the price remains unchanged until a "shock" occurs, at which time the price is multiplied by a random factor. If we let $N(t)$ denote the number of shocks by time t, and let X_i denote the ith multiplicative factor, then this model supposes that

$$S(t) = S(0) \prod_{i=1}^{N(t)} X_i$$

where $\prod_{i=1}^{N(t)} X_i$ is equal to 1 when $N(t) = 0$. Suppose that the X_i are independent exponential random variables with rate μ; that $\{N(t), t \geqslant 0\}$ is a Poisson process with rate λ; that $\{N(t), t \geqslant 0\}$ is independent of the X_i; and that $S(0) = s$.
    (a) Find $E[S(t)]$.
    (b) Find $E[S^2(t)]$.

37. A machine works for an exponentially distributed time with rate μ and then fails. A repair crew checks the machine at times distributed according to a Poisson process with rate λ; if the machine is found to have failed then it is immediately replaced. Find the expected time between replacements of machines.

38. Let $\{M_i(t), t \geqslant 0\}, i = 1, 2, 3$ be independent Poisson processes with respective rates $\lambda_i, i = 1, 2$, and set

$$N_1(t) = M_1(t) + M_2(t), \quad N_2(t) = M_2(t) + M_3(t)$$

The stochastic process $\{(N_1(t), N_2(t)), t \geqslant 0\}$ is called a bivariate Poisson process.
(a) Find $P\{N_1(t) = n, N_2(t) = m\}$.
(b) Find $\text{Cov}\big(N_1(t), N_2(t)\big)$.

39. A certain scientific theory supposes that mistakes in cell division occur according to a Poisson process with rate 2.5 per year, and that an individual dies when 196 such mistakes have occurred. Assuming this theory, find
(a) the mean lifetime of an individual,
(b) the variance of the lifetime of an individual.
Also approximate
(c) the probability that an individual dies before age 67.2,
(d) the probability that an individual reaches age 90,
(e) the probability that an individual reaches age 100.

*40. Show that if $\{N_i(t), t \geqslant 0\}$ are independent Poisson processes with rate $\lambda_i, i = 1, 2$, then $\{N(t), t \geqslant 0\}$ is a Poisson process with rate $\lambda_1 + \lambda_2$ where $N(t) = N_1(t) + N_2(t)$.

41. In Exercise 40 what is the probability that the first event of the combined process is from the N_1 process?

42. Let $\{N(t), t \geqslant 0\}$ be a Poisson process with rate λ. Let S_n denote the time of the nth event. Find
(a) $E[S_4]$,
(b) $E[S_4|N(1) = 2]$,
(c) $E[N(4) - N(2)|N(1) = 3]$.

43. Customers arrive at a two-server service station according to a Poisson process with rate λ. Whenever a new customer arrives, any customer that is in the system immediately departs. A new arrival enters service first with server 1 and then with server 2. If the service times at the servers are independent exponentials with respective rates μ_1 and μ_2, what proportion of entering customers completes their service with server 2?

44. Cars pass a certain street location according to a Poisson process with rate λ. A woman who wants to cross the street at that location waits until she can see that no cars will come by in the next T time units.
(a) Find the probability that her waiting time is 0.
(b) Find her expected waiting time.

Hint: Condition on the time of the first car.

45. Let $\{N(t), t \geqslant 0\}$ be a Poisson process with rate λ that is independent of the nonnegative random variable T with mean μ and variance σ^2. Find
(a) $\text{Cov}(T, N(T))$,
(b) $\text{Var}(N(T))$.

46. Let $\{N(t), t \geqslant 0\}$ be a Poisson process with rate λ that is independent of the sequence $X_1, X_2, \ldots$ of independent and identically distributed random variables with mean μ and variance σ^2. Find

$$\mathrm{Cov}\left(N(t), \sum_{i=1}^{N(t)} X_i\right)$$

47. Consider a two-server parallel queuing system where customers arrive according to a Poisson process with rate λ, and where the service times are exponential with rate μ. Moreover, suppose that arrivals finding both servers busy immediately depart without receiving any service (such a customer is said to be lost), whereas those finding at least one free server immediately enter service and then depart when their service is completed.
    (a) If both servers are presently busy, find the expected time until the next customer enters the system.
    (b) Starting empty, find the expected time until both servers are busy.
    (c) Find the expected time between two successive lost customers.

