A Course in Applied Stochastic Processes

Richard Serfozo

The following are chapters on renewal and regenerative processes, and Poisson processes. They are two of the chapters in a book under preparation by Richard Serfozo. The sections marked with a star cover topics that could be deferred for later reading.

Contents

CHAPTER 1

Markov Chains

This chapter is under development.

CHAPTER 2

Renewal and Regenerative Processes

Renewal and regenerative processes are used to model stochastic phenomena in which an event, or combination of events, occurs repeatedly over time, and the times between occurrences are i.i.d. Models of such phenomena typically focus on (i) determining limiting averages for costs or other system parameters, via strong laws of large numbers (SLLN's), and (ii) establishing whether certain probabilities or expected values for a system converge over time, and evaluating their limits.

Accordingly, a significant part of this chapter is devoted to SLLN's for several types of regenerative processes, including Markov chains, Poisson processes, and Markov-renewal processes. These laws are based on the classical SLLN for sums of independent, identically distributed random variables. The main part of this chapter, however, covers Blackwell's renewal theorem, and an equivalent key renewal theorem. These results are important tools for characterizing the limiting behavior of probabilities and expectations associated with regenerative processes.

1. Renewal Processes

In this section, we introduce renewal processes and discuss several examples. The discussion covers Poisson processes and renewal processes that are "embedded" in more intricate processes.

We begin with some point process notation and terminology, which we use throughout this chapter. Suppose $0 \le T_1 \le T_2 \le \dots$ are random times at which a certain event occurs; or the T_n are locations in $\Re_+$ at which some phenomenon occurs. The number of the times T_n in the interval $(0, t]$ is

$$N(t) = \sum_{n=1}^{\infty} \mathbf{1}(T_n \le t), \qquad t \ge 0.$$

We assume this counting process is finite valued for each t, which is equivalent to $T_n \to \infty$ a.s. as $n \to \infty$.

The process $\{N(t) : t \ge 0\}$, denoted simply by $N(t)$, is a *point process* on $\Re_+$. The T_n are its *occurrence times* (or point locations). The point process $N(t)$ is *simple* if its occurrence times are distinct: $0 < T_1 < T_2 < \dots$ a.s. (there is at most one occurrence at any instant).

Definition 1. A simple point process $N(t)$ is a *renewal process* if the inter-occurrence times $\xi_n = T_n - T_{n-1}$, for $n \ge 1$, are independent with a common distribution F, where $F(0) = 0$ and $T_0 = 0$. The T_n are called *renewal times* because of the independent or renewed stochastic information at these times. The ξ_n are the *inter-renewal times*, and $N(t)$ is the *number of renewals* in $(0, t]$.

To establish the "existence" of a renewal process, one must justify that there exists a probability space and independent random variables $\xi_1, \xi_2, \ldots$ defined on it that have the distribution F. The existence of such a probability space is given by Theorem 6.1 in the Appendix.

Examples of renewal processes include the random times at which: customers enter a queue for service, insurance claims are filed, accidents or emergencies happen, or a stochastic process enters a special state of interest. In addition, T_n might be the location of the nth vehicle on a highway, or the location of the nth flaw along a pipeline or cable, or the cumulative quantity of a product processed in n production cycles. A discrete-time renewal process is one whose renewal times T_n are integer-valued. Such processes are used for modelling systems in discrete time, or for modelling sequential phenomena such as the occurrence of a certain character (or special data packet) in a string of characters (or packets), such as in DNA sequences.

Example 1.1. *Scheduled Maintenance.* An automobile is lubricated when its owner has driven it L miles or every M days, whichever comes first. Let $N(t)$ denote the number of lubrications up to time t. Suppose the numbers of miles driven in disjoint time periods are independent, and the number of miles in any time interval has the same distribution, regardless of where the interval begins. Then it is reasonable that $N(t)$ is a renewal process. The inter-renewal distribution is $F(t) = P\{\tau \wedge M \le t\}$, where τ denotes the time to accumulate L miles on the automobile.

This scheduled maintenance model applies to many types of systems where maintenance is performed when the system usage exceeds a certain level L or when a time M has elapsed. For instance, in reliability theory, the *Age Replacement* model of components or systems, replaces a component with lifetime τ if it fails or reaches a certain age M (see Exercise 19).

Example 1.2. *Service Times.* An operator in a call center answers calls one at a time. The calls are independent and homogeneous in that the callers, the call durations, and the nature of the calls are independent and homogeneous. Also, the time needed to process a typical call (which may include post-call processing) has a distribution F. Then one would be justified in modelling the number of calls $N(t)$ that the operator can process in time t as a renewal process. The time scale here refers to the time that the operator is actually working; it is not the real time scale that includes intervals with no calls, operator work-breaks, etc.

We begin with elementary properties of renewal processes. Unless specified otherwise, $N(t)$ will denote a renewal process with inter-renewal distribution F. Important relations between the times T_n and counts $N(t)$

are

$$\{N(t) \geq n\} = \{T_n \leq t\},$$
$$T_{N(t)} \leq t < T_{N(t)+1}.$$

In addition, $N(T_n) = n$ and

$$N(t) = \max\{n : T_n \leq t\} = \min\{n : T_{n+1} > t\}.$$

These relations (which also hold for simple point processes) are used to derive properties of $N(t)$ in terms of T_n, and vice versa.

We have a good understanding of $T_n = \sum_{k=1}^n \xi_k$, since it is a sum of independent variables with distribution F. In particular, by properties of convolutions of distributions (see the Appendix), we know that

$$P\{T_n \leq t\} = F^{n\star}(t),$$

which is the n-fold convolution of F. Then $\{N(t) \geq n\} = \{T_n \leq t\}$ yields

$$P\{N(t) \leq n\} = 1 - F^{(n+1)\star}(t). \tag{1.3}$$

Also, using $EN(t) = \sum_{n=1}^\infty P\{N(t) \geq n\}$ (a well-known formula for means; see Exercise 1), we have

$$EN(t) = \sum_{n=1}^{\infty} F^{n\star}(t). \tag{1.4}$$

Here is a justification that this mean and all moments of $N(t)$ are finite. Properties of moment generating functions are in the Appendix.

PROPOSITION 1.5. *For each $t \geq 0$, the moment generating function $E[e^{\alpha N(t)}]$ exists for some α in a neighborhood of* 0*, and hence*

$$E[N(t)^m] < \infty, \qquad m \geq 1.$$

PROOF. Choose $x > 0$ such that $p \equiv P\{\xi_1 > x\} > 0$. Consider the sum $S_n = \sum_{k=1}^n 1(\xi_k > x)$, which is the number of successes in n independent Bernoulli trials with probability of success p. The number of trials until the mth success is $Z_m = \min\{n : S_n = m\}$

Clearly $xS_n \leq T_n$, and so

$$N(t) = \max\{n : T_n \leq t\} \leq \max\{n : S_n = \lfloor t/x \rfloor\} \leq Z_{\lfloor t/x \rfloor + 1}.$$

Now Z_m has a negative binomial distribution with parameters m and p, and its moment generating function $E[e^{\alpha Z_m}]$ is given in (20.1) in Exercise 8. From these observations, we have

$$E[e^{\alpha N(t)}] \leq E[e^{\alpha Z_{\lfloor t/x \rfloor + 1}}] = \Big(\frac{pe^\alpha}{1 - qe^\alpha} \Big)^{\lfloor t/x \rfloor + 1}, \qquad 0 < \alpha < -\log q.$$

Thus, the moment generating function of $N(t)$ exists. This existence ensures that all moments of $N(t)$ exist (a well-known property of moment generating functions for nonnegative random variables). □

Keep in mind that the preceding properties of the renewal process $N(t)$ are true for any distribution F. When this distribution has a finite mean μ and variance σ^2, the distribution of $N(t)$, for large t, is approximately a normal distribution with mean t/μ and variance $t\sigma^2/\mu^3$ (this follows by the central limit theorem in Example 14.8 below). Refined asymptotic approximations for the mean of $N(t)$ is given in Example 16.5.

The rest of this section is devoted to examples of renewal processes. The most prominent renewal process is as follows.

EXAMPLE 1.6. *Poisson Process.* The renewal process $N(t)$ is a *Poisson process* with rate λ if the inter-renewal times have the exponential distribution $F(t) = 1 - e^{-\lambda t}$ with rate λ (its mean is λ^{-1}). In this case,

$$P\{T_n \le t\} = F^{n\star}(t) = \int_0^t \lambda^n x^{n-1} \frac{e^{-\lambda x}}{(n-1)!} dx.$$

This is a gamma distribution with parameters n and λ. Alternatively,

$$P\{T_n \le t\} = 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

This is justified by noting that the derivative of this function is clearly the integrand (the gamma density) in the preceding integral. Then by $\{N(t) \ge n\} = \{T_n \le t\}$, we arrive at

$$P\{N(t) \le n\} = \sum_{k=0}^{n} \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

This is the Poisson distribution with mean $EN(t) = \lambda t$.

Poisson processes are very important in the theory and applications of stochastic processes. We will discuss them further in Chapter 3. The discrete-time analogue of a Poisson process is the Bernoulli process described in Exercise 8.

EXAMPLE 1.7. *Delayed Renewal Process.* Many applications involve a renewal process $N(t)$ with the slight difference that the first renewal time ξ_1 does not have the same distribution as the other ξ_n, for $n \ge 2$. We call $N(t)$ a *delayed renewal process.* Elementary properties of delayed renewal processes are similar to those for renewal processes with the obvious changes (e.g., if ξ_1 has distribution G, then the time T_n of the nth renewal has the distribution $G \star F^{(n-1)\star}(t)$). More important, we will see that most of the limit theorems we present for renewal processes also apply to delayed renewal processes.

In addition to being of interest by themselves, renewal processes play an important role in analyzing more complex stochastic processes. Specifically, as a stochastic process evolves over time, it is natural for some event associated with its realization to occur again and again. When the "embedded" occurrence times of the event are renewal times, they may be useful for gleaning properties about the parent process. Stochastic processes with embedded renewal times include Markov chains, Markov processes, Markov-Renewal processes and more general regenerative processes. The Markov chains are described next and the other processes are introduced in later sections.

EXAMPLE 1.8. *Ergodic Markov Chain.* Let X_n denote a discrete-time ergodic (aperiodic, irreducible and positive recurrent) Markov chain on a countable state space. Consider any state i and let $0 < \nu_1 < \nu_2 < \ldots$ denote the (discrete) times at which X_n enters state i. The Markov property ensures that the times ν_n form a discrete-time renewal process when $X_0 = i$. These times form a delayed renewal process when $X_0 \neq i$. The Bernoulli process in Exercise 8 is a special case.

EXAMPLE 1.9. *Cyclic Renewal Process.* Consider a continuous-time stochastic process $X(t)$ that cycles through states $0, 1, \ldots, K-1$ in that order, again and again. That is, it starts at $X(0) = 0$, and its nth state is j if $n = mK + j$ for some m. For instance, in modelling the status of a machine or system, $X(t)$ might be the amount of deterioration of a system, or the number of shocks (or services) it has had, and the system is renewed whenever it ends a sojourn in state $K-1$.

Assume the sojourn times in the states are independent, and let F_j denote the sojourn time distribution for state j, where $F_j(0) = 0$. The time for the process $X(t)$ to complete a cycle from state 0 back to 0 has the distribution $F = F_0 \star F_1 \star \cdots \star F_{K-1}$. Then it is clear that the times at which $X(t)$ enters state 0 form a renewal process with inter-renewal distribution F. We call $X(t)$ a *cyclic* renewal process.

There are many other renewal processes embedded in $X(t)$. For instance, the times at which the process enters any fixed state i form a delayed renewal process with the same distribution F. Another more subtle delayed renewal process is the sequence of times at which the processes $X(t)$ bypasses state 0 by jumping from state $K-1$ to state 1 (assuming $F_0(0) > 0$); see Exercise 6. It is quite natural for a single stochastic process to contain several such embedded renewal processes associated with particular features of the process.

EXAMPLE 1.10. *Alternating Renewal Process.* An *alternating* renewal process is a cyclic renewal process with only two states, say 0 and 1. This might be appropriate for indicating whether a system is working (state 1) or not working (state 0), or whether a library book is available or unavailable for use.

2. Strong Laws of Large Numbers

This section covers strong laws of large numbers (SLLN's) for renewal processes and related stochastic processes. We will use these strong laws throughout this chapter. Background material on convergence a.s. is in the Appendix.

We begin with the classical law for independent random variables, which is proved in standard texts on probability.

THEOREM 2.1. (Classical SLLN) *If $X_1, X_2, \ldots$ are independent, identically distributed random variables with a finite expectation μ, then*

$$n^{-1}\sum_{k=1}^{n} X_k \to \mu, \qquad \textit{a.s. as } n \to \infty.$$

Throughout this section, we assume $N(t)$ is a point process on $\Re_+$ with occurrence times T_n. The following result says that a SLLN for $N(t)$ is equivalent to a SLLN for T_n. Here and below, the limit statements are a.s., but for convenience, we will sometimes suppress this phrase.

THEOREM 2.2. *For a positive constant μ (or random variable that is positive a.s.), the following statements are equivalent:*

$$\lim_{n\to\infty} n^{-1}T_n = \mu, \quad \textit{a.s.} \tag{2.3}$$

$$\lim_{t\to\infty} t^{-1}N(t) = 1/\mu, \quad \textit{a.s.} \tag{2.4}$$

PROOF. Suppose (2.3) holds. We know $T_{N(t)} \le t < T_{N(t)+1}$. Dividing these terms by $N(t)$, we have

$$\frac{T_{N(t)}}{N(t)} \le \frac{t}{N(t)} < \frac{T_{N(t)+1}}{N(t)+1}\frac{N(t)+1}{N(t)}.$$

Supposition (2.3) along with $N(t) \uparrow \infty$ and $(N(t)+1)/N(t) \to 1$ ensure that the first and last terms in this display converge to μ. Since $t/N(t)$ is sandwiched between these terms, it must also converge to their limit μ. This proves (2.4).

Conversely, suppose (2.4) holds. When $N(t)$ is simple, $N(T_n) = n$, and so $T_n/n = T_n/N(T_n) \to \mu$. When $N(t)$ is not simple, $N(T_n) \ge n$ and one can prove (2.3) as suggested in Exercise 18. □

COROLLARY 2.5. (SLLN for Renewal Processes). *If $N(t)$ is a renewal process whose inter-renewal times have a finite mean μ, then*

$$t^{-1}N(t) \to 1/\mu, \quad \textit{a.s. as } t \to \infty.$$

PROOF. This follows by Theorem 2.2, since the classical SLLN ensures that $n^{-1}T_n \to \mu$. □

This result has the extension that $t^{-1}N(t) \to 0$ a.s. as $t \to \infty$, when the inter-renewal times have an infinite mean. Indeed, the classical SLLN extends to $n^{-1}T_n \to \infty$, when $ET_1 = \infty$, and this in turn implies $t^{-1}N(t) \to 0$ by the argument above for proving (2.3) implies (2.4).

EXAMPLE 2.6. *Statistical Estimation.* Suppose $N(t)$ is a Poisson process with rate λ, but this rate is not known, and one wants to estimate it. One approach is to observe the process for a fixed time interval of length t and record $N(t)$. Then an estimator for λ is

$$\hat{\lambda}_t \equiv t^{-1}N(t).$$

This estimator is unbiased in that $E\hat{\lambda}_t = \lambda$. It is also a *consistent estimator* since $\hat{\lambda}_t \to \lambda$ by Corollary 2.5. Similarly, if $N(t)$ is a renewal process whose inter-renewal distribution has a finite mean μ, then $\hat{\mu}_t \equiv t/N(t)$ is a consistent estimator for μ (but it is not unbiased). Of course, if it is practical to observe a fixed number n of renewals (rather than observing over a "fixed" time), then $n^{-1}T_n$ is an unbiased and consistent estimator of μ.

We now present a framework for obtaining SLLN's for a variety of stochastic processes. Consider a real-valued stochastic process $\{Z(t) : t \geq 0\}$ on the same probability space as the point process $N(t)$. The next result describes natural conditions under which the limit of its average value $t^{-1}Z(t)$ exists. For instance, $Z(t)$ might denote a cumulative utility (e.g., cost or reward) associated with a system, and one is interested in the utility per unit time $t^{-1}Z(t)$ for large t. The following theorem relates the limit of the *time average* $t^{-1}Z(t)$ to the limit of the embedded *interval average* $n^{-1}Z(T_n)$. The main assumption is that the maximum fluctuation M_n of $Z(t)$ in the interval $(T_{n-1}, T_n]$ does not increase faster than n as $n \to \infty$, a rather weak assumption.

THEOREM 2.7. *Suppose the real-valued stochastic process $Z(t)$ is either increasing, or it satisfies*

$$n^{-1}M_n \equiv n^{-1} \sup_{T_{n-1}<t\leq T_n} |Z(t) - Z(T_{n-1})| \to 0, \qquad \textit{a.s. as } n \to \infty.$$

Assume the limit $\mu \equiv \lim_{n\to\infty} n^{-1}T_n$ exists a.s. and is a positive constant. Then the following statements are equivalent, for $a \in \Re$:

$$n^{-1}Z(T_n) \to a, \qquad \textit{a.s. as } n \to \infty. \tag{2.8}$$

$$t^{-1}Z(t) \to a/\mu, \qquad \textit{a.s. as } t \to \infty. \tag{2.9}$$

PROOF. Suppose (14.6) holds, and consider

$$t^{-1}Z(t) = t^{-1}Z(T_{N(t)}) + t^{-1}[Z(t) - Z(T_{N(t)})].$$

By Theorem 2.2, we know $N(t)/t \to 1/\mu$, and so $N(t) \to \infty$. These properties and supposition (14.6) yield

$$t^{-1}Z(T_{N(t)}) = [Z(T_{N(t)})/N(t)][N(t)/t] \to a/\mu.$$

In light of these observations, to prove $t^{-1}Z(t) \to a/\mu$, it remains to show $r(t) \equiv t^{-1}[Z(t) - Z(T_{N(t)})] \to 0$. In case $Z(t)$ is increasing, (14.6) and $N(t)/t \to 1/\mu$ ensure that

$$|r(t)| \leq \frac{[Z(T_{N(t)+1}) - Z(T_{N(t)})]}{N(t)} \frac{N(t)}{t} \to 0.$$

Also, in the other case in which $n^{-1}M_n \to 0$,

$$|r(t)| \le [M_{N(t)+1}/(N(t)+1)][(N(t)+1)/t] \to 0.$$

Thus $r(t) \to 0$, which completes the proof that (14.6) implies (14.7).

Conversely, if (14.7) is true then (14.6) follows since

$$n^{-1}Z(T_n) = T_n^{-1}Z(T_n)(T_n/n) \to a.$$

□

We will see a number of applications of Theorem 2.7 throughout this chapter. Here are two elementary examples.

EXAMPLE 2.10. *Renewal Reward Process.* Suppose $N(t)$ is a renewal process associated with a system in which a reward Y_n (or cost or utility value) is received at time T_n, for $n \ge 1$. Then the total reward in $(0,t]$ is

$$Z(t) = \sum_{n=1}^{\infty} Y_n \mathbf{1}(T_n \le t) = \sum_{n=1}^{N(t)} Y_n, \qquad t \ge 0.$$

For the last sum, one uses the convention $\sum_{n=1}^{0}(\cdot) = 0$. For instance, Y_n might be claims received by an insurance company at times T_n, and $Z(t)$ would represent the cumulative claims.

The process $Z(t)$ is a *renewal reward process* if the pairs (ξ_n, Y_n), $n \ge 1$, are i.i.d. (ξ_n and Y_n may be dependent). Under this assumption, it follows by Theorem 2.7 that the average reward per unit time is

$$\lim_{t\to\infty} t^{-1}Z(t) = EY_1/E\xi_1, \qquad \text{a.s.,}$$

provided the expectations are finite. This result is very useful for practitioners in many diverse contexts. One only has to justify the renewal conditions and evaluate the expectations. In complicated systems with many activities, a little thought may be needed to identify the renewal times as well as the associated rewards.