48. Consider an n-server parallel queuing system where customers arrive according to a Poisson process with rate λ, where the service times are exponential random variables with rate μ, and where any arrival finding all servers busy immediately departs without receiving any service. If an arrival finds all servers busy, find
    (a) the expected number of busy servers found by the next arrival,
    (b) the probability that the next arrival finds all servers free,
    (c) the probability that the next arrival finds exactly i of the servers free.

49. Events occur according to a Poisson process with rate λ. Each time an event occurs, we must decide whether or not to stop, with our objective being to stop at the last event to occur prior to some specified time T, where $T > 1/\lambda$. That is, if an event occurs at time t, $0 \leqslant t \leqslant T$, and we decide to stop, then we win if there are no additional events by time T, and we lose otherwise. If we do not stop when an event occurs and no additional events occur by time T, then we lose. Also, if no events occur by time T, then we lose. Consider the strategy that stops at the first event to occur after some fixed time s, $0 \leqslant s \leqslant T$.
    (a) Using this strategy, what is the probability of winning?
    (b) What value of s maximizes the probability of winning?
    (c) Show that one's probability of winning when using the preceding strategy with the value of s specified in part (b) is $1/e$.

50. The number of hours between successive train arrivals at the station is uniformly distributed on $(0, 1)$. Passengers arrive according to a Poisson process with rate 7 per hour. Suppose a train has just left the station. Let X denote the number of people who get on the next train. Find
    (a) $E[X]$,
    (b) $\mathrm{Var}(X)$.

51. If an individual has never had a previous automobile accident, then the probability he or she has an accident in the next h time units is $\beta h + o(h)$; on the other hand, if he or she has ever had a previous accident, then the probability is $\alpha h + o(h)$. Find the expected number of accidents an individual has by time t.

52. Teams 1 and 2 are playing a match. The teams score points according to independent Poisson processes with respective rates λ_1 and λ_2. If the match ends when one of the teams has scored k more points than the other, find the probability that team 1 wins.

    Hint: Relate this to the gambler's ruin problem.

53. The water level of a certain reservoir is depleted at a constant rate of 1000 units daily. The reservoir is refilled by randomly occurring rainfalls. Rainfalls occur according to a Poisson process with rate 0.2 per day. The amount of water added to the reservoir by a rainfall is 5000 units with probability 0.8 or 8000 units with probability 0.2. The present water level is just slightly below 5000 units.
    (a) What is the probability the reservoir will be empty after five days?
    (b) What is the probability the reservoir will be empty sometime within the next ten days?

54. A viral linear DNA molecule of length, say, 1 is often known to contain a certain "marked position," with the exact location of this mark being unknown. One approach to locating the marked position is to cut the molecule by agents that break it at points chosen according to a Poisson process with rate λ. It is then possible to determine the fragment that contains the marked position. For instance, letting m denote the location on the line of the marked position, then if L_1 denotes the last Poisson event time before m (or 0 if there are no Poisson events in $[0, m]$), and R_1 denotes the first Poisson event time after m (or 1 if there are no Poisson events in $[m, 1]$), then it would be learned that the marked position lies between L_1 and R_1. Find
    (a) $P\{L_1 = 0\}$,
    (b) $P\{L_1 < x\}$, $0 < x < m$,
    (c) $P\{R_1 = 1\}$,
    (d) $P\{R_1 > x\}$, $m < x < 1$.

    By repeating the preceding process on identical copies of the DNA molecule, we are able to zero in on the location of the marked position. If the cutting procedure is utilized on n identical copies of the molecule, yielding the data L_i, R_i, $i = 1, \ldots, n$, then it follows that the marked position lies between L and R, where

$$L = \max_i L_i, \quad R = \min_i R_i$$

    (e) Find $E[R - L]$, and in doing so, show that $E[R - L] \sim \frac{2}{n\lambda}$.

55. Consider a single server queuing system where customers arrive according to a Poisson process with rate λ, service times are exponential with rate μ, and customers are served in the order of their arrival. Suppose that a customer arrives and finds $n - 1$ others in the system. Let X denote the number in the system at the moment that customer departs. Find the probability mass function of X.

    Hint: Relate this to a negative binomial random variable.

56. An event independently occurs on each day with probability p. Let $N(n)$ denote the total number of events that occur on the first n days, and let T_r denote the day on which the rth event occurs.
    (a) What is the distribution of $N(n)$?
    (b) What is the distribution of T_1?

(c) What is the distribution of T_r?

(d) Given that $N(n) = r$, show that the set of r days on which events occurred has the same distribution as a random selection (without replacement) of r of the values $1, 2, \ldots, n$.

*57. Events occur according to a Poisson process with rate $\lambda = 2$ per hour.

(a) What is the probability that no event occurs between 8 P.M. and 9 P.M.?

(b) Starting at noon, what is the expected time at which the fourth event occurs?

(c) What is the probability that two or more events occur between 6 P.M. and 8 P.M.?

58. Consider the coupon collecting problem where there are m distinct types of coupons, and each new coupon collected is type j with probability p_j, $\sum_{j=1}^{m} p_j = 1$. Suppose you stop collecting when you have a complete set of at least one of each type. Show that

$$P\{i \text{ is the last type collected}\} = E\left[\prod_{j \neq i} (1 - U^{\lambda_j/\lambda_i})\right]$$

where U is a uniform random variable on $(0, 1)$.

59. There are two types of claims that are made to an insurance company. Let $N_i(t)$ denote the number of type i claims made by time t, and suppose that $\{N_1(t), t \geqslant 0\}$ and $\{N_2(t), t \geqslant 0\}$ are independent Poisson processes with rates $\lambda_1 = 10$ and $\lambda_2 = 1$. The amounts of successive type 1 claims are independent exponential random variables with mean \$1000 whereas the amounts from type 2 claims are independent exponential random variables with mean \$5000. A claim for \$4000 has just been received; what is the probability it is a type 1 claim?

*60. Customers arrive at a bank at a Poisson rate λ. Suppose two customers arrived during the first hour. What is the probability that

(a) both arrived during the first 20 minutes?

(b) at least one arrived during the first 20 minutes?

61. A system has a random number of flaws that we will suppose is Poisson distributed with mean c. Each of these flaws will, independently, cause the system to fail at a random time having distribution G. When a system failure occurs, suppose that the flaw causing the failure is immediately located and fixed.

(a) What is the distribution of the number of failures by time t?

(b) What is the distribution of the number of flaws that remain in the system at time t?

(c) Are the random variables in parts (a) and (b) dependent or independent?

62. Suppose that the number of typographical errors in a new text is Poisson distributed with mean λ. Two proofreaders independently read the text. Suppose that each error is independently found by proofreader i with probability p_i, $i = 1, 2$. Let X_1 denote the number of errors that are found by proofreader 1 but not by proofreader 2. Let X_2 denote the number of errors that are found by proofreader 2 but not by proofreader 1. Let X_3 denote the number of errors that are found by both proofreaders. Finally, let X_4 denote the number of errors found by neither proofreader.

(a) Describe the joint probability distribution of X_1, X_2, X_3, X_4.
(b) Show that

$$\frac{E[X_1]}{E[X_3]} = \frac{1 - p_2}{p_2} \quad \text{and} \quad \frac{E[X_2]}{E[X_3]} = \frac{1 - p_1}{p_1}$$

Suppose now that λ, p_1, and p_2 are all unknown.
(c) By using X_i as an estimator of $E[X_i]$, $i = 1, 2, 3$, present estimators of p_1, p_2, and λ.
(d) Give an estimator of X_4, the number of errors not found by either proofreader.