EXAMPLE 2.11. *Cyclic Renewal Process.* Let $X(t)$ be a cyclic renewal process on $0, \dots, K-1$ as in Example 1.9. Recall the entrance times to state 0 form a renewal process, and the mean inter-renewal time is $\mu = \mu_0 + \cdots + \mu_{K-1}$, where μ_i is the mean sojourn time in state i. Suppose a cost or value $f(i)$ per unit time is incurred whenever $X(t)$ is in state i. Then the average cost per unit time is

$$\lim_{t\to\infty} t^{-1} \int_0^t f(X(s))ds = \frac{1}{\mu}\sum_{i=0}^{K-1} f(i)\mu_i. \tag{2.12}$$

This follows by applying Theorem 2.7 to $Z(t) = \int_0^t f(X(s))ds$ and noting that $EZ(T_1) = \sum_{i=0}^{K-1} f(i)\mu_i$.

A particular case of (2.12) says that the portion of time $X(t)$ spends in a subset of states J is

$$\lim_{t\to\infty} t^{-1}\int_0^t \mathbf{1}(X(s)\in J)ds = \sum_{j\in J}\mu_j/\mu.$$

3. The Renewal Function

Throughout this section, $N(t)$ will denote a renewal process with inter-renewal distribution F and finite mean μ. Its *renewal function* is defined as

$$U(t) \equiv \sum_{n=0}^{\infty} F^{n\star}(t), \qquad t \geq 0, \tag{3.1}$$

where $F^{0\star}(t) = 1$ or 0 according as $t \geq 0$ or $t < 0$. That is, $U(t) = EN(t) + 1$, for $t > 0$, is viewed as the expected number of renewals up to time t including a fictitious renewal at time 0. Although the renewal function $U(t)$ is ostensibly very simple, it has some remarkable uses as we will now explore.

Note that $U(t)$ is similar to a distribution function in that it is nondecreasing and right-continuous; however $U(t) \uparrow \infty$ as $t \to \infty$. Keep in mind that $U(t)$ has a unit jump at $t = 0$.

The next result says that when the inter-renewal times are continuous, the renewal function has a density; this implies that the probability of a renewal at any time is 0. For example, if $N(t)$ is a Poisson process with rate λ, then $U'(t) = \lambda$, for $t > 0$, and $P\{N(t) > N(t-)\} = 0$, for each t.

PROPOSITION 3.2. *Suppose the inter-renewal distribution F has a density f. Then $U(t)$ also has a density for $t > 0$, and it is*

$$U'(t) = \sum_{n=1}^{\infty} f^{n\star}(t).$$

In addition, $P\{N(t) > N(t-)\} = 0$, for each t.

PROOF. The first assertion follows since $U(t) = \sum_{n=0}^{\infty} F^{n\star}(t)$, and the derivative of $F^{n\star}(t)$ is $f^{n\star}(t)$. The second assertion, which is equivalent to $N(t) - N(t-) = 0$ a.s., will follow if $E[N(t) - N(t-)] = 0$. But the last equality is true since by the monotone convergence theorem (see the Appendix, Theorem 8.6) and the continuity of U,

$$EN(t-) = E[\lim_{s\uparrow t} N(s)] = \lim_{s\uparrow t} EN(s) = U(t) = EN(t).$$

□

Some of the results below will be slightly different when the inter-arrival distribution is arithmetic. The distribution F is *arithmetic* (or periodic) if it is piecewise constant and its points of increase are contained in a set $\{0, d, 2d, \ldots\}$, and the largest $d > 0$ with this property is the *span*. In this case, it is clear that the distributions $F^{n\star}$ and the renewal function $U(t)$ also have this arithmetic property. If F is not arithmetic, we call it *non-arithmetic*. A distribution with a continuous part is necessarily non-arithmetic.

The expectation of a function $g : \Re_+ \to \Re$ on the interval $[a, b]$ with respect to F will be expressed as the *Riemann-Stieltjes integral*

$$\int_{[a,b]} g(t)dF(t), \qquad a < b \leq \infty.$$

The Appendix describes this type of integral and its role for representing expectations. All the functions in this book like g are assumed to be measurable (see the Appendix); for simplicity, we will not repeat this assumption unless emphasis is needed. Riemann-Stieltjes integrals with respect to U are defined similarly , since U is nondecreasing and right-continuous just like a distribution function (although $U(t) \to \infty$ as $t \to \infty$). An example is

$$\int_{[0,b]} g(t)dU(t) = g(0) + \int_{(0,b]} g(t)dU(t).$$

The right-hand side highlights that $g(0)U(0) = g(0)$ is the contribution from the unit jump of U at 0.

An important property of the renewal function $U(t)$ is that it uniquely determines the distribution F. To see this, we will use Laplace transforms. The Laplace-Stieltjes or simply the *Laplace transform* of F is defined by

$$\hat{F}(\alpha) \equiv \int_{\Re_+} e^{-\alpha t} dF(t), \qquad \alpha \geq 0.$$

A basic property is that the transform $\hat{F}$ uniquely determines F and vice versa. The Laplace transform $\hat{U}(\alpha)$ of $U(t)$ is defined similarly. Now, taking the Laplace transform of both sides in (3.1), we have

$$\hat{U}(\alpha) = \sum_{n=0}^{\infty} \widehat{F^{n\star}}(\alpha) = \sum_{n=0}^{\infty} \hat{F}(\alpha)^n = 1/(1 - \hat{F}(\alpha)).$$

This yields the following result.

Proposition 3.3. The Laplace transforms $\hat{U}(\alpha)$ and $\hat{F}(\alpha)$ determine each other uniquely by the relation $\hat{U}(\alpha) = 1/(1 - \hat{F}(\alpha))$. Hence U and F uniquely determine each other.

One can sometimes use this result for identifying that a renewal process is of a certain type, such as a Poisson process, whose renewal function has the form $U(t) = \lambda t + 1$.

Remark 3.4. A renewal process $N(t)$, whose inter-renewal times have a finite mean, is a Poisson process with rate λ if and only if $EN(t) = \lambda t$, for $t \geq 0$.

Two other renewal processes with tractable renewal functions are ones whose inter-renewal distribution is a convolution or mixture of exponential distributions; see Exercises 5 and 12. Sometimes the Laplace transform $\hat{U}(\alpha) = 1/(1 - \hat{F}(\alpha))$ can be inverted to determine $U(t)$. Unfortunately, nice expressions for renewal processes are the exception rather than the rule.

In addition to characterizing renewal processes as discussed above, renewal functions arise naturally in expressions for probabilities and expectations of functions associated with renewal processes. Such expressions are the focus of much of this chapter.

The next result describes an important family of functions of point processes as well as renewal processes. Expression (3.6) is a special case of Campbell's formula in the theory of point processes.

THEOREM 3.5. *Let $N(t)$ be a simple point process with point locations T_n such that $\eta(t) \equiv E[N(t)]$ is finite for each t. Then for any function $f : \Re_+ \to \Re$,*

$$E\Big[\sum_{n=1}^{N(t)} f(T_n)\Big] = \int_{[0,t]} f(s) d\eta(s), \qquad t \geq 0, \tag{3.6}$$

provided the integral exists.

PROOF. The following is a standard approach for proving properties of integrals. For convenience, denote the equality (3.6) by $\Sigma(f) = I(f)$. First, consider the simple piecewise-constant function

$$f(s) = \sum_{k=1}^{m} a_k \mathbf{1}(s \in (s_k, t_k]),$$

for fixed $0 \leq s_1 < t_1 < \ldots < s_m < t_m \leq t$. In this case,

$$\begin{aligned} \Sigma(f) &= E\Big[\sum_{k=1}^{m} a_k[N(t_k) - N(s_k)]\Big] \\ &= \sum_{k=1}^{m} a_k[\eta(t_k) - \eta(s_k)] \;=\; I(f). \end{aligned}$$

Next, for any nonnegative function f one can define simple functions f_m as above such that $f_m(s) \uparrow f(s)$ as $m \to \infty$ for each s. For instance, $f_m(s) = m \wedge (\lfloor 2^m f(s) \rfloor / 2^m)$. Then by the monotone convergence theorem (see the Appendix, Theorem 8.6) and the first part of this proof,

$$\Sigma(f) = \lim_{m\to\infty} \Sigma(f_m) = \lim_{m\to\infty} I(f_m) = I(f).$$

Thus, (3.6) is true for nonnegative f.

Finally, (3.6) is true for a general function f, since $f(s) = f(s)^+ - f(s)^-$ and the preceding part of the proof for nonnegative functions yield

$$\Sigma(f) = \Sigma(f^+) - \Sigma(f^-) = I(f^+) - I(f^-) = I(f).$$

□

Some applications of formula (3.6) take the following form.

COROLLARY 3.7. *Let $N(t)$ be a simple point process with point locations T_n such that $\eta(t) \equiv E[N(t)]$ is finite for each t. Suppose $X_1, X_2, \dots$ are random variables defined on the same probability space as the process $N(t)$, and there is a function $f : \Re_+ \to \Re$ such that $E[X_n|T_n = s] = f(s)$, independent of n. Then*

$$E\Big[\sum_{n=1}^{N(t)} X_n\Big] = \int_{[0,t]} f(s)d\eta(s), \qquad t \geq 0, \tag{3.8}$$

provided the integral exists.

PROOF. It suffices to prove the assertion for nonnegative X_n. Conditioning on T_n, we have

$$\begin{aligned} E\Big[\sum_{n=1}^{N(t)} X_n\Big] &= \sum_{n=1}^{\infty} E[X_n \mathbf{1}(T_n \leq t)] \\ &= \sum_{n=1}^{\infty} \int_{[0,t]} f(s)dF^{n\star}(s) \;=\; E\Big[\sum_{n=1}^{N(t)} f(T_n)\Big]. \end{aligned}$$

Then applying (3.6) to the last term yields (3.8). □

REMARK 3.9. Theorem 3.5 and Corollary 3.7 apply to renewal processes. For a renewal process $N(t)$, it is convenient to express (3.6) as

$$E\Big[\sum_{n=0}^{N(t)} f(T_n)\Big] = \int_{[0,t]} f(s)dU(s).$$

This includes a value $f(0)$ on both sides of the equality — recall $U(t)$ has a unit jump at $t = 0$. Similarly, (3.8) would be $E\Big[\sum_{n=0}^{N(t)} X_n\Big] = \int_{[0,t]} f(s)dU(s)$.

Here is a particular application of (3.8) for renewal processes. It is a special case of a general Wald identity for stopping times, which is covered in the theory of martingales.

COROLLARY 3.10. (Wald Identity for Renewals) *The renewal process $N(t)$ satisfies*

$$E[T_{N(t)+1}] = \mu E[N(t) + 1], \qquad t \geq 0.$$

PROOF. Since $T_{N(t)+1} = \sum_{n=0}^{N(t)} \xi_{n+1}$, an application of (3.8) with $X_n \equiv \xi_{n+1}$, where $E[\xi_{n+1}|T_n] = \mu$, yields

$$E[T_{N(t)+1}] = \mu U(t) = \mu E[N(t) + 1].$$

□

In light of this result, one might suspect that $E[T_{N(t)}] = \mu E[N(t)]$. However, this is not the case. In fact, $ET_{N(t)} \le \mu EN(t)$; and this is a strict inequality for a Poisson process; see Exercise 22.

EXAMPLE 3.11. *Discounted Rewards.* Suppose a renewal process $N(t)$ has rewards associated with it such that a reward (or cost) Y_n is obtained at the nth renewal time T_n. The rewards are discounted continuously over time and if a reward y occurs a time t, it has a discounted value of $ye^{-\alpha t}$. Then the total discounted reward up to time t is

$$\sum_{n=1}^{N(t)} Y_n e^{-\alpha T_n}.$$

As in Corollary 3.7, assume there is a function $f : \Re_+ \to \Re$ such that $E[Y_n|T_n = s] = f(s)$, independent of n. Then applying (3.8) to $X_n = Y_n e^{-\alpha T_n}$ yields

$$E\Big[\sum_{n=1}^{N(t)} Y_n e^{-\alpha T_n}\Big] = \int_{[0,t]} e^{-\alpha s} f(s) dU(s).$$

The next examples describe several systems modelled by renewal processes with the same type of inter-renewal distribution shown in (3.13) below (also see Exercise 16).

EXAMPLE 3.12. *Single-Server System.* Pallets are scheduled to arrive at an automatically guided vehicle (AGV) station according to a renewal process $N(t)$ with inter-arrival distribution F. The station is attended by a single AGV, which can transport only one pallet at a time. Pallets scheduled to arrive when the AGV is already busy transporting a pallet are diverted to another station. Assume the transportation times are independent with common distribution G.

Let us consider the times $\tilde{T}_n$ at which the AGV begins to transport a pallet (the times at which pallets arrive and the AGV is idle). For simplicity, assume a transport starts at time 0. To describe $\tilde{T}_1$, let τ denote a transport time for the first pallet. Then $\tilde{T}_1$ equals τ plus the waiting time $T_{N(\tau)+1} - \tau$ for the next pallet to arrive after transporting the first pallet. That is, $\tilde{T}_1 = T_{N(\tau)+1}$. When the next pallet arrives at time $T_{N(\tau)+1}$, the system is renewed and these cycles are repeated indefinitely. Thus $\tilde{T}_n$ are renewal times.

The inter-renewal distribution of $\tilde{T}_1$ and its mean have reasonable expressions in terms of the arrival process. Indeed, conditioning on τ, which is independent of $N(t)$, yields

$$P\{\tilde{T}_1 \le t\} = \int_{\Re_+} P\{T_{N(x)+1} \le t\} dG(x). \tag{3.13}$$

Also, if F has a finite mean μ, then by Wald's identity,

$$E[\tilde{T}_1] = \int_{\Re_+} E[T_{N(x)+1}]dG(x) = \mu \int_{\Re_+} U(x)dG(x).$$

EXAMPLE 3.14. *$GI/G/1/1$ System.* Consider a system in which customers arrive at a processing station according to a renewal process with inter-arrival distribution F and are processed by a single server. The processing or service times are independent with the common distribution G, and are independent of the arrival process. Also, customer arrivals during a service time are blocked from being served — they either go elsewhere or go without service. In this context the times $\tilde{T}_n$ at which customers begin services are renewal times as in the preceding example with inter-renewal distribution (3.13). This system is called a $GI/G/1/1$ system: GI/G means the inter-arrival and service times are i.i.d. (with general distributions) and $1/1$ means there is one server and at most one customer in the system.

EXAMPLE 3.15. *Geiger Counters.* A classical model of a Geiger counter assumes that electronic particles arrive at the counter according to a Poisson or renewal process. Upon recording an arrival of a particle, the counter is locked for a random time during which arrivals of new particles are not recorded. The times of being locked are i.i.d. and independent of the arrivals. Under these assumptions, it follows that the times $\tilde{T}_n$ at which particles are recorded are renewal times, and have the same structure as those for the $GI/G/1/1$ system described above. This so-called Type I model assumes that particles arriving while the counter is locked do not affect the counter.

A slightly different Type II Geiger counter model assumes that whenever the counter is locked and a particle arrives, that particle is not recorded, but it extends the locked period by another independent locking time. The times at which particles are registered are renewal times, but the inter-renewal distribution is more intricate than that for the Type I counter.

4. Future Expectations

We have just seen the usefulness of the renewal function for characterizing a renewal process and for describing some expected values of the process. In the following sections, we will discuss the major role a renewal function plays in describing the limiting behavior of probabilities and expectations associated with renewal and regenerative phenomena. This section outlines what to expect in the next three sections, which cover the heart of renewal theory.

The analysis to follow will use convolutions of functions with respect to the renewal function $U(t)$, such as

$$U \star h(t) = \int_{[0,t]} h(t-s)dU(s) = h(0) + \int_{(0,t]} h(t-s)dU(s),$$

where h is bounded on finite intervals and equals 0 for $t < 0$. As with distributions, we extend the domain of U to the entire real line, setting $U(t) = 0$ for $t < 0$.

Many probabilities and expectations associated with a renewal process $N(t)$ can be expressed as a function $H(t)$ that satisfies a recursive equation of the form

$$H(t) = h(t) + \int_0^t H(t-s)dF(s), \qquad t \geq 0.$$

This "renewal equation", under minor technical conditions given in the next section, has a unique solution of the form $H(t) = U \star h(t)$.

The next topic we address is the limiting behavior of such functions as $t \to \infty$. We will present Blackwell's theorem, and an equivalent key renewal theorem, which establish

$$\lim_{t\to\infty} U \star h(t) = \frac{1}{\mu} \int_{\Re_+} h(s)ds.$$

This is for non-arithmetic F; an analogous result holds for arithmetic F. Also, the integral is slightly different than the standard Riemann integral.

We cover the topics outlined above — Renewal Equations, Blackwell's Theorem and the Key Renewal Theorem — in the next three sections. Thereafter, we discuss applications of these theorems that describe the limiting behavior of probabilities and expectations associated with renewal, regenerative and Markov processes.

5. Renewal Equations

We begin our discussion of renewal equations with a concrete example.

Example 5.1. Let $X(t)$ be a cyclic renewal process on $0, 1, \ldots, K-1$, and consider the probability $H(t) \equiv P\{X(t) = i\}$ as a function of time, for a fixed state i. To show $H(t)$ satisfies a renewal equation, the standard approach is to condition on the time T_1 of the first renewal (the first entrance to state 0). The result is

$$H(t) = P\{X(t) = i, T_1 > t\} + P\{X(t) = i, T_1 \leq t\}, \tag{5.2}$$

where the last probability, conditioning on the renewal at T_1, is

$$\int_0^t P\{X(t) = i | T_1 = s\}dF(s) = \int_0^t H(t-s)dF(s).$$

Therefore, the recursive equation (5.2) that $H(t)$ satisfies is the renewal equation

$$H(t) = h(t) + F \star H(t),$$

where $h(t) = P\{X(t) = i, T_1 > t\}$.

With this example in mind, we are now ready for a formal definition of a renewal equation. Let $h(t)$ be a real-valued function on $\Re$ that is bounded on finite intervals and equals 0 for $t < 0$. The *renewal equation* for $h(t)$ and the distribution F is

$$H(t) = h(t) + \int_0^t H(t-s)dF(s), \qquad t \geq 0, \tag{5.3}$$

where $H(t)$ is a real-valued function. That is $H = h + F \star H$. We say $H(t)$ is a *solution of this equation* if it satisfies the equation, and is bounded on finite intervals and equals 0 for $t < 0$.

PROPOSITION 5.4. *The function $U \star h(t)$ is the unique solution to the renewal equation (5.3).*

PROOF. Clearly $U \star h(t) = 0$ for $t < 0$, and it is bounded on finite intervals since

$$\sup_{s \leq t} |U \star h(s)| \leq \sup_{s \leq t} |h(s)| U(t) < \infty, \qquad t \geq 0.$$

Also, $U \star h$ is a solution to the renewal equation, since by the definition of U and $F^{0\star} \star h = h$,

$$U \star h = \Big(F^{0\star} + F \star \sum_{n=1}^{\infty} F^{(n-1)\star}\Big) \star h = h + F \star (U \star h).$$

To prove $U \star h$ is the unique solution, let $H(t)$ be any solution to the renewal equation, and consider the difference $D(t) = H(t) - U \star h(t)$. From the renewal equation, we have $D = F \star D$, and so iterating this yields $D = F^{n\star} \star D$. Now, the finiteness of $U(t)$ implies $F^{n\star}(t) \to 0$, as $n \to \infty$, and hence $D(t) = 0$ for each t. This proves that $U \star h(t)$ is the unique solution of the renewal equation. □

The standard approach for deriving a renewal equation is by conditioning on the first renewal time, as illustrated above in Example 5.1. Upon establishing that a function $H(t)$ satisfies a renewal equation, one immediately obtains $H(t) = U \star h(t)$ by Proposition 5.4. For instance, Example 5.1 showed the probability $P\{X(t) = i\}$ for a cyclic renewal process satisfies a renewal equation, and so by Proposition 5.4, $P\{X(t) = i\} = U \star h(t)$, where $h(t) = P\{X(t) = i, T_1 > t\}$.

Note that $P\{X(t) = i\} = U \star h(t)$ is a "recursive" formula for the probability since part of the probability information appears in $h(t)$. Only in special cases are such recursive formulas for stochastic processes tractable enough for computations. On the other hand, $H(t) = U \star h(t)$ does provide a framework for analyzing the limit of $H(t)$ as we will soon see.

We will shed more light on these functions by answering the following questions: "What types of time-dependent probabilities and expected values can be expressed as $H(t) = U \star h(t)$?" Can one derive such functions directly without using a renewal equation? In summary, the answers are $H(t) = U \star h(t)$ is "universally" applicable, and one can derive such functions without using a renewal equation. Here are more details.

REMARK 5.5. If $H(t)$ is bounded on finite intervals and is 0 for $t < 0$, then $H(t) = U \star h(t)$, where $h(t) = H(t) - F \star H(t)$. This follows since $U = F^{0\star} + U \star F$, and so

$$H = F^{0\star} \star H = U \star (H - F \star H).$$

Although $H(t) = U \star h(t)$ is universally applicable, it may not be useful (e.g., for limit theorems) without linking it to a renewal process. A rather general linking condition is given by (5.7) below, which we call a "crude-regeneration" property (its importance will be seen in Section 8).

REMARK 5.6. *Bypassing a Renewal Equation.* Suppose $X(t)$ is a real-valued stochastic process such that $H(t) = E[X(t)]$ is bounded on finite intervals, and

$$E[X(T_1 + t)|T_1] = E[X(t)], \qquad t \geq 0. \tag{5.7}$$

Then $E[X(t)] = U \star h(t)$ where $h(t) = E[X(t)\mathbf{1}(T_1 > t)]$. This follows by the preceding remark since

$$\begin{aligned} h(t) &= H(t) - F \star H(t) = E[X(t)] - \int_0^t E[X(t-s)]dF(s) \\ &= E[X(t)] - \int_0^t E[X(t)|T_1 = s]dF(s) = E[X(t)\mathbf{1}(T_1 > t)]. \end{aligned}$$

6. Blackwell's Theorem

The next issue is to characterize the limit of functions of the form $H(t) = U \star h(t)$ as $t \to \infty$. Their limiting behavior is intimately related to the limiting behavior of $U(t)$, which we now consider.

Throughout this section, we assume $N(t)$ is a renewal process with renewal function $U(t)$ and mean inter-renewal time μ, which we allow to be finite or infinite. In Section 2, we saw that $N(t)/t \to 1/\mu$ a.s., and so $N(t)$ behaves asymptotically like t/μ as $t \to \infty$. This suggests $U(t) = EN(t) + 1$ should also behave asymptotically like t/μ. Here is a confirmation.

THEOREM 6.1. (Elementary Renewal Theorem)

$$t^{-1}U(t) \to 1/\mu, \qquad \textit{as } t \to \infty.$$

Here, $1/\mu = 0$ *when* $\mu = \infty$.

PROOF. For finite μ, using $t < T_{N(t)+1}$ and Wald's identity (Corollary 3.10), we have

$$t < E[T_{N(t)+1}] = \mu U(t).$$

This yields the lower bound $1/\mu < t^{-1}U(t)$. Also, this inequality hold trivially when $\mu = \infty$. With this bound in hand, to finish proving the assertion it suffices to show

$$\limsup_{t \to \infty} t^{-1}U(t) \leq 1/\mu. \tag{6.2}$$

To this end, for a constant b, define a renewal process $\overline{N}(t)$ with inter-renewal times $\overline{\xi}_n = \xi_n \wedge b$. Define $\overline{T}_n$ and $\overline{U}(t)$ accordingly. Clearly, $U(t) \leq \overline{U}(t)$. Also, by Wald's identity and $\overline{T}_{\overline{N}(t)+1} \leq t+b$ (since the $\overline{\xi}_n$ are bounded by b),

$$E[\xi_1 \wedge b]\overline{U}(t) = E[\overline{T}_{\overline{N}(t)+1}] \leq t + b.$$

Consequently,

$$t^{-1}U(t) \leq t^{-1}\overline{U}(t) \leq \frac{1 + b/t}{E[\xi_1 \wedge b]}.$$

Letting $t \to \infty$ and then letting $b \to \infty$ (whereupon the last fraction tends to $1/\mu$, even when $\mu = \infty$), we obtain (6.2), which finishes the proof. $\square$

A more definitive description of the asymptotic behavior of $U(t)$ is given in the following major result. Its proof is sketched in Section 15 below.

THEOREM 6.3. (Blackwell) *For non-arithmetic F and $a > 0$,*

$$U(t+a) - U(t) \to a/\mu, \qquad \textit{as } t \to \infty.$$

If F is arithmetic with span d, the preceding limit holds with $a = md$ for any integer m.

This theorem says that the renewal function $U(t)$ is asymptotically linear. Note that when $\mu = \infty$, the increments of the renewal function are asymptotically 0. We now address the issue: "Does the asymptotic linearity of $U(t)$ determine a nice limit for functions of the form $U \star h(t)$?"

Before addressing this question, let us investigate the limit of $U \star h(t)$ for a simple piecewise-constant function

$$h(s) = \sum_{k=1}^{m} a_k \mathbf{1}(s \in (s_k, t_k]),$$

where $0 \leq s_1 < t_1 < s_2 < t_2 < \ldots s_m < t_m < \infty$. In this case,

$$\begin{aligned} U \star h(t) &= \int_{[0,t]} h(t-s)dU(s) = \sum_{k=1}^{m} a_k \int_0^t \mathbf{1}(t - s \in (s_k, t_k])dU(s) \\ &= \sum_{k=1}^{m} a_k [U(t-s_k) - U(t-t_k)]. \end{aligned} \tag{6.4}$$

The last equality follows since the integral is over $s \in (t - t_k, t - s_k]$, and $U(t) = 0$ when $t < 0$. By Theorem 6.3, we know

$$U(t - s_k) - U(t - t_k) \to (t_k - s_k)/\mu.$$

Applying this to (6.4) yields

$$\lim_{t\to\infty} U \star h(t) = \frac{1}{\mu} \sum_{k=1}^{m} a_k (t_k - s_k) = \frac{1}{\mu} \int_{\Re_+} h(s)ds. \tag{6.5}$$

This result suggests that a limit of this form would also be true for general functions $h(t)$. That is what we will establish in the next section.

7. Key Renewal Theorem

This section will complete our development of renewal functions and solutions of renewal equations. As above, we will consider a function $U \star h(t)$, typical of those satisfying a renewal equation. The issue is to determine the limit of this function as $t \to \infty$.

To proceed, we will need a subtle condition concerning the integral of the function h on $\Re_+$. For such a function the Riemann integral $\int_0^t h(s)ds$, when it exits, is the limit of the Riemann sums in Definition 2 below on $[0,t]$ as $\delta \to 0$. The integral exists when h is continuous on $[0,t]$, or it has at most a countable number of discontinuity points. Furthermore, the usual Riemann integral of h on $\Re_+$ is defined as the limit

$$\int_{\Re_+} h(s)ds = \lim_{t\to\infty} \int_0^t h(s)ds, \tag{7.1}$$

provided the limit exits.

For our purposes, we need a stronger notion of a Riemann integral defined "directly" on $\Re_+$ as follows.

Definition 2. Similarly to the definition of a Riemann integral on a finite interval, it is natural to approximate the integral of a real-valued function $h(t)$ on the entire domain $\Re_+$ over a grid $0, \delta, 2\delta, \ldots$ by the upper and lower Riemann sums

$$\begin{aligned} I^\delta(h) &= \delta \sum_{k=0}^{\infty} \sup\{h(s) : k\delta \le s < (k+1)\delta\}, \\ I_\delta(h) &= \delta \sum_{k=0}^{\infty} \inf\{h(s) : k\delta \le s < (k+1)\delta\}. \end{aligned}$$

The function $h(t)$ is *directly Riemann integrable* (DRI) if $I^\delta(h)$ and $I_\delta(h)$ are finite for each δ, and they both converge to the same limit as $\delta \to 0$. The limit is necessarily the usual Riemann integral $\int_{\Re_+} h(s)ds$, since this integral is defined by (15.5), where $\int_0^t h(s)ds$ is the limit of the Riemann sums on $[0,t]$.

A DRI function is clearly Riemann integrable in the usual sense, but the converse is not true; see Exercise 30. From the definition, it is clear that $h(t)$ is DRI if it is Riemann integrable and it is 0 outside a finite interval. Also, $h(t)$ is DRI if and only if its positive and negative parts $h^+(t)$ and $h^-(t)$ are both DRI. Further criteria for DRI are given in Proposition 7.8 below and Exercise 31.

We are now ready for the main result.

Theorem 7.2. (Key Renewal Theorem) *If $h(t)$ is DRI and F is non-arithmetic, then*

$$\lim_{t\to\infty} U \star h(t) = \frac{1}{\mu} \int_{\Re_+} h(s)ds. \tag{7.3}$$

PROOF. Fix $\delta > 0$ and define $\overline{h}_k = \sup\{h(s) : k\delta \le s < (k+1)\delta\}$ and

$$\overline{h}(t) = \sum_{k=0}^{\infty} \overline{h}_k \mathbf{1}(k\delta \le t < (k+1)\delta).$$

Define $\underline{h}(t)$ and $\underline{h}_k$ similarly, with sup replaced by inf. Obviously,

$$U \star \underline{h}(t) \le U \star h(t) \le U \star \overline{h}(t). \tag{7.4}$$

Letting $d_k(t) \equiv U(t - k\delta) - U(t - (k+1)\delta)$, we can write (like (6.4))

$$U \star \overline{h}(t) = \sum_{k=0}^{\infty} \overline{h}_k d_k(t).$$

Now $\lim_{t\to\infty} d_k(t) = \delta/\mu$ by Theorem 6.3, and $d_k(t) \le U(\delta)$ by Exercise 27. Then by the dominated convergence theorem (see the Appendix, Theorem 8.7) and the DRI property of h,

$$\begin{aligned}\lim_{\delta\to 0}\lim_{t\to\infty} U \star \overline{h}(t) &= \lim_{\delta\to 0} \frac{\delta}{\mu} \sum_{k=0}^{\infty} \overline{h}_k \\ &= \lim_{\delta\to 0} \frac{1}{\mu} I^{\delta}(h) = \frac{1}{\mu}\int_{\Re_+} h(s)ds.\end{aligned}$$

This (double) limit is the same with the replacement of $\overline{h}(t)$ and $I^{\delta}(h)$ by $\underline{h}(t)$ and $I_{\delta}(h)$. Therefore, the upper and lower bounds in (7.4) for $U \star h(t)$ have the same limit $1/\mu \int_{\Re_+} h(s)ds$, and so $U \star h(t)$ must also have this limit. □

REMARK 7.5. The key renewal theorem is equivalent to Blackwell's renewal theorem. Indeed, the proof above showed that Blackwell's theorem implies the key renewal theorem, and the reverse implication follows by applying (7.3) with $h(t) \equiv \mathbf{1}(0 \le t + a < a)$.

An analogous key renewal theorem for arithmetic F is as follows. It can also be proved by Blackwell's renewal theorem — with fewer technicalities — as suggested in Exercise 29.

THEOREM 7.6. (Arithmetic Key Renewal Theorem) *If F is arithmetic with span d, then for any $u < d$,*

$$\lim_{n\to\infty} U \star h(u + nd) = \frac{d}{\mu}\sum_{k=0}^{\infty} h(u + kd), \tag{7.7}$$

provided the sum exists.

We end this section with criteria for a function to be DRI.

PROPOSITION 7.8. Any one of the following conditions is sufficient for $h(t)$ to be DRI.

(a) $h(t) \ge 0$ is decreasing and is Riemann integrable on $\Re+$.

(b) $h(t)$ is Riemann integrable on $[0, a]$ for each a, and $I^{\delta}(h) < \infty$ for some

$\delta > 0$.
(c) $h(t)$ is continuous except possibly on a set of Lebesgue measure 0, and $|h(t)| \le b(t)$, where $b(t)$ is DRI.

PROOF. Suppose condition (a) holds. Since the usual Riemann integral of h on $\Re+$ exists, we have

$$I_\delta(h) \le \int_{\Re_+} h(s)ds \le I^\delta(h).$$

Also, the decreasing property of $h(t)$ implies $I^\delta(h) - I_\delta(h) = \delta h(0) \to 0$ as $\delta \to 0$. These observations prove $h(t)$ is DRI.

Next, suppose (b) holds. We will write

$$I^\delta(h) = \mathcal{I}^\delta[0, a/\delta) + \mathcal{I}^\delta[a/\delta, \infty),$$

where $\mathcal{I}^\delta[x, \infty) = \delta \sum_{k=\lceil x-1 \rceil}^{\infty} \sup\{h(s) : k\delta \le s < (k+1)\delta\}$. We will use a similar expression for $I_\delta(h)$. Since $h(t)$ is Riemann integrable on $[0, a]$, it follows that $\mathcal{I}^\delta[0, a/\delta)$ and $\mathcal{I}_\delta[0, a/\delta)$ both converge to $\int_0^a h(s)ds$ as $\delta \to 0$. Therefore,

$$I^\delta(h) - I_\delta(h) = o(1) + \mathcal{I}^\delta[a/\delta, \infty) - \mathcal{I}_\delta[a/\delta, \infty), \qquad \text{as } \delta \to 0. \tag{7.9}$$

Let γ be such that $I^\gamma(h) < \infty$. Then for any $\epsilon > 0$, there is a large enough a such that $\mathcal{I}^\gamma[a/\gamma, \infty) < \epsilon$. Then clearly, for sufficiently small δ,

$$\mathcal{I}_\delta[a/\delta, \infty) \le \mathcal{I}^\delta[a/\delta, \infty) \le \mathcal{I}^\gamma[a/\gamma, \infty) < \epsilon.$$

Using this in (7.9), we have

$$I^\delta(h) - I_\delta(h) \le o(1) + 2\epsilon, \qquad \text{as } \delta \to 0.$$

Since this holds for any ϵ, it follows that $h(t)$ is DRI.

Finally, if (c) holds, then (b) is satisfied since $I^\delta(h) \le I^\delta(b)$. Thus $h(t)$ is DRI. □

8. Regenerative Processes

The primary use of the key renewal theorem is in providing limit theorems for regenerative processes and their relatives. Such theorems will be our focus for a while. This section covers regenerative and crudely regenerative processes, and the next three sections cover Markov processes, Markov-renewal processes, and processes with regenerative increments.

Loosely speaking, a continuous- or discrete-time stochastic process is regenerative if there is a renewal process such that the segments of the process between successive renewal times are i.i.d.

We will describe regenerative processes with the following notation. Let $\{X(t) : t \ge 0\}$ denote a continuous-time stochastic process with a state space S that is a metric space (e.g., the Euclidean space $\Re^d$ or a Polish space; see the Appendix). This process need not be a jump process like the Markov processes we have been discussing. However, we assume that the sample paths of $X(t)$ are right-continuous with left-hand limits a.s. This ensures that the sample paths are continuous except possibly on a set of Lebesgue measure 0.

Let $N(t)$ denote a renewal process on $\Re_+$, defined on the same probability space as $X(t)$, with renewal times T_n and inter-renewal times $\xi_n = T_n - T_{n-1}$. For simplicity, assume throughout this section that the inter-renewal distribution F is non-arithmetic and has a finite mean μ.

DEFINITION 3. For the two-dimensional process $\{(N(t), X(t)) : t \geq 0\}$, its sample path in the time interval $[T_{n-1}, T_n)$, called the nth *segment* of the process, is

$$\zeta_n = (\xi_n, \{X(T_{n-1} + t) : 0 \leq t < \xi_n\}). \tag{8.1}$$

The $X(t)$ is a *regenerative process* over the times T_n if its segments ζ_n are independent and identically distributed.

In the preceding definition, the process $X(t)$ is a *delayed* regenerative process if ζ_n are independent, and $\zeta_2, \zeta_3, \dots$ have the same distribution, which is different from the distribution of ζ_1. We discuss more general regenerative-like processes with stationary segments in Section 19.

Classic examples of regenerative processes are ergodic Markov chains and continuous-time ergodic Markov and Markov-renewal processes. We will discuss these examples in the next section. Keep in mind that any theorem for regenerative processes will apply to these Markovian processes.

Regenerative process have the useful "inheritance property" that a function of a regenerative process is also regenerative. Specifically, if $\tilde{X}(t)$ with state space $\tilde{S}$ is regenerative over T_n, then $X(t) = f(\tilde{X}(t))$ is also regenerative over T_n, for any $f : \tilde{S} \to S$. Because of this fact, we can describe the limit of many probabilities and expectations of regenerative processes in terms of real-valued regenerative processes (e.g., $P\{\tilde{X}(t) \in B\} = EX(t)$, where $X(t) = \mathbf{1}(\tilde{X}(t) \in B)$).

We will also use the following regeneration-like property for expectations.

DEFINITION 4. A real-valued process $X(t)$ is *crudely regenerative* at T_1 if

$$E[X(T_1 + t)|T_1] = EX(t), \qquad t \geq 0, \tag{8.2}$$

and these expectations are finite.

If $X(t)$ is regenerative over T_n, then (8.2) holds, but the converse is clearly not true. We will now see that (8.2) is the essential property of regenerative processes needed for applying the key renewal theorem. Keep in mind that throughout this section T_1 is assumed to have a non-arithmetic distribution.

THEOREM 8.3. *Suppose $X(t)$ is a real-valued process that is crudely regenerative at T_1, and define $M \equiv \sup\{|X(t)| : t \leq T_1\}$. If the expectations* *of M and MT_1 are finite, then*

$$\lim_{t\to\infty} EX(t) = \frac{1}{\mu}\int_{\Re_+} h(s)ds, \tag{8.4}$$

where $h(t) = E[X(t)\mathbf{1}(T_1 > t)]$.

Proof. By Remark 5.6, $EX(t) = U \star h(t)$, where $h(t) = E[X(t)\mathbf{1}(T_1 > t)]$. Thus (8.4) will follow by the key renewal theorem provided $h(t)$ is DRI. To prove this, note that $|h(t)| \le b(t) \equiv E[M\mathbf{1}(T_1 > t)]$. Now $b(t) \downarrow 0$ since $EM < \infty$; and $\int_{\Re_+} b(s)ds = E[MT_1] < \infty$. Then $b(t)$ is DRI by Proposition 7.8 (a), and hence $h(t)$ is also DRI by Proposition 7.8 (c). □

We will now apply Theorem 8.3 to characterize the limiting distribution of a regenerative process. For a process $X(t)$ on a countable state space S, a probability measure $\mathrm{P}(\cdot)$ on S is the limiting distribution of $X(t)$ if

$$\lim_{t\to\infty} P\{X(t) \in B\} = \mathrm{P}(B), \qquad B \subset S. \tag{8.5}$$

This definition, however, is too restrictive for uncountable S, where (8.5) is not needed for all subsets B. In particular, when the state space S is the Euclidean space $\Re^d$, then $\mathrm{P}(\cdot)$ on $S = \Re^d$ is defined to be the limiting distribution of $X(t)$ if (8.5) holds for $B \in \mathcal{E}$ (the Borel sets of S) such that $\mathrm{P}(\delta B) = 0$, where δB is the boundary of B. An equivalent definition is that $\mathrm{P}(\cdot)$ on S is the *limiting distribution* of $X(t)$ if

$$\lim_{t\to\infty} E[f(X(t))] = \int_S f(x)\mathrm{P}(dx), \tag{8.6}$$

for any continuous function $f : S \to [0,1]$. This means the distribution of $X(t)$ converges weakly to P (see Section 8 in the Appendix for more details on weak convergence).