63. Consider an infinite server queuing system in which customers arrive in accordance with a Poisson process with rate λ, and where the service distribution is exponential with rate μ. Let $X(t)$ denote the number of customers in the system at time t. Find
(a) $E[X(t + s)|X(s) = n]$;
(b) $\text{Var}[X(t + s)|X(s) = n]$.

Hint: Divide the customers in the system at time $t + s$ into two groups, one consisting of "old" customers and the other of "new" customers.

(c) Consider an infinite server queuing system in which customers arrive according to a Poisson process with rate λ, and where the service times are all exponential random variables with rate μ. If there is currently a single customer in the system, find the probability that the system becomes empty when that customer departs.

*64. Suppose that people arrive at a bus stop in accordance with a Poisson process with rate λ. The bus departs at time t. Let X denote the total amount of waiting time of all those who get on the bus at time t. We want to determine $\text{Var}(X)$. Let $N(t)$ denote the number of arrivals by time t.
(a) What is $E[X|N(t)]$?
(b) Argue that $\text{Var}[X|N(t)] = N(t)t^2/12$.
(c) What is $\text{Var}(X)$?

65. An average of 500 people pass the California bar exam each year. A California lawyer practices law, on average, for 30 years. Assuming these numbers remain steady, how many lawyers would you expect California to have in 2050?

66. Policyholders of a certain insurance company have accidents at times distributed according to a Poisson process with rate λ. The amount of time from when the accident occurs until a claim is made has distribution G.
(a) Find the probability there are exactly n incurred but as yet unreported claims at time t.
(b) Suppose that each claim amount has distribution F, and that the claim amount is independent of the time that it takes to report the claim. Find the expected value of the sum of all incurred but as yet unreported claims at time t.

67. Satellites are launched into space at times distributed according to a Poisson process with rate λ. Each satellite independently spends a random time (having distribution G) in space before falling to the ground. Find the probability that none of the satellites in the air at time t was launched before time s, where $s < t$.

68. Suppose that electrical shocks having random amplitudes occur at times distributed according to a Poisson process $\{N(t), t \geqslant 0\}$ with rate λ. Suppose that the

amplitudes of the successive shocks are independent both of other amplitudes and of the arrival times of shocks, and also that the amplitudes have distribution F with mean μ. Suppose also that the amplitude of a shock decreases with time at an exponential rate α, meaning that an initial amplitude A will have value $Ae^{-\alpha x}$ after an additional time x has elapsed. Let $A(t)$ denote the sum of all amplitudes at time t. That is,

$$A(t) = \sum_{i=1}^{N(t)} A_i e^{-\alpha(t-S_i)}$$

where A_i and S_i are the initial amplitude and the arrival time of shock i.

(a) Find $E[A(t)]$ by conditioning on $N(t)$.

(b) Without any computations, explain why $A(t)$ has the same distribution as does $D(t)$ of Example 5.21.

69. Let $\{N(t), t \geqslant 0\}$ be a Poisson process with rate λ. For $s < t$, find

(a) $P(N(t) > N(s))$;

(b) $P(N(s) = 0, N(t) = 3)$;

(c) $E[N(t)|N(s) = 4]$;

(d) $E[N(s)|N(t) = 4]$.

70. For the infinite server queue with Poisson arrivals and general service distribution G, find the probability that

(a) the first customer to arrive is also the first to depart.

Let $S(t)$ equal the sum of the remaining service times of all customers in the system at time t.

(b) Argue that $S(t)$ is a compound Poisson random variable.

(c) Find $E[S(t)]$.

(d) Find $\text{Var}(S(t))$.