Theorem 8.7. *Suppose the process $X(t)$ with a metric state space S (e.g. E^d) is a regenerative process over T_n. If $f : S \to \Re$ is such that the expectations of $M \equiv \sup\{|f(X(t))| : t \le T_1\}$ and MT_1 are finite, then*

$$\lim_{t\to\infty} E[f(X(t))] = \frac{1}{\mu}E\Big[\int_0^{T_1} f(X(s))ds\Big]. \tag{8.8}$$

In particular, the limiting distribution of $X(t)$ is

$$\mathrm{P}(B) \equiv \frac{1}{\mu}E\Big[\int_0^{T_1} \mathbf{1}(X(s) \in B)ds\Big], \qquad B \in \mathcal{E}. \tag{8.9}$$

Proof. Assertion (8.8) follows by Theorem 8.3, since $f(X(t))$ is regenerative over T_n and therefore it satisfies the crude-regeneration property.

In particular, (8.8) holds for any continuous function $f : S \to [0,1]$. Then by (8.6), we know that (8.9) is the limiting distribution of $X(t)$. □

Theorems 8.3 and 8.7 provide a framework for characterizing limits of expectations and probabilities of regenerative processes. For expectations, one must check that the maximum M of the process during an inter-renewal interval has a finite mean. The main step in applying these theorems, however, is to evaluate the integrals $\int_{\Re_+} h(s)ds$ or $\int_S f(x)\mathrm{P}(dx)$. Keep in mind that one need not set up a renewal equation or check the DRI property for each application — these properties have already been verified in the proof of Theorem 8.3.

Theorem 8.7 and most of those to follow, are true, with slight modifications, for delayed regenerative processes. This is due to the property in Exercise 40 that the limiting behavior of a delayed regenerative process is the same as the limiting behavior of the process after its first regeneration time T_1. Here is an immediate consequence of Theorem 8.7 and Exercise 40.

COROLLARY 8.10. (Delayed Regenerations) *Suppose the process $X(t)$ with a metric state space S is a delayed regenerative process over T_n. If $f : S \to \Re$ is such that the expectations of $M \equiv \sup\{|f(X(t))| : T_1 \le t \le T_2\}$ and $M\xi_2$ are finite, then*

$$\lim_{t\to\infty} E[f(X(t))] = \frac{1}{\mu} E\Big[\int_{T_1}^{T_2} f(X(s))\, ds\Big].$$

In particular, the limiting distribution of $X(t)$ is

$$\mathrm{P}(B) \equiv \frac{1}{\mu} E\Big[\int_{T_1}^{T_2} \mathbf{1}(X(s) \in B) ds\Big], \qquad B \in \mathcal{E}.$$

We end this section with applications of Theorem 8.7 to three regenerative processes associated with a renewal process.

DEFINITION 5. *Renewal Process Trinity.* For a renewal process $N(t)$, we define

$$A(t) = t - T_{N(t)}, \quad \text{the } \textit{backward recurrence time} \text{ at } t,$$

(or the *age*), which is the time since the last renewal prior to t;

$$B(t) = T_{N(t)+1} - t, \quad \text{the } \textit{forward recurrence time} \text{ at } t,$$

(or the *residual renewal time*), which is the time to the next renewal after t; and

$$L(t) \equiv \xi_{N(t)+1} = A(t) + B(t), \quad \text{length of the renewal interval covering } t.$$

For instance, a person arriving at a bus stop at time t would have to wait $B(t)$ minutes for the next bus to arrive, or a call-center operator returning to answer calls at time t would have to wait for a time $B(t)$ before the next call. Also, if a person begins analyzing an information string at a location t looking for a certain character (or pattern), then $A(t)$ and $B(t)$ would be the distances to the left and right of t where the next character occurs.

Note that the three-dimensional process $(A(t), B(t), L(t))$ is regenerative over T_n, and so is each process by itself. Each of the processes $A(t)$ and $B(t)$ is a continuous-time Markov process with piece-wise deterministic paths on the state space $\Re_+$; see Exercises 32 and 33. A convenient expression for their joint distribution is, for $0 \le x < t$, $y \ge 0$,

$$P\{A(t) > x, B(t) > y\} = P\{N(t+y) - N((t-x)-) = 0\}. \tag{8.11}$$

This is simply the probability of no renewals in $[t-x, t+y]$. Although this probability is generally not tractable, one can show it is the solution of a renewal equation, and hence it has the form $U \star h(t)$; see Exercises 34 and 35.

One can obtain the limiting distributions of $A(t)$ and $B(t)$ separately from Theorem 8.7. Instead, we will derive their joint limiting distribution. Since $(A(t), B(t))$ is regenerative over T_n, Theorem 8.3 yields

$$\lim_{t\to\infty} P\{A(t) > x, B(t) > y\} = 1 - \frac{1}{\mu}\int_0^{x+y} [1 - F(s)]ds, \tag{8.12}$$

since

$$h(t) = P\{A(t) > x, B(t) > y, T_1 > t\} = P\{T_1 > t + y\}\mathbf{1}(t > x).$$

From (8.12), it immediately follows that

$$\lim_{t\to\infty} P\{A(t) \le x\} = \lim_{t\to\infty} P\{B(t) \le x\} = \frac{1}{\mu}\int_0^{x} [1 - F(s)]ds. \tag{8.13}$$

This limiting distribution is the *equilibrium distribution* associated with F. We will see its significance in Section 15, for stationary renewal processes.

One can also obtain the limiting distribution of $L(t) = A(t) + B(t)$ by Theorem 8.7. Namely,

$$\lim_{t\to\infty} P\{L(t) \le x\} = \frac{1}{\mu}\int_0^{x} sdF(s), \tag{8.14}$$

since

$$h(t) = P\{L(t) \le x, T_1 > t\} = P\{T_1 \le x, T_1 > t\} = (F(x) - F(t))\mathbf{1}(x > t).$$

Alternatively, one can derive (8.14) directly from (8.12).

Additional properties of the three regenerative processes $A(t)$, $B(t)$ and $L(t)$ are in Exercises 32– 39. These processes are especially nice for a Poisson process.

Example 8.15. *Poisson Recurrence Times.* If $N(t)$ is a Poisson process with rate λ, then from (8.11)

$$P\{A(t) > x, B(t) > y\} = e^{-\lambda(x+y)}, \qquad 0 \le x < t, y \ge 0, \tag{8.16}$$

which is the Poisson probability of no renewals in an interval of length $x+y$. In particular, setting $x = 0$, and then $y = 0$, yields

$$P\{B(t) > y\} = e^{-\lambda y}, \qquad P\{A(t) > x\} = e^{-\lambda x}\mathbf{1}(x < t).$$

Thus $B(t)$ is exponentially distributed with rate λ; this also follows by the memoryless property of the exponential distribution. Note that $A(t)$ has the same exponential distribution, but it is truncated at $x = t$. The limiting distribution of each of these processes, however, is exponential with rate λ.

Since $L(t) = A(t) + B(t)$, its distribution can be obtained from (8.16); its mean is shown in Exercise 37.

Even though recurrence time processes $A(t)$ and $B(t)$ are typically not tractable for a fixed t, their equilibrium distribution may be nice.

EXAMPLE 8.17. *Uniform Distribution.* Suppose $N(t)$ is a renewal process with uniform inter-renewal distribution $F(x) = x$, for $x \in [0, 1]$. It associated equilibrium distribution (8.13) is simply $F_e(x) = 2x - x^2$. Interestingly, $F_e(x) \geq F(x)$ for each x. That is, the distribution F_e for the forward recurrence time $B(t)$ in equilibrium is greater than the distribution F of the forward recurrence time $B(0) = \xi_1$ at time 0. This means that $B(t)$ in equilibrium is stochastically smaller than $B(0)$. This is due to fact that the failure rate $F'(x)/(1 - F(x)) = 1/(1 - x)$ of F is increasing. Compare this property with the inspection paradox in Exercise 37.

9. Limiting Distributions for Markov Processes

This section covers the important role of renewal theory in characterizing the limiting distributions of discrete- and continuous-time Markov processes. The results here follow from the limit theorems in the preceding section, which are manifestations of the key renewal theorem for regenerative processes.

For the next two results, assume X_n is an ergodic Markov chain on a countable state space S. For each $i \in S$, let $0 < \nu_1(i) < \nu_2(i) < \dots$ denote the times at which X_n enters state i. The Markov property ensures that X_n is a (discrete-time) delayed regenerative process over $\nu_n(i)$. The following theorem establishes the existence of the limiting distribution

$$\pi_j \equiv \lim_{n\to\infty} P\{X_n = j\}, \qquad j \in S,$$

which does not depend on X_0. Here we write $E_i[\,\cdot\,] \equiv E[\,\cdot\,|X_0 = i]$.

THEOREM 9.1. (Markov Chains) *The ergodic Markov chain X_n has a unique limiting distribution given as follows: for a fixed $i \in S$,*

$$\pi_j = \frac{1}{E_i[\nu_1(i)]} E_i\Big[\sum_{n=0}^{\nu_1(i)-1} \mathbf{1}(X_n = j)\Big], \qquad j \in S. \tag{9.2}$$

Another expression for this probability is

$$\pi_j = \frac{1}{E_j[\nu_1(j)]}, \qquad j \in S. \tag{9.3}$$

PROOF. Since X_n is a delayed regenerative process over $\nu_n(i)$, assertion (9.2) follows by Corollary 8.10. Setting $i = j$ in (9.2) yields (9.3), since the sum in (9.2) is the sojourn time in state j, which is 1. □

Many expected costs or utility values that the Markov chain X_n accrues between successive entrances to a fixed state, e.g., $E_i[\sum_{n=0}^{\nu_1(i)-1} f(X_n)]$ can be expressed in terms of its limiting distribution.

PROPOSITION 9.4. (Cycle Values) *Suppose V_n is a random variable associated with X_n and there is a $g : S \to \Re$ such that*

$$E_i[V_n|X_n, \gamma_n] = g(X_n)\gamma_n, \qquad n \geq 0,$$

where $\gamma_n = \mathbf{1}(\nu_1(i) > n)$. Then, providing the following sums exist,

$$E_i[\sum_{n=0}^{\nu_1(i)-1} V_n] = \pi_i^{-1} \sum_{j \in S} g(j)\pi_j, \qquad i \in S. \tag{9.5}$$

In particular, if Y_n, $n \geq 0$, are i.i.d. random variables (or random elements) independent of the X_n and $f : S \times \Re \to \Re$, then

$$E_i\Big[\sum_{n=0}^{\nu_1(i)-1} f(X_n, Y_n)\Big] = \pi_i^{-1} \sum_{j \in S} E[f(j, Y_0)]\pi_j, \qquad i \in S, \tag{9.6}$$

provided the sum exists.

PROOF. Conditioning the nth term on X_n, γ_n, the left-hand side of (9.5) equals

$$\sum_{j \in S} E_i\Big[\sum_{n=0}^{\infty} E_i[V_n \mathbf{1}(X_n = j)\gamma_n \Big| X_n, \gamma_n]\Big] = \sum_{j \in S} g(j) E_i\Big[\sum_{n=0}^{\nu_1(i)-1} \mathbf{1}(X_n = j)\Big].$$

But this equals the right-hand side of (9.5) by (9.2). □

We will now introduce continuous-time Markov processes and describe their limiting distributions. Consider a continuous-time stochastic process $\{X(t) : t \geq 0\}$ with a countable state space S. Assume the process changes its state at times τ_n that form a simple point process on $\Re_+$. That is,

$$X(t) = X_n \qquad \text{if } \tau_n \leq t < \tau_{n+1} \text{ for some } n, \tag{9.7}$$

where X_n is the state of the process during the time interval $[\tau_n, \tau_{n+1})$. Then $X(t)$ is a *jump process* with piecewise-constant, right-continuous paths.

Assume the states X_n it visits is an ergodic Markov chain on S with transition probabilities p_{ij}. Furthermore, assume that a sojourn time $S_{n+1} \equiv \tau_{n+1} - \tau_n$ in state $X_n = i$ has an exponential distribution with rate q_i, independent of everything else. In other words, (X_n, S_n) is a Markov chain on the space $I \times \Re_+$ with transition probabilities

$$P\{X_{n+1} = j, S_{n+1} \leq t | X_n = i, S_n = s\} = p_{ij}(1 - e^{-q_i t}).$$

Finally, for a fixed state i, denote the first entrance time to state i by $T_1(i) = \inf\{t > \tau_1 : X(t) = i\}$ and assume $\mu_i \equiv E[T_1(i)|X(0) = i]$ (the expected time between entrances to i) is finite.

DEFINITION 6. The process $X(t)$ described above is an *ergodic Markov process*. The sequence of states X_n it visits is the *embedded Markov chain*, and q_i are the rates of the exponential sojourn times in the states. More details on this important class of Markov jump processes will be covered in Chapters 4 and 5.

For a fixed state i, let $T_1(i) < T_2(i) < \dots$ denote the times at which $X(t)$ enters state i. The Markov property of (X_n, S_n) ensures that $X(t)$ is a delayed regenerative process over $T_n(i)$.

The rest of this section is devoted to describing the limiting distribution of the ergodic Markov process $X(t)$ defined above. For this development, we let π_j denote the limiting distribution of X_n described in Theorem 9.1, and define $\mu_i = E_i[T_1(i)]$, the mean time between entrances to i. We begin by giving expressions for μ_i and other cycle values in the spirit of Proposition 9.4.

EXAMPLE 9.8. *More Cycle Values.* Suppose $\{V(t) : t \geq 0\}$ is a cost (or value) process associated with the ergodic Markov process $X(t)$ such that

$$E_i\Big[\int_{\tau_n}^{\tau_{n+1}} V(t)\,dt \Big| X_k,\ k \geq 0\Big] = h(X_n), \qquad n \geq 0,$$

for some $h : S \to \Re$. Then the expected cost between visits of $X(t)$ to state i is

$$E_i\Big[\int_0^{T_1(i)} V(t)\,dt\Big] = \pi_i^{-1} \sum_{j\in S} h(j)\pi_j. \tag{9.9}$$

This follows from Proposition 9.4 since

$$\int_0^{T_1(i)} V(t)\,dt = \sum_{n=0}^{\nu_1(i)-1} \int_{\tau_n}^{\tau_{n+1}} V(t)\,dt.$$

Special cases of (9.9) are

$$E_i\Big[\int_0^{T_1(i)} \mathbf{1}(X(t) = j)\,dt\Big] = \pi_i^{-1}\pi_j q_j^{-1}, \tag{9.10}$$

$$\mu_i = \pi_i^{-1} \sum_{j\in S} \pi_j q_j^{-1}. \tag{9.11}$$

Expression (9.10) follows from (9.9) since

$$\begin{aligned} h(X_n) &= E_i\Big[\int_{\tau_n}^{\tau_{n+1}} \mathbf{1}(X(t) = j)\,dt \Big| X(s),\ s \geq 0\Big] \\ &= \mathbf{1}(X_n = j)E_i[\tau_{n+1} - \tau_n | X_n] \ = \ \mathbf{1}(X_n = j)q_{X_n}^{-1}. \end{aligned}$$

Also, (9.11) follows by (9.10) since

$$\mu_i = E_i[T_1(i)] = \sum_{j\in S} E_i\Big[\int_0^{T_1(i)} \mathbf{1}(X(s) = j)ds\Big].$$

The next result gives three expressions for the limiting distribution

$$p_j \equiv \lim_{t\to\infty} P\{X(t) = j\}$$

of the Markov process $X(t)$. Expressions (9.13) and (9.14) below are useful in theoretical arguments, or for applications when the expressions can be evaluated. The main formula, (9.15), involves only q_i and the limiting distribution π_i for the embedded Markov chain X_n.

THEOREM 9.12. (Markov Processes) *The ergodic Markov process $X(t)$ described above has a unique limiting distribution given as follows: for a fixed $i \in S$,*

$$p_j = \frac{1}{\mu_i} E_i\Big[\int_0^{T_1(i)} \mathbf{1}(X(s) = j)ds\Big], \qquad j \in S. \tag{9.13}$$

In other words,

$$p_j = \frac{1}{q_j \mu_j}, \qquad j \in S. \tag{9.14}$$

Furthermore, this distribution is

$$p_j = \pi_j q_j^{-1} / \sum_{\ell \in S} \pi_\ell q_\ell^{-1}, \qquad j \in S. \tag{9.15}$$

PROOF. Since $X(t)$ is a delayed regenerative process over the $T_n(i)$, and $T_n(i)$ has a non-arithmetic distribution due to the exponential sojourn times in states, the first assertion follows by Corollary 8.10. Then (9.13) and (9.10) yield (9.14). Also, (9.15) follows by (9.14) with $\mu_j = \pi_j^{-1} \sum_{\ell \in S} \pi_\ell q_\ell^{-1}$ from (9.11). □

10. Markov-Renewal Processes

A distinguishing feature of an ergodic Markov jump process is that the sojourn time in a state i has an exponential distribution with rate q_i depending on the state. In many applications, however, the sojourn time in a state has a general non-exponential distribution depending on the state and the next state at the end of the sojourn. We will now describe a class of processes with this feature called Markov-renewal processes, and present limit theorems for them.

Throughout this section, $\{X(t) : t \geq 0\}$ will denote a continuous-time stochastic jump process with a countable state space S that changes its state at times τ_n that form a simple point process on $\Re_+$. Assume the states it visits forms an ergodic Markov chain X_n on S with transition probabilities p_{ij} and limiting distribution π_j. Also, assume the sojourn time $S_{n+1} = \tau_{n+1} - \tau_n$ in state $X_n = i$ ending with a jump to state $X_{n+1} = j$ has a distribution F_{ij}, independent of everything else. In this case, (X_n, S_n) is a Markov chain on the space $I \times \Re_+$ with transition probabilities

$$P\{X_{n+1} = j, S_{n+1} \leq t | X_n = i, S_n = s\} = p_{ij} F_{ij}(t).$$

Next, assume the sojourn time distributions F_{ij} have finite means μ_{ij}. Then the mean sojourn time in state i is

$$E[S_{n+1} | X_n = i] = \sum_{j \in S} p_{ij} \mu_{ij}.$$

Let $T_1(i)$ denote the time of the first entrance to state i. Since $T_1(i)$ is the sum of the sojourn times in all states until i is reached, it follows by arguing as in Example 9.8 (Exercise 45) that

$$\mu_i \equiv E_i[T_1(i)] = \pi_i^{-1} \sum_{j\in S} \pi_j \sum_{\ell\in S} p_{j\ell}\mu_{j\ell}.$$

Assume this is finite. Finally, assume for simplicity that at least one of these distributions F_{ij} is non-arithmetic, which ensures that the distribution of $T_1(i)$ is also non-arithmetic.

Under these assumptions, the process $X(t)$ is an *ergodic Markov-Renewal Process* with transition probabilities p_{ij} and sojourn distributions F_{ij}. (It is sometimes called a semi-Markov process.) Of course, $X(t)$ is an ergodic Markov process when each F_{ij} is an exponential distribution independent of j. A non-Markovian example is a cyclic renewal process.

The Markov-renewal process $X(t)$ is a delayed regenerative process, and so all the results in this chapter for delayed regenerative processes apply to $X(t)$. For instance, the following is analogous to Theorem 9.12 for Markov processes. The proof is Exercise 45.

THEOREM 10.1. (Markov-renewal Processes) *The ergodic Markov-renewal process $X(t)$ has a unique limiting distribution p_j given by (9.13). An alternative expression is*

$$p_j = \frac{1}{\mu_j} \sum_{\ell\in S} p_{j\ell}\mu_{j\ell}, \qquad j \in S. \tag{10.2}$$

Furthermore,

$$p_j = \pi_j \sum_{\ell\in S} p_{j\ell}\mu_{j\ell} \Big/ \sum_{i,\ell\in S} \pi_i p_{i\ell}\mu_{i\ell}, \qquad j \in S. \tag{10.3}$$

These expressions for the limiting distribution are essentially the same form as (9.14) and (9.15) for Markov processes. The only difference is that the mean sojourn time in state i is now $\sum_{j\in S} p_{ij}\mu_{ij}$ instead of q_i^{-1} for Markov processes.

11. Processes with Regenerative Increments

In Section 2, we discussed laws of large numbers for cumulative processes $Z(t)$ associated with renewal processes. We will now continue that discussion for cumulative utility processes (e.g., reward or cost processes) associated with Markovian and regenerative processes.

The focus will be on the following processes.

DEFINITION 7. Let $Z(t)$ be a real-valued process with $Z(0) = 0$ defined on the same probability space as a renewal process $N(t)$. The *increments of $Z(t)$ are regenerative* over T_n if the increments of $(N(t), Z(t))$ after time T_n, namely

$$(\{\xi_k : k \geq n+1\}, \{Z(T_n + t) - Z(T_n) : t \geq 0\}),$$

are independent of the past $\{N(t), Z(t) : t \leq T_n\}$, and the distribution of this set of increments is independent of n. A process with delayed regenerative increments is defined in the obvious way.

Primary examples of such processes are cumulative functionals of regenerative processes, such as $Z(t) = \int_0^t f(X(s))ds$, where $X(t)$ is regenerative over T_n and $f(i)$ is a utility rate when the process $X(t)$ is in state i.

Throughout this section, we assume $Z(t)$ is a process with regenerative increments over T_n such that $\mu \equiv ET_1$ and $a \equiv E[Z(T_1)]$ are finite. Keep in mind that $Z(0) = 0$. Note that $(\xi_n, Z(T_n) - Z(T_{n-1}))$ are i.i.d. This is the main property leading to the following results.

The distribution and mean of $Z(t)$ are generally not tractable for computations. However, we do have a Wald identity for some expectations.

PROPOSITION 11.1. (Wald Identity for Regenerations) *For the process $Z(t)$ with regenerative increments described above,*

$$E[Z(T_{N(t)+1})] = aE[N(t) + 1], \qquad t \geq 0. \tag{11.2}$$

PROOF. Apply Corollary 3.7 to $Z(T_{N(t)+1}) = \sum_{n=0}^{N(t)} [Z(T_{n+1}) - Z(T_n)]$. □

The next result describes the limiting behavior of $t^{-1}Z(t)$. By the classical SLLN, we know

$$n^{-1}Z(T_n) = n^{-1} \sum_{k=1}^{n} \left[Z(T_k) - Z(T_{k-1})\right] \to a, \qquad \text{a.s. as } n \to \infty.$$

This extends to $Z(t)$ as follows.

THEOREM 11.3. *For the process $Z(t)$ with regenerative increments described above, assume the mean of $M_n \equiv \sup_{T_{n-1}<t\leq T_n} |Z(t) - Z(T_{n-1})|$ is finite. Then*

$$t^{-1}Z(t) \to a/\mu, \qquad \textit{a.s. as } t \to \infty.$$

PROOF. It suffices by Theorem 2.7 to show $n^{-1}Z(T_n) \to a$ and $n^{-1}M_n \to 0$. However, we have noted above that $n^{-1}Z(T_n) \to a$, and by the classical SLLN

$$n^{-1}M_n = n^{-1}[\sum_{k=1}^{n} M_k - \sum_{k=1}^{n-1} M_k] \to 0.$$

□

The next result is a special case of Theorem 11.3 for a functional of a regenerative process, where the limiting average is expressible in terms of the limiting distribution of the regenerative process. The convergence of the expected value per unit time is also shown in (11.6); a refinement of this is given by Theorem 16.7.

THEOREM 11.4. *Let $X(t)$ be a regenerative process over T_n with state space $S = \Re^d$, and let P denote the limiting distribution of $X(t)$ given by (8.9), where $\mu = ET_1$. Suppose $f : S \to \Re$ is such that $E[\int_0^{T_1} |f(X(s))|ds]$ and $E|f(\overline{X})|$ are finite, where $\overline{X}$ has the distribution P. Then*

$$\lim_{t\to\infty} t^{-1}\int_0^t f(X(s))ds = E[f(\overline{X})], \qquad a.s.. \tag{11.5}$$

If, in addition, $E[T_1 \int_0^{T_1} |f(X(s))|ds]$ is finite, and T_1 has a non-arithmetic distribution, then

$$\lim_{t\to\infty} t^{-1}E\Big[\int_0^t f(X(s))ds\Big] = E[f(\overline{X})]. \tag{11.6}$$

PROOF. Assertion (11.5) follows since Theorem 11.3 with $Z(t) = \int_0^t f(X(s))ds$ and $EM_n \le E[\int_0^{T_1} |f(X(s))|ds]$ yields $t^{-1}Z(t) \to EZ(T_1)/\mu$; and by expressions (8.9) and (**??**) for P,

$$\begin{aligned} EZ(T_1)/\mu &= \frac{1}{\mu}E\Big[\int_0^{T_1} f(X(s))ds\Big] \\ &= \int_S f(x)dP(x) \;=\; E[f(\overline{X})]. \end{aligned}$$

To prove (11.6), note that $E[f(X(t))] \to E[f(\overline{X})]$ by Theorem 8.7. Then (11.6) follows by the property $t^{-1}\int_0^t g(s)ds \to c$ if $g(t) \to c$. □

REMARK 11.7. *Limiting Averages as Expected Values.* The limit (11.5) as an expected value is a common feature of many SLLN's when $f(X(t)) \xrightarrow{d} f(\overline{X})$. However, limit statements for expectations like (11.6) may not be true for a non-regenerative process.

Here is an example of Theorem 11.4 for Markov processes.

EXAMPLE 11.8. *Functionals of Markov Processes.* Let $X(t)$ denote an ergodic Markov process as in Theorem 9.12 with transition probabilities p_{ij}, exponential sojourn rates q_i and limiting distribution p_i. Let $f(i)$ denote a value (reward, cost utility) per unit time when the process $X(t)$ is in state i. Then by Theorem 11.4, the average value per unit time is

$$\lim_{t\to\infty} t^{-1}\int_0^t f(X(s))ds = \sum_{i\in S} p_i f(i), \quad a.s., \tag{11.9}$$

provided the sum exists.

Next, suppose $g(i,j)$ is a value for the process to jump from state i to state j. Then the value up to time t is

$$Z(t) = \sum_n g(X_{n-1}, X_n)\mathbf{1}(\tau_n \le t),$$

where X_n is the embedded markov chain with limiting distribution π_i. Therefore, by Theorem 11.4 and Proposition 9.4, the average value is

$$\lim_{t\to\infty} t^{-1}Z(t) = \frac{1}{\mu_i}E[Z(T_1)] = \sum_{j\in S} p_j q_j \sum_{\ell\in S} p_{j\ell}\, g(j,\ell), \quad a.s., \tag{11.10}$$

provided the sums exist. For the last equality, see Exercise 44.

REMARK 11.11. *Functionals of Markov-Renewal Processes.* The two limit statements in Example 11.8 for functionals of Markov processes are also valid for Markov-renewal processes. One simply uses (10.3) as the limiting distribution, and uses $1/\mu_{j\ell}$ instead of q_j in (11.10).

12. Average Sojourn Times via Little Laws

This section describes Little laws that apply to a large class of queueing processes and other non-queueing contexts as well. A Little law is essentially a SLLN for a queueing process, which follows by SLLN's like those discussed above.

Consider a general service system or input-output system where discrete items (e.g., customers, jobs, data packets) are processed, or simply visit for a while. The items arrive to the system at times τ_n that form a point process $N(t)$ on $\Re_+$ (it need not be a renewal process). Let W_n denote the total time the nth item spends in the system. In a production system, the waiting or sojourn time W_n includes the item's service time plus any delay waiting in queue for service. The item exits the system at time $\tau_n + W_n$. Then the quantity of items in the system at time t is

$$Q(t) = \sum_{n=1}^{\infty} \mathbf{1}(\tau_n \le t < \tau_n + W_n), \qquad t \ge 0.$$

There are no assumptions concerning the processing or visits of the items or the stochastic nature of the variables W_n and τ_n, other than their existence. For instance, items may arrive and depart in batches, an item may reenter for multiple services, or the items may be part of a larger network that affects their sojourns.

We will consider the following three standard system performance parameters:

$$\begin{aligned}
L &\equiv \lim_{t\to\infty} t^{-1}\int_0^t Q(s)ds && \text{average quantity in the system}\\
\lambda &\equiv \lim_{t\to\infty} t^{-1}N(t) && \text{arrival rate}\\
W &\equiv \lim_{n\to\infty} n^{-1}\sum_{k=1}^{n} W_k && \text{average waiting time}
\end{aligned}$$

There are many diverse systems in which λ and L exist, and the issue is whether W exists. The waiting times W_n typically have a complicated stochastic structure and may not even be regenerative. We will consider the existence of W under the following assumption, which is very natural for most queueing systems.

Empty-System Assumption. Let T_n denote the nth time at which an item arrives to an empty system, i.e., $Q(T_n-) = 0$ and $Q(T_n) > 0$. Assume the times T_n form a point process on $\Re_+$ such that the limit

$$\mu \equiv \lim_{n\to\infty} n^{-1}T_n$$

exists and is positive. This says the system empties out infinitely often, at an asymptotically constant rate.

THEOREM 12.1. (Little Law) *Suppose the system described above satisfies the empty-system assumption. If L and λ exist, then W exists, and $L = \lambda W$.*

PROOF. With no loss in generality, we may assume the system is empty at time 0 and an item arrives. First observe that in the time interval $[0, T_n)$, all of the $\nu_n \equiv N(T_n-)$ items that arrive in the interval also depart by the empty-system time T_n, and their total waiting time is

$$\sum_{k=1}^{\nu_n} W_k = \sum_{k=1}^{\infty} \int_0^{T_n} \mathbf{1}(\tau_k \le s < \tau_k + W_k)ds = \int_0^{T_n} Q(s)ds. \tag{12.2}$$

The first equality follows since the system is empty just prior to T_n, and the second equality follows from the definition of $Q(t)$.

We will apply a discrete-time version of Theorem 2.7 to $Z(n) \equiv \sum_{k=1}^{n} W_k$ and the discrete-time indices ν_n. Since $n^{-1}T_n \to \mu$ (by the empty-system assumption) and L exists, it follows, in light of (12.2), that

$$n^{-1}Z(\nu_n) = (n^{-1}T_n)\Big(T_n^{-1}\int_0^{T_n} Q(s)ds\Big) \to \mu L, \qquad \text{a.s. as } t \to \infty.$$

Also, since $t^{-1}N(t) \to \lambda$, we have

$$n^{-1}\nu_n = T_n^{-1}N(T_n-)(n^{-1}T_n) \to \lambda\mu.$$

Then by Theorem 2.7, $W = \lim_{n\to\infty} n^{-1}Z(n) = L/\lambda$. Thus $L = \lambda W$. $\square$

Here are some examples.

EXAMPLE 12.3. *Regenerative Queueing System.* Suppose the system described above satisfies the empty-system assumption, and $Q(t)$ is a regenerative process over the empty-system times T_n. Assume ET_1 and $E[\int_0^{T_1} Q(s)ds]$ are finite. Then by Theorem 2.7 with $Z(t) = \int_0^t Q(s)ds$, we have $L = E[\int_0^{T_1} Q(s)ds]/ET_1$. Assume the arrival process is a renewal process with a finite mean μ. By the SLLN for renewal processes, the arrival rate is $\lambda = 1/\mu$. Therefore, the average waiting time W exists by Theorem 12.1 and $L = \lambda W$.

In some queueing systems, this Little law for averages is a law of expectations. Specifically, in addition to the preceding assumptions, assume the sequence of sojourn times W_n is regenerative over the discrete times $\nu_n = N(T_n-)$. Since $Q(t)$ and W_n are regenerative,

$$Q(t) \xrightarrow{d} \overline{Q} \text{ as } t \to \infty, \quad \text{and} \quad W_n \xrightarrow{d} \overline{W} \text{ as } n \to \infty,$$

where the distributions of $\overline{Q}$ and $\overline{W}$ are described in Theorem 8.7. Furthermore, by Theorem 11.4,

$$L = \lim_{t\to\infty} t^{-1} \int_0^t Q(s)ds = E\overline{Q},$$

$$W = \lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} W_k = E\overline{W}, \quad \text{a.s.}$$

Also, the renewal arrival rate λ can be represented as $\lambda = E\tilde{N}(1)$, where $\tilde{N}(t)$ is a stationary version of $N(t)$ as described in Theorem 15.3. Then the Little law

$$L = \lambda W \quad \text{is equivalent to} \quad E\overline{Q} = E\tilde{N}(1)E\overline{W}.$$

EXAMPLE 12.4. *$GI/G/1$ Queueing System.* A general example of a regenerative queueing system is the $GI/G/1$ system, where arrivals form a renewal process with mean inter-arrival time $1/\lambda$, the services times are i.i.d., independent of the arrivals, and customers are served by a single server under a first-in-first-out (FIFO) discipline. When the mean service time is less than the mean inter-arrival time, ET_1 and $E[\int_0^{T_1} Q(s)ds]$ are finite, but this will not be verified here. Also, the sojourn times W_n are regenerative over $\nu_n = N(T_n-)$. Then the Little laws in the preceding example are true.

Special cases of the $GI/G/1$ system are a $M/G/1$ system when the arrival process is a Poisson process, a $GI/M/1$ system when the service times are exponentially distributed, and a $M/M/1$ system when the arrivals are Poisson and the service times are exponential.

Theorem 12.1 also yields expected waiting times in Jackson networks, which we discuss in Chapter 5.

There are several Little laws for input-output systems and general utility processes that need not be related to queueing [**32**]. The next result is an elementary but very useful example. For this, we let $X(t)$ be a regenerative process over T_n with state space S. Assume $X(t)$ is a pure jump process (piecewise constant paths, etc.) with a limiting distribution $P(B) = \lim_{n\to\infty} P\{X(t) \in B\}$, which is known. Let B denote a fixed subset of the state space whose complement B^c is not empty. The expected number of times that $X(t)$ enters B between regenerations is

$$\alpha(B) = E\Big[\sum_n \mathbf{1}(X(\tau_{n-1}) \in B^c, X(\tau_n) \in B, \tau_n \in (T_1, T_2]\Big],$$

where τ_n is the time of the nth jump of $X(t)$. The expected number of transitions of $X(t)$ between regenerations is $\alpha(S)$, which we assume if finite.

Consider the average sojourn time of $X(t)$ in B defined by

$$W(B) \equiv \lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} W_k(B)$$

where $W_n(B)$ is its sojourn time in B at its nth visit to the set.

PROPOSITION 12.5. (Sojourns in Regenerative Processes) *For the regenerative process $X(t)$ defined above, its average sojourn time in B exists and equals $W(B) = P(B)\alpha(S)/\alpha(B)$.*

PROOF. Consider $Q(t) = \mathbf{1}(X(t) \in B)$ as an artificial queueing process that only takes values 0 or 1. Clearly $Q(t)$ is regenerative over T_n, since $X(t)$ is regenerative over T_n; and $Q(t)$ satisfies the empty-system assumption. Now, the limiting average of $Q(t)$ is

$$L = \lim_{t\to\infty} t^{-1} \int_0^t \mathbf{1}(X(s) \in B)\, ds = P(B).$$

The arrival rate λ is the rate $\alpha(B)/\alpha(S)$ at which $X(t)$ enters B. Thus, Theorem 12.1 yields $P(B) = \lambda W(B) = (\alpha(B)/\alpha(S))W(B)$, which proves the assertion. $\square$

EXAMPLE 12.6. *Sojourns in Markov Processes.* Suppose the process $X(t)$ in Proposition 12.5 is an ergodic Markov process with a countable state space and limiting distribution p_i. Then its average sojourn time in the set B is

$$W(B) = \sum_{j\in B} p_j \Big/ \sum_{j\in B^c} p_j q_j \sum_{\ell\in B} p_{j\ell},$$

where p_{ij} are the transition probabilities of the embedded chain and q_i is the exponential sojourn rate in state i. The double summation in the denominator equals $\alpha(B)\alpha(S)$ by Exercise 44.

13. Batch-Service Queueing System

This section describes a model that illustrates the use of SLLN's for evaluating averages of queueing processes.

For service systems that process items in batches, a basic problem is to determine when to serve batches and how many items should be in the batches. This is a dynamic control problem or a Markov decision problem. We will address this problem for a particular setting and show how to obtain certain control parameters by using renewal processes.

Consider a single-server station that serves items or customers in batches as follows. Items arrive to the station according to a Poisson process with rate λ and they enter a queue where they wait to be served. The server can serve items in batches, and the number of items in a batch can be any number less than or equal to a fixed number $K \le \infty$ (the service capacity). The service times of the batches are independent, identically distributed and do not depend on the arrival process or the batch size (think of a computer, bus, or truck). Only one batch can be served at a time and, during a service, additional arrivals join the queue.

The server observes the queue length at the times at which an arrival occurs and the server is idle, or whenever a service is completed. At each of these observation times, the server takes one of the following actions:

- No items are served.
- A batch consisting of all or a portion of the items waiting is served (the batch size cannot exceed $i \wedge K$, where i is the queue length).

These actions control the batch sizes and the timing of the services. If the server takes the first action, the next control action is taken when the next item arrives, and if the server takes the second action to serve a batch, the next control action is taken when the service is completed. A *control policy* is a rule for selecting one of these actions at each observation time. The general problem is to find a control policy that minimizes the average cost (or discounted cost) of serving items over an infinite time horizon.

This Markov decision problem was solved in [**13**] for natural holding and service cost functions for both the average-cost and discounted-cost criteria. In either case, the main result is that there is an optimal M-policy of the following form: At each observation time when the queue length is i, do not serve any items if $i < M$, and serve a batch of $i \wedge K$ items if $i \geq M$. Here M is an "optimal" level that is a function of the costs.

We will now describe an optimal level M for a special case. Suppose the system is to operate under the preceding m-policy ($M = m$ for notational convenience), where the capacity K is infinite, and the service times are exponentially distributed with rate γ. Assume there is cost C for serving a batch and a cost hi per unit time for holding i items in the queue.

THEOREM 13.1. *Under the preceding assumptions, the average cost per unit time is minimized by setting the level M to be*

$$M = \min\{m \geq 0 : m(m+1) \geq 2[(\lambda/\gamma)^2 p^m + C\lambda/h]\}, \tag{13.2}$$

where $p = \lambda/(\lambda + \gamma)$.

PROOF. Let $X_m(t)$ denote the number of items in the queue at time t (where m refers to the m-policy). Let T_n denote the time at which the server completes the nth service, and let $N(t)$ denote the associated counting process. For simplicity, assume that a service has just been completed at time 0, and let $T_0 = 0$.

We will show that, under the m-policy with exponential service times, the T_n are renewal times; and the service plus holding cost in $[0, t]$ is

$$Z_m(t) \equiv CN(t) + h\int_0^t X_m(s)\, ds.$$

Next, we will establish the existence of the average cost

$$f(m) \equiv \lim_{t\to\infty} t^{-1} Z_m(t),$$

and then show that $f(m)$ is minimized at the M specified in (13.2).

Let $Q_n \equiv X_m(T_n)$ denote the queue length at the nth service completion time T_n, for $n \geq 0$. Note that Q_n is just the number of arrivals that occur during the nth service period, since all the waiting items are served in the batch. Because of the exponential service times, Q_n, for $n \geq 1$, are i.i.d. with

$$P\{Q_n = i\} = \int_{\Re_+} \frac{(\lambda t)^i e^{-\lambda t}}{i!} \gamma e^{-\gamma t}\, dt = p^i(1-p), \qquad i \geq 0. \tag{13.3}$$

For notational convenience, assume the initial queue length Q_0 has this distribution and is independent of everything else.