71. Let S_n denote the time of the nth event of the Poisson process $\{N(t),\ t \geqslant 0\}$ having rate λ. Show, for an arbitrary function g, that the random variable $\sum_{i=1}^{N(t)} g(S_i)$ has the same distribution as the compound Poisson random variable $\sum_{i=1}^{N(t)} g(U_i)$, where $U_1, U_2, \ldots$ is a sequence of independent and identically distributed uniform $(0, t)$ random variables that is independent of N, a Poisson random variable with mean λt. Consequently, conclude that

$$E\left[\sum_{i=1}^{N(t)} g(S_i)\right] = \lambda \int_0^t g(x)\,dx \qquad \text{Var}\left(\sum_{i=1}^{N(t)} g(S_i)\right) = \lambda \int_0^t g^2(x)\,dx$$

72. A cable car starts off with n riders. The times between successive stops of the car are independent exponential random variables with rate λ. At each stop one rider gets off. This takes no time, and no additional riders get on. After a rider gets off the car, he or she walks home. Independently of all else, the walk takes an exponential time with rate μ.

(a) What is the distribution of the time at which the last rider departs the car?

(b) Suppose the last rider departs the car at time t. What is the probability that all the other riders are home at that time?

73. Shocks occur according to a Poisson process with rate λ, and each shock independently causes a certain system to fail with probability p. Let T denote the time at which the system fails and let N denote the number of shocks that it takes.
   (a) Find the conditional distribution of T given that $N = n$.
   (b) Calculate the conditional distribution of N, given that $T = t$, and notice that it is distributed as 1 plus a Poisson random variable with mean $\lambda(1 - p)t$.
   (c) Explain how the result in part (b) could have been obtained without any calculations.

74. The number of missing items in a certain location, call it X, is a Poisson random variable with mean λ. When searching the location, each item will independently be found after an exponentially distributed time with rate μ. A reward of R is received for each item found, and a searching cost of C per unit of search time is incurred. Suppose that you search for a fixed time t and then stop.
   (a) Find your total expected return.
   (b) Find the value of t that maximizes the total expected return.
   (c) The policy of searching for a fixed time is a static policy. Would a dynamic policy, which allows the decision as to whether to stop at each time t, depend on the number already found by t be beneficial?

   Hint: How does the distribution of the number of items not yet found by time t depend on the number already found by that time?

75. Suppose that the times between successive arrivals of customers at a single-server station are independent random variables having a common distribution F. Suppose that when a customer arrives, he or she either immediately enters service if the server is free or else joins the end of the waiting line if the server is busy with another customer. When the server completes work on a customer, that customer leaves the system and the next waiting customer, if there are any, enters service. Let X_n denote the number of customers in the system immediately before the nth arrival, and let Y_n denote the number of customers that remain in the system when the nth customer departs. The successive service times of customers are independent random variables (which are also independent of the interarrival times) having a common distribution G.
   (a) If F is the exponential distribution with rate λ, which, if any, of the processes $\{X_n\}$, $\{Y_n\}$ is a Markov chain?
   (b) If G is the exponential distribution with rate μ, which, if any, of the processes $\{X_n\}$, $\{Y_n\}$ is a Markov chain?
   (c) Give the transition probabilities of any Markov chains in parts (a) and (b).

76. For the model of Example 5.27, find the mean and variance of the number of customers served in a busy period.

77. Suppose that customers arrive to a system according to a Poisson process with rate λ. There are an infinite number of servers in this system so a customer begins service upon arrival. The service times of the arrivals are independent exponential random variables with rate μ, and are independent of the arrival process. Customers depart the system when their service ends. Let N be the number of arrivals before the first departure.
   (a) Find $P(N = 1)$.
   (b) Find $P(N = 2)$.

(c) Find $P(N = j)$.
(d) Find the probability that the first to arrive is the first to depart.
(e) Find the expected time of the first departure.

78. A store opens at 8 A.M. From 8 until 10 A.M. customers arrive at a Poisson rate of four an hour. Between 10 A.M. and 12 P.M. they arrive at a Poisson rate of eight an hour. From 12 P.M. to 2 P.M. the arrival rate increases steadily from eight per hour at 12 P.M. to ten per hour at 2 P.M.; and from 2 to 5 P.M. the arrival rate drops steadily from ten per hour at 2 P.M. to four per hour at 5 P.M.. Determine the probability distribution of the number of customers that enter the store on a given day.