Next, observe that the quantity Q_n determines the time $\xi_{n+1} = T_{n+1} - T_n$ until the next service completion. Specifically, if $Q_n \geq m$, then ξ_{n+1} is simply a service time; and if $Q_n = i < m$, then ξ_{n+1} is the time it takes for $m - i$ more item to arrive plus a service time. Since the Q_n are i.i.d., it follows that T_n are renewal times. Furthermore, conditioning on Q_0, the inter-arrival distribution is

$$P\{\xi_1 \leq t\} = P\{Q_0 \geq m\} G_\gamma(t) + \sum_{i=0}^{m-1} P\{Q_0 = i\} G_\lambda^{(m-i)\star} \star G_\gamma(t),$$

where G_λ is an exponential distribution with rate λ.

Then using the distribution (13.3), the inter-arrival distribution and its mean (indexed by m) are:

$$\begin{aligned} P\{\xi_1 \leq t\} &= p^m G_\gamma(t) + (1-p) \sum_{i=0}^{m-1} p^i G_\lambda^{(m-i)\star} \star G_\gamma(t), \\ \mu_m &= \gamma^{-1} + m\lambda^{-1} - (1-p^m)\gamma^{-1}. \end{aligned} \tag{13.4}$$

($\sum_{i=1}^k ip^{i-1} = \frac{d}{dp}(\sum_{i=0}^k p^i)$ is used for the last formula.)

Now, the increasing process $Z_m(t)$ is such that $Z_m(T_n) - Z(T_{n-1})$, for $n \geq 1$, are i.i.d. with mean

$$E[Z_m(T_1)] = C + hE\Big[\int_0^{T_1} X_m(s)\, ds\Big].$$

Then by Theorem 2.7, the average cost, as a function of m, is

$$f(m) \equiv \lim_{t\to\infty} t^{-1} Z_m(t) = \mu_m^{-1} E[Z_m(T_1)].$$

To evaluate this limit, let $\tilde{N}(t)$ denote the Poisson arrival process with exponential inter-arrival times $\tilde{\xi}_n$, and let τ denote an exponential service time with rate γ. Then we can write

$$\int_0^{T_1} X_m(s)\, ds = Q_0\tau + \int_0^\tau \tilde{N}(s)\, ds + \sum_{i=0}^{m-1} \mathbf{1}(Q_0 = i) \sum_{k=1}^{m-i} (i+k-1)\tilde{\xi}_k. \tag{13.5}$$

The first two terms on the right-hand side represent the holding time of items during the service period, and the last term represents the holding time of items (which is 0 if $Q_0 \geq m$) prior to the service period. Then from the independence of Q_0 and τ and Exercise 14,

$$E\Big[\int_0^{T_1} X_m(s)\,ds\Big] = \Big[1/(1-p)\gamma + \lambda/\gamma^2 + (1-p)\lambda^{-1}\sum_{i=0}^{m-1} p^i \sum_{k=1}^{m-i}(i+k-1)\Big].$$

Substituting this in the expression above for $f(m)$, it follows from lengthy algebraic manipulations that

$$f(m+1) - f(m) = h(1-p^{m+1})D_m/(\lambda^2\mu_m\mu_{m+1}),$$

where $D_m = m(m+1) - 2[(\lambda/\gamma)^2 p^m + C\lambda/h]$. Now, D_m is increasing in m and the other terms in the preceding display are positive. Therefore $f(m)$ is monotone decreasing and then increasing and has a unique minimum at $M \equiv \min\{m : D_m \geq 0\}$, which is equivalent to (13.2).

□

Analysis similar to that above yields a formula for the optimal level M when the service capacity K is finite. In this case, the service completion times T_n are "not" renewal times. However, (Q_n, T_n) is a Markov-renewal process. The Q_n is a Markov chain with transition probabilities

$$p_{ij} = \begin{cases} p^j(1-p) & \text{if } i < K \\ p^{j-i-K}(1-p) & \text{if } K \leq i \leq j-K \end{cases}$$

and $p_{ij} = 0$ otherwise. One would then evaluate the average cost by applying a Markov-renewal result as mentioned in Remark 11.11.

14. Central Limit Theorems

For a real-valued process $Z(t)$ with regenerative increments over T_n, we know that under the conditions in Theorem 11.3,

$$Z(t)/t \to a \equiv EZ(T_1)/ET_1 \quad \text{a.s.} \quad \text{as } t \to \infty.$$

In other words, $Z(t)$ behaves asymptotically like at. This section presents a complementary central limit theorem that describes an approximate normal distribution for the difference $Z(t) - at$, when t is large. Special cases are CLT's for renewal and Markovian processes.

As one might suspect, CLT's for regenerative processes are basically applications of the following classical CLT for sums of independent random variables (which is proved in standard probability texts). The analysis in this section involves the notion of convergence in distribution of random variables; further properties of this mode of convergence are discussed in Section 8 in the Appendix.

THEOREM 14.1. (Classical CLT) *Suppose $X_1, X_2, \ldots$ are independent, identically distributed random variables with a finite mean μ and variance σ^2, and define $S_n = \sum_{k=1}^n X_k - n\mu$. Then*

$$P\{S_n/n^{1/2} \leq x\} \to \int_{-\infty}^x \frac{e^{-y^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}\,dy, \qquad x \in \Re.$$

This convergence in distribution is denoted by $S_n/n^{1/2} \xrightarrow{d} N(0,\sigma^2)$, *as* $n \to \infty$, *where* $N(0,\sigma^2)$ *is a normal random variable with mean* 0 *and variance* σ^2.

We will also use the following result for randomized sums; see for instance p.216 in [**10**].

THEOREM 14.2. (Anscombe) *In the context of Theorem 14.1, let* $N(t)$ *be an integer-valued process defined on the same probability space as the* X_n, *where* $N(t)$ *may depend on the* X_n. *If* $t^{-1}N(t) \xrightarrow{d} c$, *where* c *is a positive constant, then*

$$S_{N(t)}/t^{1/2} \xrightarrow{d} N(0, c\sigma^2), \text{ as } t \to \infty.$$

The following is a regenerative analogue of the classical CLT.

THEOREM 14.3. (Regenerative CLT) *Suppose* $Z(t)$ *is a real-valued process with regenerative increments over* T_n *such that* $\mu \equiv ET_1$, ET_1^2, *and* $a \equiv EZ(T_1)/\mu$ *are finite. In addition, let* $M_n \equiv \sup_{T_{n-1}<t\le T_n} |Z(t) - Z(T_{n-1})|$, *and assume* EM_n, $E[T_1M_1]$, *and* $\sigma^2 \equiv Var(Z(T_1) - aT_1)$ *are finite, and* $\sigma > 0$. *Then*

$$(Z(t) - at)/t^{1/2} \xrightarrow{d} N(0, \sigma^2/\mu), \qquad \text{as } t \to \infty. \tag{14.4}$$

PROOF. The process $Z(t)$ is "asymptotically close" to $Z(T_{N(t)})$ because their difference is bounded by $M_{N(t)+1}$, which is a regenerative process that is 0 at regeneration times. Consequently, the normalized process

$$\tilde{Z}(t) \equiv (Z(t) - at)/t^{1/2}$$

should have the same limit as the process

$$Z'(t) \equiv (Z(T_{N(t)}) - aT_{N(t)})/t^{1/2}.$$

Based on this conjecture, we will prove

$$Z'(t) \xrightarrow{d} N(0, \sigma^2/\mu), \qquad \text{as } t \to \infty, \tag{14.5}$$

$$|\tilde{Z}(t) - Z'(t)| \xrightarrow{d} 0, \qquad \text{as } t \to \infty. \tag{14.6}$$

Then it will follow by a standard property of convergence in distribution that

$$\tilde{Z}(t) = Z'(t) + (\tilde{Z}(t) - Z'(t)) \xrightarrow{d} N(0, \sigma^2/\mu).$$

To prove (14.5), note that $Z'(t) = S_{N(t)}/t^{1/2}$ where

$$S_n \equiv \sum_{k=1}^n X_k \equiv \sum_{k=1}^n [Z(T_k) - Z(T_{k-1}) - a\xi_k].$$

Since $Z(t)$ has regenerative increments over T_n, the X_k are i.i.d. with mean 0 and variance σ^2. Also, $t^{-1}N(t) \to 1/\mu$ by the SLLN for renewal processes. In light of these observations, Anscombe's theorem above yields (14.5).

The prove (14.6), we use the inequality

$$|\tilde{Z}(t) - Z'(t)| \le t^{-1/2} X(t) \equiv t^{-1/2}[M_{N(t)+1} + a\xi_{N(t)+1}]. \tag{14.7}$$

Clearly $X(t)$ is regenerative over T_n, and by Theorem 8.7,

$$\lim_{t\to\infty} EX(t) = \frac{1}{\mu} E\Big[\int_0^{T_1} X(s)ds\Big] < \infty.$$

Then by Markov's inequality,

$$P\{t^{-1/2}X(t) > \epsilon\} \le \frac{1}{t^{1/2}\epsilon} EX(t).$$

This bound tends to 0 as $t \to \infty$, and so $t^{-1/2}X(t) \xrightarrow{d} 0$. Applying this limit to (14.7) yields (14.6). □

To apply the preceding result to a regenerative process that satisfies the main assumptions, one only needs to evaluate the normalization constants a and σ. Here are some examples.

EXAMPLE 14.8. *CLT for Renewal Processes.* If $N(t)$ is a renewal process whose inter-renewal distribution has a finite mean μ and variance σ^2, then

$$(N(t) - t/\mu)/t^{1/2} \xrightarrow{d} N(0, \sigma^2/\mu^3), \qquad \text{as } t \to \infty.$$

This follows by Theorem 14.3 with $Z(t) = N(t)$, $a = 1/\mu$, and $\mathrm{Var}(Z(T_1) - aT_1) = \sigma^2\mu^{-2}$.

EXAMPLE 14.9. *CLT for Markov Chains.* Let X_n be an ergodic Markov chain with limiting distribution π_j, $j \in S$. Consider the sum

$$Z_n = \sum_{k=1}^{n} f(X_k), \qquad n \ge 0,$$

where $f(j)$ is a real-valued cost or utility for the process being in state j. For simplicity, fix an $i \in S$ and assume $X_0 = i$ a.s. Then Z_n has regenerative increments over the discrete times ν_n at which X_n enters state i. We will apply a discrete-time version of Theorem 14.3 to Z_n.

Accordingly, assume $\mu_i = E\nu_1$, $E\nu_1^2$, and $E\Big[\max_{1\le n\le \nu_1} |Z_n|\Big]$ are finite. The latter is true when $E\Big[\sum_{n=1}^{\nu_1} |f(X_n)|\Big]$ is finite. In addition, assume

$$a \equiv \frac{1}{\mu_i} E_i[Z_{\nu_1}] = \sum_{j\in S} \pi_j f(j), \quad \text{and} \quad \sigma^2 \equiv \frac{1}{\mu_i}\mathrm{Var}(Z_{\nu_1} - a\nu_1)$$

are finite and $\sigma > 0$. Letting $\tilde{f}(j) = f(j) - a$, Exercise 53 shows that

$$\sigma^2 = E[\tilde{f}(X_0)^2] + 2\sum_{n=1}^{\infty} E[\tilde{f}(X_0)\tilde{f}(X_n)], \tag{14.10}$$

where $P\{X_0 = i\} = \pi_i$. Then Theorem 14.3 (in discrete time) yields

$$(Z_n - an)/n^{1/2} \xrightarrow{d} N(0, \sigma^2), \qquad \text{as } n \to \infty. \tag{14.11}$$

This result also applies to Markov chains in random environments as follows (see Exercise 54 for a related continuous-time version). Suppose

$$Z_n = \sum_{k=1}^{n} f(X_k, Y_k), \qquad n \geq 0,$$

where $f : S \times S' \to \Re$, and Y_k are conditionally independent given X_n $(n \geq 0)$, and $P\{Y_k \in B | X_n, n \geq 0\}$ only depends on X_k and $B \in \mathcal{S}'$. In this setting, the cost or utility $f(X_k, Y_k)$ at time k is partially determined by the auxiliary or environmental variable Y_k. Then the argument above yields the CLT (14.11). In this case, $a = \sum_{j \in S} \pi_j m(j)$ where $m(j) = E[f(j, Y_1)]$, and

$$\sigma^2 = E[(f(X_0, Y_1) - m(X_0))^2] + 2 \sum_{n=1}^{\infty} E[m(X_0, X_n)],$$

where $m(j,k) = E[(f(j, Y_1) - m(j))(f(k, Y_2) - m(k))]$.

15. Stationary Renewal Processes

Recall that a basic property of an ergodic Markov chain is that it is stationary if the distribution of its state at time 0 is its stationary distribution (which is also its limiting distribution). This section addresses the analogous issue of determining an appropriate starting condition for a delayed renewal process so that its increments are stationary in time. Part of the analysis can be viewed as an introduction to the use of time-shift operators for establishing properties of stationary processes.

We begin by defining the notion of stationarity for stochastic processes and point processes. A continuous-time stochastic process $X \equiv \{X(t) : t \geq 0\}$ on a general space is *stationary* if its finite-dimensional distributions are invariant under any shift in time: for each $0 \leq s_1 < \ldots < s_k$ and $t \geq 0$,

$$(X(s_1 + t), \ldots, X(s_k + t)) \stackrel{d}{=} (X(s_1), \ldots, X(s_k)). \tag{15.1}$$

Now, consider a point process $N(t) = \sum_n \mathbf{1}(\tau_n \leq t)$ on $\Re_+$. Another way of representing this process is by the family $N \equiv \{N(B) : B \in \mathcal{B}(\Re_+)\}$, where $N(B) = \sum_n \mathbf{1}(\tau_n \in B)$ is the number of points τ_n in the set B. We also define $B + t = \{s + t : s \in B\}$. The process N is *stationary* (i.e., it has *stationary increments*) if, for any $B_1, \ldots, B_k \in \mathcal{B}(\Re_+)$,

$$(N(B_1 + t), \ldots, N(B_k + t)) \stackrel{d}{=} (N(B_1), \ldots, N(B_k)), \qquad t \geq 0. \tag{15.2}$$

We are now ready to characterize stationary renewal processes. Assume $N(t)$ is a delayed renewal process, where the distribution of ξ_1 is G, and the distribution of ξ_n, $n \geq 2$, is F, which has a finite mean μ. The issue is how to select the initial distribution G such that N is stationary. The answer, according to (iv) below, is to select G to be F_e, which is the limiting distribution of the forward and backward recurrence times for a renewal process with inter-arrival distribution F. The following result also shows that the stationarity of N is equivalent to the stationarity of its forward recurrence time process.

THEOREM 15.3. *The following statements are equivalent.*
(i) *The delayed renewal process N is stationary.*
(ii) *The forward recurrence time process $B(t) = T_{N(t)+1} - t$ is stationary.*
(iii) $EN(t) = t/\mu$, *for* $t \geq 0$.
(iv) $G(x) = F_e(x) \equiv \frac{1}{\mu}\int_0^x [1 - F(s)]ds$.
When these statements are true, $P\{B(t) \leq x\} = F_e(x)$, *for* $t \geq 0$.

EXAMPLE 15.4. Suppose the inter-renewal distribution for the delayed renewal process N is the beta distribution

$$F(x) = 30 \int_0^x y^2(1-y)2\,dy, \qquad x \in [0,1].$$

Then by Theorem 15.3, N is stationary if and only if

$$G(x) = 2x - 5x^4 + 6x^5 - 2x^6, \qquad x \in [0,1],$$

which is the equilibrium distribution F_e associated with F.

One consequence of Theorem 15.3 is that Poisson processes are the only non-delayed renewal processes (whose inter-renewal times have a finite mean) that are stationary.

COROLLARY 15.5. *A renewal process $N(t)$ with no delay is stationary if and only if it is a Poisson process.*

PROOF. By Theorem 15.3, $N(t)$ is stationary if and only if $EN(t) = t/\mu$, $t \geq 0$, which is equivalent to $N(t)$ being a Poisson process by Remark 3.4. $\square$

Here is another useful stationarity property.

REMARK 15.6. If $N(t)$ is a stationary renewal process, then

$$E\Big[\sum_{n=1}^{N(t)} f(T_n)\Big] = \frac{1}{\mu}\int_0^t f(s)ds.$$

This follows by Theorem 3.5 and $EN(t) = t/\mu$.

The rest of this section is devoted to proving Theorem 15.3. To simplify our discussion, we will use time-shift operators to describe stationarity. Specifically, the defining property (15.1) for a process $X(t)$ to be stationary will be expressed as

$$S_t X \stackrel{d}{=} X, \quad t \geq 0,$$

where $S_t X \equiv \{X(s+t) : s \geq 0\}$ is the process X with the time shifted by t. Analogously, the defining property (15.2) for a point process N to be stationary will be expressed as

$$S_t N \stackrel{d}{=} N, \qquad t \geq 0,$$

where $S_tN = \{N(B+t) : B \in \mathcal{B}(\Re_+)\}$ is the process N with the time shifted by t.

We will need two properties of stationary processes. The first deals with functions of stationary processes. From the definition of stationarity, it is clear that if $X(t)$ is stationary, then so is the process $Y(t) = f(X(t))$, for any function f on the state space of X to another space. This inheritance of stationarity also holds under more general functions described as follows.

PROPOSITION 15.7. *Suppose X is stationary and*

$$Y(t) = g(S_tX), \qquad t \geq 0,$$

where g is a function on the space of sample paths of X to some space. Then Y is stationary.

PROOF. By the nature of the time-shift operator S_t, and the stationarity of X, we know $S_{u+t}X = S_u(S_tX) \stackrel{d}{=} S_uX$. Therefore,

$$S_tY = \{g(S_{u+t}X) : u \geq 0\} \stackrel{d}{=} \{g(S_uX) : u \geq 0\} = Y.$$

Thus Y is stationary. □

The second property we need is that the mean value function of a stationary point process is linear.

PROPOSITION 15.8. If N is a stationary point process and $EN(1)$ is finite, then

$$EN(t) = tEN(1), \qquad t \geq 0.$$

PROOF. To see this, consider

$$EN(s+t) = EN(s) + E[N(s+t) - N(s)] = EN(s) + EN(t).$$

This is a linear equation $f(s+t) = f(s)+f(t)$, $s, t \geq 0$. The only measurable function that satisfies this linear equation is $f(t) = ct$ for some c. In our case, $c = f(1) = EN(1)$, and hence $EN(t) = tEN(1)$. □

We are now ready to prove Theorem 15.3 above, which we restate here.

THEOREM 15.9. *The following statements are equivalent.*
(i) *The delayed renewal process N is stationary.*
(ii) *The forward recurrence time process $B(t) = T_{N(t)+1} - t$ is stationary.*
(iii) *$EN(t) = t\mu$, for $t \geq 0$.*
(iv) *$G(x) = F_e(x) \equiv \frac{1}{\mu}\int_0^x [1 - F(s)]ds$.*
When these statements are true, $P\{B(t) \leq x\} = F_e(x)$, for $t \geq 0$.

PROOF. We will show (i) ⇔ (ii), and then (i) ⇒ (iii) ⇒ (iv) ⇒ (ii).

(i) ⇒ (ii): Using $T_n = \inf\{u : N(u) = n\}$, we have

$$\begin{aligned} B(t) &= T_{N(t)+1} - t = \inf\{u - t : N(u) = N(t) + 1\} \\ &= \inf\{t' : N(t' + t) - N(t) = 1\} = \inf\{t' : S_tN(t') = 1\}. \end{aligned}$$

That is, $B(t) = g(S_t N)$, where $g(N) = \inf\{t' : N(t') = 1\}$. Then the stationarity of N implies that B is stationary by Proposition 15.7 (with N in place of X).

(ii) $\Rightarrow$ (i): Since N counts the number of times $B(t)$ jumps upward, we can write

$$S_t N(A) = \sum_{u \in A} \mathbf{1}(S_t B(u) > S_t B(u-)).$$

That is $S_t N = g(S_t B)$, where g maps each sample path $b(t)$ of B to the counting measure $g(b)(A) = \sum_{u \in A} \mathbf{1}(b(u) > b(u-))$. Then the stationarity of B implies that N is stationary by Proposition 15.7 (with N in place of Y).

(i) $\Rightarrow$ (iii): If N is stationary, Proposition 15.8 ensures $EN(t) = tEN(1)$. Also, $EN(1) = 1/\mu$ since $t^{-1}EN(t) \to 1/\mu$ by Proposition 6.1. Therefore, $EN(t) = t/\mu$.

(iii) $\Rightarrow$ (iv): Assume $EN(t) = t/\mu$. Exercise 52 shows $U \star F_e(t) = t/\mu$, and so $EN(t) = U \star F_e(t)$. Another expression for this expectation is

$$EN(t) = \sum_{n=1}^{\infty} G \star F^{(n-1)\star}(t) = G \star U(t).$$

Equating these expressions we have $U \star F_e(t) = G \star U(t)$. Taking the Laplace transform of this equality, we have

$$\hat{U}(\alpha)\hat{F}_e(\alpha) = \hat{G}(\alpha)\hat{U}(\alpha), \tag{15.10}$$

where the hat symbol denotes Laplace transform; e.g., $\hat{G}(\alpha) = \int_0^\infty e^{-\alpha t} dG(t)$. By Remark 3.3, we know $\hat{U}(\alpha) = 1/(1 - \hat{F}(\alpha))$ is positive. Using this in (15.10) yields $\hat{F}_e(\alpha) = \hat{G}(\alpha)$. Since these Laplace transforms uniquely determine the distributions, we obtain $G = F_e$.

(iv) $\Rightarrow$ (ii): By direct computation as in Exercise 35, it follows that

$$P\{B(t) > x\} = 1 - G(t + x) + \int_{[0,t]} [1 - F(t + x - s)] dV(s), \tag{15.11}$$

where $V(t) \equiv EN(t) = G \star U(t)$. Now, the assumption $G = F_e$, along with $U \star F_e(t) = t/\mu$ from Exercise 52, yield

$$V(t) = G \star U(t) = U \star G(t) = U \star F_e(t) = t/\mu.$$

Using this in (15.11), along with a change of variable in the integral, we have

$$P\{B(t) > x\} = 1 - G(t + x) + F_e(x + t) - F_e(x). \tag{15.12}$$

Since $G = F_e$, this expression is simply $P\{B(t) > x\} = 1 - F_e(x)$, $t \geq 0$.

We will use this property to show $B(t)$ is stationary. At this point, one could justify $B(t)$ is a Markov process and then invoke the basic result that a Markov process is stationary if its distribution at any time t is independent of t. Instead, we will prove the stationarity of $B(t)$ directly (which amounts to proving the two properties in the preceding sentence).

To prove $B(t)$ is stationary, it suffices to show, for any s_i, x_i and t, the probability of the event

$$\Gamma_t \equiv \{B(s_1+t) \le x_1, \dots, B(s_k+t) \le x_k\}$$

is independent of t. Conditioning on $B(t)$, we have

$$P\{\Gamma_t\} = \int_0^\infty P\{\Gamma_t | B(t) = x\} dF_e(x).$$

Knowing $B(t) = x$, the delayed renewal process on $[t,\infty)$ behaves like a delayed renewal process starting at time 0 with x as its forward recurrence time. Consequently, $P\{\Gamma_t|B(t) = x\} = P\{\Gamma_0|B(0) = x\}$, independent of t, and so $P\{\Gamma_t\}$ is also independent of t. □

Since a stationary renewal process $N(t)$ has a stationary forward recurrence time process, it seems reasonable that the backward recurrence time process $A(t) = t - T_{N(t)}$ would also be stationary. This is not true, since the distribution of $A(t)$ is not independent of t; in particular, $A(t) = t$, for $t < T_1$. However, there is stationarity in the following sense.

REMARK 15.13. *Stationary Backward Recurrence Time Process.* Suppose the stationary renewal process is extended to the negative time axis with (artificial or virtual) renewals at times $\dots < T_{-1} < T_0 < 0$. One can think of the renewals occurring since the beginning of time at $-\infty$. Consistent with the definition above, the backward recurrence process is

$$A(t) = t - T_n, \qquad \text{if } t \in [T_n, T_{n+1}), \text{ for some } n \in \mathbb{N}.$$

Assuming N is stationary on $\Re_+$, the time $A(0) = T_1$ to the first renewal has the distribution F_e. Then one can show, as we proved (i) ⇔ (ii) in Theorem 15.3, that the process $\{A(t) : t \in \Re\}$ is stationary with distribution F_e.

16. Refined Limit Laws*

We will now describe applications of the key renewal theorem for functions that are not asymptotically constant.

The applications of the renewal theorem we have been discussing are for limits of functions $H(t) = U \star h(t)$ that converge to a constant (i.e., $H(t) = c + o(1)$). However, there are many situation in which $H(t)$ tends to infinity, but the key renewal theorem can still be used to describe limits of the form $H(t) = v(t) + o(1)$ as $t \to \infty$, where the function $v(t)$ is the asymptotic value of $H(t)$.

For instance, a SLLN $Z(t)/t \to b$ suggests $EZ(t) = bt + c + o(1)$ might be true, where the constant c gives added information on the convergence. In this section, we discuss such limit theorems.

We first note that an approach for considering limits $H(t) = v(t)+o(1)$ is simply to consider a renewal equation for the function $H(t)-v(t)$ as follows.

LEMMA 16.1. *Suppose $H(t) = U \star h(t)$ is a solution of a renewal equation for a non-arithmetic distribution F, and $v(t)$ is a real-valued function on $\Re$ that is bounded on finite intervals and is 0 for negative t. Then*

$$H(t) = v(t) + \frac{1}{\mu} \int_{\Re_+} \overline{h}(s)\, ds + o(1), \quad \text{as } t \to \infty, \tag{16.2}$$

provided $\overline{h}(t) \equiv h(t) - v(t) + F \star v(t)$ is DRI. In particular, for a linear function $v(t) = bt$,

$$H(t) = bt + \frac{b(\sigma^2 + \mu^2)}{2\mu} + \frac{1}{\mu} \int_{\Re_+} (h(s) - b\mu)\, ds + o(1), \quad \text{as } t \to \infty, \tag{16.3}$$

where σ^2 is the variance of F, provided $h(t) - b\mu$ is DRI.

PROOF. Clearly $H - v$ satisfies the renewal equation

$$H - v = (h - v + F \star v) + F \star (H - v).$$

Then $H - v = U \star \overline{h}$ by Proposition 5.4, and its limit (16.2) is given by the key renewal theorem.

Next, suppose $v(t) = bt$ and $h(t) - b\mu$ is DRI. Then using $\mu = \int_{\Re_+} [1 - F(x)]\, dx$ and the change of variable $x = t - s$ in the integral below, we have

$$\begin{aligned} \overline{h}(t) &= h(t) - bt + b \int_0^t F(t - s)\, ds \\ &= h(t) - b\mu + bg(t). \end{aligned}$$

where $g(t) = \int_t^\infty [1 - F(x)]\, dx$. Now $g(t)$ is continuous and decreasing and

$$\int_0^\infty g(t)\, dt = \frac{1}{2} \int_{\Re_+} t^2 dF(t) = \frac{\sigma^2 + \mu^2}{2}. \tag{16.4}$$

Then $g(t)$ is DRI by Proposition 7.8 (a), and hence $\overline{h}(t) = h(t) - b\mu + bg(t)$ is DRI. Thus, by what we already proved, (16.2) is true but it reduces to (16.3) in light of (16.4). □

Our first use of the preceding remark is a refinement of the result $t^{-1}U(t) \to 1/\mu$ from Proposition 6.1.

PROPOSITION 16.5. *If $N(t)$ is a renewal process whose inter-renewal times have a non-arithmetic distribution with mean μ and variance σ^2, then*

$$U(t) = t/\mu + (\sigma^2 + \mu^2)/2\mu^2 + o(1), \qquad \text{as } t \to \infty.$$

PROOF. This follows by Lemma 16.1 with $H(t) = U(t)$, $h(t) = 1$, and $v(t) = t/\mu$ (here $h(t) - b\mu = 0$ being DRI is not an issue). □

We will now apply Lemma 16.1 to a rather general stochastic process. Suppose T_n are renewal times whose inter-renewal distribution F is non-arithmetic with finite mean μ and variance σ^2. Consider a real-valued process $Z(t)$ that has regenerative increments over T_n, or, more generally, has *crude regenerative increments at* T_1 in the sense that

$$E[Z(T_1+t)-Z(T_1)|T_1] = EZ(t), \qquad t \geq 0. \tag{16.6}$$

In addition, assume the sample paths of $Z(t)$ are right-continuous with left-hand limits a.s.

THEOREM 16.7. *For the process $Z(t)$ defined above, let*

$$M \equiv \sup\{|Z(T_1)-Z(t)| : t \leq T_1\}.$$

If the expectations of M, MT_1, T_1^2, $|Z(T_1)|$, and $\int_0^{T_1} |Z(s)|ds$ are finite, then

$$EZ(t) = at/\mu + a(\sigma^2+\mu^2)/2\mu^2 + c + o(1), \qquad \textit{as } t \to \infty, \tag{16.8}$$

where $a = EZ(T_1)$ and $c = \frac{1}{\mu}E\Big[\int_0^{T_1} Z(s)ds - T_1 Z(T_1)\Big]$.

PROOF. Because $Z(t)$ has crude regenerative increments, we anticipate that $t^{-1}EZ(t) \to a/\mu$. So to prove (16.8), we will apply Lemma 16.1 with $v(t) = at/\mu$.

The first step is to derive a renewal equation for $EZ(t)$. Conditioning on T_1,

$$EZ(t) = E[Z(t)\mathbf{1}(T_1 > t)] + \int_0^t E[Z(t)|T_1 = s]dF(s).$$

Using $E[Z(t)|T_1 = s] = EZ(t-s) + E[Z(s)|T_1 = s]$, from assumption (16.6), and some algebra, it follows that the preceding is a renewal equation $H = h + F \star H$, where $H(t) \equiv EZ(t)$ and

$$h(t) = a + E\Big[(Z(t) - Z(T_1))\mathbf{1}(T_1 > t)\Big].$$

Now, by Lemma 16.1 for $v(t) = at/\mu$, we have

$$EZ(t) = at/\mu + \frac{\sigma^2+\mu^2}{2\mu^2} + \frac{1}{\mu}\int_{\Re_+} g(s)\,ds + o(1), \quad \text{as } t \to \infty, \tag{16.9}$$

provided $g(t) \equiv h(t) - a = E\Big[(Z(t)-Z(T_1))\mathbf{1}(T_1 > t)\Big]$ is DRI. Clearly

$$|g(t)| \leq b(t) \equiv E[M\mathbf{1}(T_1 > t)].$$

Now, $b(t) \downarrow 0$ and

$$\int_{\Re_+} b(s)ds = E\Big[\int_0^{T_1} M\,ds\Big] = E[MT_1] < \infty.$$

Then $b(t)$ is DRI by Proposition 7.8 (a). Hence $g(t)$ is also DRI by Proposition 7.8 (c). Finally, observe that

$$\int_{\Re_+} g(t)\,dt = E\Big[\int_0^{T_1} Z(s)ds - T_1 Z(T_1)]\Big].$$

Substituting this formula in (16.9) proves (16.8). □

Here is a refinement of the Markov process limit in Example 11.8.

EXAMPLE 16.10. Suppose $X(t)$ is an ergodic Markov process on a countable state space S, and let p_i denote its limiting distribution. Assume the time between entrances to a fixed state i has a finite variance σ_i^2; the mean of this time, by (9.14), is $\mu_i = 1/q_i p_i$, where q_i is the rate of the exponential sojourn time in state i. Let $f : S \to \Re$ be such that $\int_0^t E|f(X(s))|ds$ and $\sum_{i\in S} |f(i)|p_i$ are finite. Then Theorem 16.7 for $Z(t) = \int_0^t f(X(s))ds$ says

$$\begin{aligned} E\Big[\int_0^t f(X(s))ds\Big] &= [t + (\sigma_i^2 + \mu_i^2)/2\mu_i] \sum_{i\in S} f(i)p_i \\ &\quad - \sum_{i\in S} f(i)p_i/q_i + o(1), \quad \text{as } t \to \infty. \end{aligned} \tag{16.11}$$

Here, the constants in Theorem 16.7 are $a = \mu \sum_{i\in S} f(i)p_i$ and

$$c = -\frac{1}{\mu} E\Big[\int_0^{T_1} s f(X(s))ds\Big] = -\sum_{i\in S} f(i)p_i/q_i.$$

The first equality follows by calculus manipulations and the second equality follows by (9.9) in Example 9.8 and $p_i = 1/q_i\mu_i$.

As a particular case of (16.11), the expected time spent in state i has the asymptotic behavior

$$E\Big[\int_0^t \mathbf{1}(X(s) = i)ds\Big] = tp_i + p_i[(\sigma^2 + \mu^2)/2\mu - 1/q_i] + o(1), \quad \text{as } t \to \infty.$$

17. Terminating Renewal Processes*

In this section, we discuss renewal processes that terminate after a random number of renewals. Analysis of these terminating (or transient) renewal processes can be done with renewal equations and the key renewal theorem applied a little differently than above.

Consider a sequence of renewal times T_n with inter-arrival distribution F. Suppose that at each time T_n (including $T_0 = 0$), the renewals terminate with probability $1 - p$, or continue until the next renewal epoch with probability p. These events are independent of the preceding renewal times, but may depend on the future renewal times.

Under these assumptions, the total number of renewals ν over the entire time horizon $\Re_+$ has the distribution

$$P\{\nu \ge n\} = p^n, \qquad n \ge 0,$$

and $E\nu = p/(1-p)$. Now, the number of renewals in $[0, t]$ is given by

$$N(t) = \sum_{n=1}^{\infty} \mathbf{1}(T_n \le t, \nu \ge n), \qquad t \ge 0.$$

Of course $N(t) \to \nu$ a.s. Another quantity of interest is the time T_ν at which the renewals terminate.

An equivalent formulation of this terminating renewal process is to assume $N(t)$ counts renewals in which the independent inter-renewal times have an *improper* distribution $G(t)$, with $p \equiv G(\infty) < 1$. Then $1 - p = 1 - G(\infty)$ is the probability that an inter-renewal time is "infinite", which terminates the renewals, and p is the probability of another renewal. In this setting, $G(t) = pF(t)$, where F as described above is the conditional distribution of an inter-renewal time given that it is allowed (or is finite).

Similarly to renewal processes, we will address issues about the process $N(t)$ with the use of its *renewal function*

$$V(t) \equiv \sum_{n=0}^{\infty} G^{n\star}(t) = \sum_{n=0}^{\infty} p^n F^{n\star}(t).$$

Note that $V(t) \to 1/(1-p)$ as $t \to \infty$ (Example 17.7 below describes the convergence rate); this differs from the infinite limit for regular renewal functions.

We first observe that the distribution and mean of the counting process, and of the termination time (which is finite a.s.) are

$$\begin{aligned} &P\{N(t) \geq n\} = G^{n\star}(t), \qquad && EN(t) = V(t) - 1, \\ &P\{T_\nu \leq t\} = (1-p)V(t), \qquad && ET_\nu = p\mu/(1-p). \end{aligned} \tag{17.1}$$

To establish these formulas, recall that the events $\nu = n$ (to terminate at n) and $\nu > n$ (to continue to the $n+1$st renewal) are assumed to be independent of $T_1, \ldots, T_n$. Then

$$\begin{aligned} P\{N(t) \geq n\} &= P\{\nu \geq n, T_n \leq t\} = p^n F^{n\star}(t) = G^{n\star}(t), \\ EN(t) &= \sum_{n=1}^{\infty} P\{N(t) \geq n\} = V(t) - 1. \end{aligned}$$

Similarly, using the independence and $T_\nu = \sum_{n=0}^{\infty} \mathbf{1}(\nu = n) T_n$,

$$\begin{aligned} P\{T_\nu \leq t\} &= \sum_{n=0}^{\infty} P\{\nu = n, T_n \leq t\} = (1-p) \sum_{n=0}^{\infty} p^n F^{n\star}(t), \\ ET_\nu &= \sum_{n=1}^{\infty} P\{\nu = n\} ET_n = \mu p/(1-p). \end{aligned}$$

We will now discuss limits of certain functions associated with the terminating renewal process. As in Proposition 5.4, it follows that $H(t) = V \star h(t)$ is the unique solution to the renewal equation

$$H(t) = h(t) + G \star H(t).$$

We will consider the limiting behavior of $H(t)$ for the case in which the limit

$$h(\infty) \equiv \lim_{t \to \infty} h(t)$$

exists, which is common in applications. Since $V(t) \to 1/(1-p)$ and $h(t)$ is bounded on compact sets and converges to $h(\infty)$, it follows by dominated convergence that

$$
\begin{aligned}
(17.2)\quad H(t) &= h\star V(t) = h(\infty)V(t) + \int_{[0,t]} [h(t-s) - h(\infty)]dV(s) \\
&= h(\infty)/(1-p) + o(1), \qquad \text{as } t \to \infty.
\end{aligned}
$$

The next result describes the rate of this convergence under a few more technical conditions. Assume there is a positive β such that

$$\int_{\Re_+} e^{\beta t} dG(t) = 1.$$

The existence of a unique β is guaranteed under the weak condition that $g(\beta) \equiv \int_{\Re_+} e^{\beta t} dG(t)$ is finite for $\beta \geq 0$. Indeed, since g is continuous and increasing to ∞, it must equal 1 for some β. We also assume the distribution

$$F^{\#}(t) \equiv \int_{[0,t]} e^{\beta s} dG(s)$$

is non-arithmetic and has a mean $\mu^{\#}$.

THEOREM 17.3. *In addition to the preceding assumptions, assume the function $e^{\beta t}[h(t) - h(\infty)]$ is DRI. Then*

$$H(t) = h(\infty)/(1-p) + ce^{-\beta t}/\mu^{\#} + o(e^{-\beta t}), \quad \textit{as } t \to \infty, \tag{17.4}$$

where $c = \int_{\Re_+} e^{\beta s}(h(s) - h(\infty))\, ds - h(\infty)/\beta$.

PROOF. Multiplying the renewal equation $H = h + G \star H$ by $e^{\beta t}$ yields the renewal equation $H^{\#} = h^{\#} + F^{\#} \star H^{\#}$ where $H^{\#}(t) = e^{\beta t} H(t)$ and $h^{\#}(t) = e^{\beta t} h(t)$.

We can now describe the limit of $H(t) - h(\infty)/(1-p)$ by the limit of $H^{\#}(t) - v(t)$ where $v(t) \equiv e^{\beta t} h(\infty)/(1-p)$. From Lemma 16.1, we know

$$H^{\#}(t) = v(t) + \frac{1}{\mu^{\#}} \int_{\Re_+} \overline{h}(s)\, ds + o(1), \quad \text{as } t \to \infty, \tag{17.5}$$

provided $\overline{h}(t) \equiv h^{\#}(t) - v(t) + F^{\#} \star v(t)$ is DRI. In this case,

$$\overline{h}(t) = e^{\beta t}[h(t) - h(\infty)] - \Big[\frac{h(\infty)e^{\beta t}}{1-p}(p - G(t))\Big]. \tag{17.6}$$

Now, the first term on the right-hand side is DRI by assumption. Also,

$$e^{\beta t}(p - G(t)) \leq \int_{(t,\infty)} e^{\beta s} dG(s) = 1 - F^{\#}(t).$$

This bound is decreasing to 0 and its integral is $\mu^{\#}$, and so the last term in brackets in (17.6) is DRI. Thus $\overline{h}(t)$ is DRI. Finally, an easy check shows that $\int_{\Re_+} \overline{h}(s)\, ds = c$, the constant in (17.4). Substituting this in (17.5) and dividing by $e^{\beta t}$ yields (17.4). $\square$

EXAMPLE 17.7. Under the assumptions preceding Theorem 17.3,

$$\begin{aligned} V(t) &= 1/(1-p) - e^{-\beta t}/(\beta\mu^{\#}) + o(e^{-\beta t}), \\ P\{T_\nu > t\} &= (1-p)e^{-\beta t}/(\beta\mu^{\#}) + o(e^{-\beta t}), \quad \text{as } t \to \infty. \end{aligned}$$

The first line follows by Theorem 17.3 with $h(t) = 1$, since by its definition, $V(t) = 1 + G \star V(t)$. The second follows from the first line and (17.1).