*79. Consider a nonhomogeneous Poisson process whose intensity function $\lambda(t)$ is bounded and continuous. Show that such a process is equivalent to a process of counted events from a (homogeneous) Poisson process having rate λ, where an event at time t is counted (independent of the past) with probability $\lambda(t)/\lambda$; and where λ is chosen so that $\lambda(s) < \lambda$ for all s.

80. Let $T_1, T_2, \ldots$ denote the interarrival times of events of a nonhomogeneous Poisson process having intensity function $\lambda(t)$.
(a) Are the T_i independent?
(b) Are the T_i identically distributed?
(c) Find the distribution of T_1.

81. (a) Let $\{N(t), t \geqslant 0\}$ be a nonhomogeneous Poisson process with mean value function $m(t)$. Given $N(t) = n$, show that the unordered set of arrival times has the same distribution as n independent and identically distributed random variables having distribution function

$$F(x) = \begin{cases} \dfrac{m(x)}{m(t)}, & x \leqslant t \\ 1, & x \geqslant t \end{cases}$$

(b) Suppose that workmen incur accidents in accordance with a nonhomogeneous Poisson process with mean value function $m(t)$. Suppose further that each injured man is out of work for a random amount of time having distribution F. Let $X(t)$ be the number of workers who are out of work at time t. By using part (a), find $E[X(t)]$.

82. Let $X_1, X_2, \ldots$ be independent positive continuous random variables with a common density function f, and suppose this sequence is independent of N, a Poisson random variable with mean λ. Define

$$N(t) = \text{number of } i \leqslant N : X_i \leqslant t$$

Show that $\{N(t), t \geqslant 0\}$ is a nonhomogeneous Poisson process with intensity function $\lambda(t) = \lambda f(t)$.

83. Suppose that $\{N_0(t),\ t \geqslant 0\}$ is a Poisson process with rate $\lambda = 1$. Let $\lambda(t)$ denote a nonnegative function of t, and let

$$m(t) = \int_0^t \lambda(s)\, ds$$

Define $N(t)$ by

$$N(t) = N_0(m(t))$$

Argue that $\{N(t), t \geqslant 0\}$ is a nonhomogeneous Poisson process with intensity function $\lambda(t), t \geqslant 0$.

Hint: Make use of the identity

$$m(t+h) - m(t) = m'(t)h + o(h)$$

*84. Let $X_1, X_2, \ldots$ be independent and identically distributed nonnegative continuous random variables having density function $f(x)$. We say that a record occurs at time n if X_n is larger than each of the previous values $X_1, \ldots, X_{n-1}$. (A record automatically occurs at time 1.) If a record occurs at time n, then X_n is called a *record value*. In other words, a record occurs whenever a new high is reached, and that new high is called the record value. Let $N(t)$ denote the number of record values that are less than or equal to t. Characterize the process $\{N(t), t \geqslant 0\}$ when
(a) f is an arbitrary continuous density function.
(b) $f(x) = \lambda e^{-\lambda x}$.

Hint: Finish the following sentence: There will be a record whose value is between t and $t + dt$ if the first X_i that is greater than t lies between ...

85. An insurance company pays out claims on its life insurance policies in accordance with a Poisson process having rate $\lambda = 5$ per week. If the amount of money paid on each policy is exponentially distributed with mean \$2000, what is the mean and variance of the amount of money paid by the insurance company in a four-week span?

86. In good years, storms occur according to a Poisson process with rate 3 per unit time, while in other years they occur according to a Poisson process with rate 5 per unit time. Suppose next year will be a good year with probability 0.3. Let $N(t)$ denote the number of storms during the first t time units of next year.
(a) Find $P\{N(t) = n\}$.
(b) Is $\{N(t)\}$ a Poisson process?
(c) Does $\{N(t)\}$ have stationary increments? Why or why not?
(d) Does it have independent increments? Why or why not?
(e) If next year starts off with three storms by time $t = 1$, what is the conditional probability it is a good year?

87. Determine

$$\text{Cov}[X(t), X(t+s)]$$

when $\{X(t), t \geqslant 0\}$ is a compound Poisson process.

88. Customers arrive at the automatic teller machine in accordance with a Poisson process with rate 12 per hour. The amount of money withdrawn on each transaction is a random variable with mean \$30 and standard deviation \$50. (A negative withdrawal means that money was deposited.) The machine is in use for 15 hours daily. Approximate the probability that the total daily withdrawal is less than \$6000.

89. Some components of a two-component system fail after receiving a shock. Shocks of three types arrive independently and in accordance with Poisson processes. Shocks of the first type arrive at a Poisson rate λ_1 and cause the first component to fail. Those of the second type arrive at a Poisson rate λ_2 and cause the second component to fail. The third type of shock arrives at a Poisson rate λ_3 and causes both components to fail. Let X_1 and X_2 denote the survival times for the two components. Show that the joint distribution of X_1 and X_2 is given by

$$P\{X_1 > s, X_1 > t\} = \exp\{-\lambda_1 s - \lambda_2 t - \lambda_3 \max(s, t)\}$$

This distribution is known as the *bivariate exponential distribution.*

90. In Exercise 89 show that X_1 and X_2 both have exponential distributions.

*91. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed exponential random variables. Show that the probability that the largest of them is greater than the sum of the others is $n/2^{n-1}$. That is, if

$$M = \max_j X_j$$

then show

$$P\left\{M > \sum_{i=1}^{n} X_i - M\right\} = \frac{n}{2^{n-1}}$$

Hint: What is $P\{X_1 > \sum_{i=2}^{n} X_i\}$?

92. Prove Equation (5.22).

93. Prove that

(a) $\max(X_1, X_2) = X_1 + X_2 - \min(X_1, X_2)$ and, in general,

(b) $$\begin{aligned}\max(X_1, \ldots, X_n) = \sum_1^n X_i - \sum\sum_{i<j} \min(X_i, X_j) \\ + \sum\sum\sum_{i<j<k} \min(X_i, X_j, X_k) + \cdots \\ + (-1)^{n-1} \min(X_i, X_j, \ldots, X_n)\end{aligned}$$

(c) Show by defining appropriate random variables X_i, $i = 1, \ldots, n$, and by taking expectations in part (b) how to obtain the well-known formula

$$P\left(\bigcup_1^n A_i\right) = \sum_i P(A_i) - \sum\sum_{i<j} P(A_i A_j) + \cdots + (-1)^{n-1} P(A_1 \cdots A_n)$$

(d) Consider n independent Poisson processes—the ith having rate λ_i. Derive an expression for the expected time until an event has occurred in all n processes.

94. A two-dimensional Poisson process is a process of randomly occurring events in the plane such that

(i) for any region of area A the number of events in that region has a Poisson distribution with mean λA, and

(ii) the number of events in nonoverlapping regions are independent.

For such a process, consider an arbitrary point in the plane and let X denote its distance from its nearest event (where distance is measured in the usual Euclidean manner). Show that

(a) $P\{X > t\} = e^{-\lambda\pi t^2}$,

(b) $E[X] = \frac{1}{2\sqrt{\lambda}}$.

95. Let $\{N(t), t \geqslant 0\}$ be a conditional Poisson process with a random rate L.

(a) Derive an expression for $E[L|N(t) = n]$.

(b) Find, for $s > t$, $E[N(s)|N(t) = n]$.

(c) Find, for $s < t$, $E[N(s)|N(t) = n]$.

96. For the conditional Poisson process, let $m_1 = E[L]$, $m_2 = E[L^2]$. In terms of m_1 and m_2, find $\mathrm{Cov}(N(s), N(t))$ for $s \leqslant t$.

97. Consider a conditional Poisson process in which the rate L is, as in Example 5.29, gamma distributed with parameters m and p. Find the conditional density function of L given that $N(t) = n$.

References

[1] H. Cramér and M. Leadbetter, "Stationary and Related Stochastic Processes," John Wiley, New York, 1966.

[2] S. Ross, "Stochastic Processes," Second Edition, John Wiley, New York, 1996.

[3] S. Ross, "Probability Models for Computer Science," Academic Press, 2002.