EXAMPLE 17.8. *Waiting Time for a Gap in a Poisson Process.* Consider a Poisson process with rate λ that terminates at the first time a gap of size $\geq c$ occurs. That is, the termination time is T_ν, where $\nu = \min\{n : \xi_{n+1} \geq c\}$, where $\xi_n = T_n - T_{n-1}$ and T_n are the occurrence times of the Poisson process. Now, at each time T_n, the process either terminates if $\xi_{n+1} \geq c$, or it continues until the next renewal epoch if $\xi_{n+1} < c$. These events are clearly independent of $T_1, \ldots, T_n$.

Then the probability of terminating is

$$1 - p \equiv P\{\xi_{n+1} \geq c\} = e^{-\lambda c}.$$

The conditional distribution of the next renewal period beginning at T_n is

$$F(t) = P\{\xi_{n+1} \leq t | \xi_{n+1} < c\} = p^{-1}(1 - e^{-\lambda t}), \qquad 0 \leq t \leq c.$$

Then from (17.1), the waiting time for a gap of size c has a mean $E[T_\nu] = (e^{\lambda c} - 1)/\lambda$, and its distribution is

$$P\{T_\nu \leq t\} = e^{-\lambda c} V(t).$$

Now, assume $\lambda c > 1$. Then the condition $\int_{\Re_+} e^{\beta t} p dF(t) = 1$ discussed above for defining β reduces to

$$\lambda e^{(\beta - \lambda)c} = \beta, \qquad \text{for } \lambda < \beta.$$

Using this formula and integration by parts, we have

$$\mu^{\#} = \int_{[0,c]} t e^{\beta t} p dF(t) = p(c\beta - 1)/(\beta - \lambda).$$

Then by Example 17.7,

$$P\{T_\nu > t\} = \Big(\frac{1 - \beta/\lambda}{1 - \beta c}\Big) e^{-\beta t} + o(e^{-\beta t}), \quad \text{as } t \to \infty.$$

EXAMPLE 17.9. *Cramér-Lundberg Risk Model.* Consider an insurance company that receives capital at a constant rate c from insurance premiums, investments, interest etc. The company uses the capital to pay claims that arrive according to a Poisson process $N(t)$ with rate λ. The claim amounts $X_1, X_2, \ldots$ are independent, identically distributed positive random variables with mean μ, and are independent of the arrival times. Then the company's capital at time t is

$$Z_x(t) = x + ct - \sum_{n=1}^{N(t)} X_n, \qquad t \geq 0,$$

where x is the capital at time 0.

An important performance parameter of the company is the probability

$$R(x) \equiv P\{Z_x(t) \geq 0, t \geq 0\},$$

that the capital does not go negative (the company is not ruined). We are interested in approximating this survival probability when the initial capital x is large. Exercise 24 shows that $R(x) = 0$, regardless of the initial capital x, when $c < \lambda\mu$ (the capital input rate is less than the payout rate).

We will now consider the opposite case $c > \lambda\mu$. Conditioning on the time and size of the first claim, one can show (e.g., see [**14, 28, 29**]) that $R(x)$ satisfies a certain differential equation whose corresponding integral equation is the renewal equation

$$R(x) = R(0) + R \star G(x), \tag{17.10}$$

where $R(0) = 1 - \lambda\mu/c$ and

$$G(y) = \lambda c^{-1} \int_0^y P\{X_1 > u\}\, du.$$

The G is a defective distribution with $G(\infty) = \lambda\mu/c < 1$. Then applying (17.2) to $R(x) = h \star V(x) = R(0)V(x)$, we have

$$R(x) \to R(0)/(1 - \lambda\mu/c) = 1, \qquad \text{as } x \to \infty.$$

We now consider the rate at which the "ruin" probability $1 - R(x)$ converges to 0 as $x \to \infty$. Assume there is a positive β such that

$$\lambda c^{-1} \int_{\Re_+} e^{\beta x} P\{X_1 > x\}\, dx = 1,$$

and that

$$\mu^{\#} \equiv \lambda c^{-1} \int_{\Re_+} x e^{\beta x} P\{X_1 > x\}\, dx < \infty.$$

Then by Theorem 17.3 (with $R(x)$, $R(0)$ in place of $H(t)$, $h(t)$), the probability of ruin has the asymptotic form

$$1 - R(x) = \frac{1}{\beta\mu^{\#}}(1 - \lambda\mu/c)e^{-\beta x} + o(e^{-\beta x}), \quad \text{as } x \to \infty.$$

18. Sketch of the Proof of Blackwell's Theorem*

This section describes a coupling proof of Blackwell's theorem. The classical proof of Blackwell's theorem based on analytical properties of the renewal function and integral equations is in Feller (1971). Lindvall (1977) and Athreya, McDonald and Ney (1978) gave another probabilistic proof involving "coupling" techniques. A nice review of various applications of coupling is in Lindvall (1992). A recent refinement of the coupling proof is given in Durrett (2005). The following is a sketch of his presentation when the inter-renewal time has a finite mean (he gives a different proof for the case of an infinite mean).

Let $N(t)$ be a renewal process with renewal times T_n whose inter-renewal times ξ_n have a non-arithmetic distribution and a finite mean μ. For convenience, we will write Blackwell's theorem (Theorem 6.3) as

$$\lim_{t\to\infty} E[N(t,t+a]] = a/\mu, \tag{18.1}$$

where $N(t,t+a] \equiv N(t+a)-N(t)$. Now, this statement would trivially hold if $N(t)$ were a stationary renewal process, since in this case $E[N(t,t+a]]$ would equal a/μ by Proposition 15.8. So if one could construct a version of $N(t)$ that approximates a stationary process as close as possible, then (18.1) would be true. That is the approach in the proof, which we now describe.

On the same probability space as $N(t)$, let $N'(t)$ be a stationary renewal process with renewal times T'_n, whose inter-renewal ξ'_n times for $n \geq 2$ have the same distribution as the ξ_n. The first and most subtle part of the proof is to construct a third renewal process $N''(t)$ on the same probability space that is equal in distribution to the original process $N(t)$ and approximates the stationary process $N'(t)$.

Specifically, for a fixed $\epsilon > 0$, the proof begins by defining random indices ν and ν' such that $|T_\nu - T_{\nu'}| < \epsilon$. Then a third renewal process $N''(t)$ is defined (on the same probability space) with inter-renewal times $\xi_1, \ldots, \xi_\nu, \xi'_{\nu'}, \xi'_{\nu'+1} \ldots$. This process has the following properties:

(a) $\{N''(t) : t \geq 0\} \stackrel{d}{=} \{N(t) : t \geq 0\}$ (i.e., their finite-dimensional distributions are equal).

(b) On the event $\{T_\nu \leq t\}$,

$$N'(t+\epsilon, t+a-\epsilon] \leq N''(t,t+a] \leq N'(t-\epsilon, t+a+\epsilon]. \tag{18.2}$$

This is an *epsilon-coupling* in that $N''(t)$ is a coupling of $N(t)$ that is within ϵ of the targeted stationary version $N'(t)$ in the sense of condition (b).

With this third renewal process in hand, the rest of the proof is as follows. Consider the expectation

$$E[N(t,t+a]] = E[N''(t,t+a]] = V_{\mathbf{1}}(t) + V_2(t), \tag{18.3}$$

where

$$V_{\mathbf{1}}(t) = E\Big[N''(t,t+a]\mathbf{1}(T_\nu \leq t)\Big], \qquad V_2(t) = E\Big[N''(t,t+a]\mathbf{1}(T_\nu > t)\Big].$$

Condition (b) and $E[N'(c,d]] = (d-c)/\mu$ (due to the stationarity) ensure

$$V_{\mathbf{1}}(t) \leq E\Big[N'(t-\epsilon, t+a+\epsilon]\mathbf{1}(T_\nu \leq t)\Big] \leq (a+2\epsilon)\mu.$$

Next, observe $E[N''(t,t+a]|T_\nu > t] \leq E[N''(a)]$, since the worse-case scenario is that there is a renewal at t. This and condition (b) yield

$$V_2(t) \leq P\{T_\nu > t\}E[N''(a)].$$

Similarly,

$$\begin{aligned} V_{\mathbf{1}}(t) \ &\geq \ E\Big[N'(t+\epsilon, t+a-\epsilon] - N''(t,t+a]\mathbf{1}(T_\nu > t)\Big] \\ &\geq \ (a-2\epsilon)/\mu - P\{T_\nu > t\}E[N''(a)]. \end{aligned}$$

Here we take $\epsilon < a/2$, so that $t + \epsilon < t + a - \epsilon$. Combining the preceding inequalities with (18.3), and using $P\{T_\nu > t\} \to 0$ as $t \to \infty$, it follows that

$$(a - 2\epsilon)/\mu + o(1) \le E[N(t, t + a]] \le (a + 2\epsilon)/\mu + o(1).$$

Since this is true for arbitrarily small ϵ, we obtain $E[N(t, t + a]] \to a/\mu$, which is Blackwell's result.

19. Stationary-Cycle Processes*

Most of the results above for regenerative processes also apply to a wider class of regenerative-like processes that we will now describe.

For this discussion, suppose $\{X(t) : t \ge 0\}$ is a continuous-time stochastic process with a general state space S, and $N(t)$ is a renewal process on the same probability space. As in Section 8, we let

$$\zeta_n = (\xi_n, \{X(T_{n-1} + t) : 0 \le t < \xi_n\})$$

denote the segment of these processes on the interval $[T_{n-1}, T_n)$. Then $\{\zeta_{n+k} : k \ge 1\}$ is the *future of* $(N(t), X(t))$ *beginning at time* T_n. This is what an observer of the processes would see beginning at time T_n.

DEFINITION 8. The process $X(t)$ is a *stationary-cycle process* over the times T_n if the future $\{\zeta_{n+k} : k \ge 1\}$ of $(N(t), X(t))$ beginning at any time T_n is independent of $T_1, \ldots, T_n$, and the distribution of this future is independent of n. Discrete-time and delayed stationary-cycle processes are defined similarly.

The defining property ensures that the segments ζ_n form a stationary sequence, whereas for a regenerative process, the segments are i.i.d. Also, for a regenerative process $X(t)$, its future $\{\zeta_{n+k} : k \ge 1\}$ beginning at any time T_n is independent of the entire past $\{\zeta_k : k \le n\}$ (rather than only $T_1, \ldots, T_n$ as in the preceding definition).

All the strong laws of large numbers in this chapter for regenerative processes also hold for stationary-cycle processes. The only difference is that a law's limiting value would be random instead of a constant when stationary segments of the process does not satisfy a technical ergodic property. We will not get into this topic.

As in Section 11, one can define processes with stationary-cycle increments. Most of the results above such as the CLT have obvious extensions to these more complicated processes.

We end this section by by commenting on limiting theorems for probabilities and expectations of stationary-cycle processes.

REMARK 19.1. Theorem 8.7 and Corollary 8.10 are also true for stationary-cycle processes. This follows since such a process satisfies the crude-regeneration property in Theorem 8.3 leading to Theorem 8.7 and Corollary 8.10.

There are many intricate stationary-cycle processes that arise naturally from systems that involve stationary and regenerative phenomena. Here is an elementary illustration.

EXAMPLE 19.2. *Regenerations in a Stationary Environment.* Consider a process $X(t) = g(Y(t), Z(t))$ where $Y(t)$ and $Z(t)$ are independent processes and g is a function on their product space. Assume $Y(t)$ is a regenerative process over the times T_n (e.g., an ergodic Markov process) with state space $S = \Re^d$ and limiting distribution P, as in Theorem 8.7. Assume $Z(t)$ is a stationary process. One can regard $X(t) = g(Y(t), Z(t))$ as a regenerative-stationary reward process, where $g(y, z)$ is the reward rate from operating a system in state y in environment z. Now, the segments ζ_n defined above form a stationary process, and hence $X(t)$ is a stationary-cycle process.

In light of Remark 19.1, we can describe the limiting behavior of $X(t)$ as we did for regenerative processes. Specifically, assuming for simplicity that g is real-valued and bounded, we have

$$\lim_{t\to\infty} EX(t) = \int_{\Re^d} E[g(y, Z(0))]P(dy). \tag{19.3}$$

Indeed, from Theorem 8.7 for stationary-cycle processes,

$$\lim_{t\to\infty} EX(t) = \frac{1}{\mu} E\Big[\int_0^{T_1} g(Y(s), Z(s))ds\Big].$$

Then conditioning on $\{N(t), Y(t) : t \geq 0\}$, which is independent of the process $Z(t)$, the last expectation equals the integral in (19.3).

Also, as in Theorem 11.4 it follows that, for any real-valued bounded function f on the state space of $X(t)$,

$$\lim_{t\to\infty} \int_0^t E[f(X(s))], ds = \int_{\Re^d} E[f(g(y, Z(0)))]P(dy).$$

20. Exercises

1. For a nonnegative random variable X with distribution F, show

$$EX = \int_{\Re_+} (1 - F(x))dx.$$

This reduces to $EX = \sum_{n=1}^{\infty} P\{X \geq n\}$ for integer-valued X. One approach is to use $EX = \int_{\Re_+} \big(\int_0^x dy\big) dF(x)$. For a general random variable X with finite mean, use $X = X^+ - X^-$ to obtain

$$EX = \int_{\Re_+} (1 - F(x))dx - \int_{-\infty}^0 F(x)dx.$$

2. An exponential random variable X satisfies the *memoryless property*

$$P\{X > s + t | X > s\} = P\{X > t\}, \qquad s, t, > 0.$$

Prove the analogue $P\{X > \tau + t | X > \tau\} = P\{X > t\}$, for $t > 0$, where τ is a positive random variable independent of X. Show that, for a Poisson process $N(t)$ with rate λ, the forward recurrence time $B(t) = T_{N(t)+1} - t$ at time t has an exponential distribution with rate λ. One approach is to condition on $T_{N(t)}$.

Consider the forward recurrence time $B(\tau)$ at a random time τ independent of the Poisson process. Show that $B(\tau)$ also has an exponential distribution with rate λ.

3. A system consists of two components with independent lifetimes X_1 and X_2, where X_1 is exponentially distributed with rate λ, and X_2 has a uniform distribution on $[0,1]$. The components operate in parallel, and the system lifetime is $\max\{X_1, X_2\}$ (the system is operational if and only if at least one component is working). When the system fails, it is replaced by another system with an identical and independent lifetime, and this is repeated indefinitely. The number of system renewals over time forms a renewal process $N(t)$. Find the distribution and mean of the system lifetime. Find the distribution and mean of $N(t)$ (reduce your formulas as much as possible). Determine the portion of time that (a) two components are working, (b) only type 1 component is working, and (c) only type 2 component is working.

4. *Continuation.* In the context of the preceding exercise, a typical system initially operates for a time $Y = \min\{X_1, X_2\}$ with two components and then operates for a time $Z = \max\{X_1, X_2\} - Y$ with one component. Thereupon it fails. Find the distributions and means of Y and Z. Find the distribution of Z conditioned that $X_1 > X_2$. You might want to use the memoryless property of the exponential distribution in Exercise 2. Find the distribution of Z conditioned that $X_2 > X_1$.

5. Let $N(t)$ denote a renewal process with inter-renewal distribution F and consider the number of renewals $N(T)$ in an interval $(0, T]$ for some random time T independent of $N(t)$. For instance, $N(T)$ might represent the number of customers that arrive at a service station during a service time T. Find general expressions for the mean and distribution of $N(T)$. Evaluate these expressions for the case in which T has an exponential distribution with rate μ and $F = G^{2\star}$, where G is an exponential distribution with rate λ.

6. Consider the cyclic renewal process $X(t)$ described in Example 1.9, where $F_0, \dots, F_{K-1}$ are the sojourn distributions in states $0, 1, \dots, K-1$. Assume $p \equiv F_0(0) > 0$, but $F_i(0) = 0$, for $i = 1, \dots, K-1$. Let T_n denote the times at which the process $X(t)$ jumps from state $K-1$ directly to state 1 (i.e., it spends no time in state 0). Justify that the T_n form a delayed renewal process with inter-renewal distribution

$$F(t) = p \sum_{j=0}^{\infty} F_1 \star \cdots \star F_{K-1} \star \tilde{F}^{j\star}(t),$$

where $\tilde{F}(t) = \tilde{F}_0 \star F_1 \star \cdots \star F_{K-1}(t)$, and $\tilde{F}_0(t)$ is the conditional distribution of the sojourn time in state 0 given it is positive. Specify a formula for $\tilde{F}_0(t)$, and describe what $\tilde{F}(t)$ represents.

7. *Large Inter-renewal Times*. Let $N(t)$ denote a renewal process with inter-renewal distribution F. Of interest are occurrences of inter-renewal times that are greater than a value c, assuming $F(c) < 1$. Let $\tilde{T}_n$ denote the subset of times T_n for which $\xi_n > c$. So $\tilde{T}_n$ equals some T_k if $\xi_k > c$. (Example 17.8 addresses a related problem of determining the waiting time for a gap of size c in a Poisson process.) Show that $\tilde{T}_n$ are delayed renewal times and the inter-renewal distribution has the form

$$\tilde{F}(t) = \sum_{k=0}^{\infty} F_c^{k\star} \star G(t),$$

where $F_c(t) = F(t)/F(c)$, $0 \leq t \leq c$ (the conditional distribution of an inter-renewal time given that it is $\leq c$), and specify the distribution $G(t)$ as a function of F.

8. *Bernoulli Process*. Consider a sequence of independent Bernoulli trials in which each trial results in a success or failure with respective probabilities p and $q = 1 - p$. Let $N(t)$ denote the number of successes in t trials, where t is an integer. Show that $N(t)$ is a discrete-time renewal process, called a Bernoulli Process. (The parameter t may denote discrete-time or any integer referring to sequential information.) Justify that the inter-renewal times have the geometric distribution $P\{\xi_1 = n\} = q^{n-1}p$, $n \geq 1$. Find the distribution and mean of $N(t)$, and do the same for the renewal time T_n. Show that the moment generating function of T_n is

$$E[e^{\alpha T_n}] = \Big(\frac{pe^{\alpha}}{1 - qe^{\alpha}}\Big)^n, \qquad 0 < \alpha < -\log q. \tag{20.1}$$

9. *Partitioning and Thinning of a Renewal Process.* Let $N(t)$ be a renewal process with inter-renewal distribution F. Suppose each renewal time is independently assigned to be a type i renewal with probability p_i, for $i = 1, \dots, m$, where $p_1 + \cdots + p_m = 1$. Let $N_i(t)$ denote the number of type i renewals up to time t. These processes form a partition of $N(t)$ in that $N(t) = \sum_{i=1}^m N_i(t)$. Each $N_i(t)$ is a thinning of $N(t)$, where p_i is the probability that a point of $N(t)$ is assigned to $N_i(t)$.

Show that $N_i(t)$ is a renewal process with inter-renewal distribution

$$F_i(t) = \sum_{k=1}^{\infty} (1 - p_i)^{k-1} p_i F^{k\star}(t).$$

Show that, for $n = n_1 + \cdots n_m$,

$$\begin{aligned} &P\{N_1(t) = n_1, \dots, N_m(t) = n_m\} \\ &\qquad = [F^{(n)\star}(t) - F^{(n+1)\star}(t)] \frac{n!}{n_1! \cdots n_m!} p_1^{n_1} \cdots p_m^{n_m}. \end{aligned}$$

Give an example of F with $m = 2$ for which $N_1(t)$ and $N_2(t)$ are not independent.

10. *Multi-type Renewals.* An infinite number of jobs are to be processed one-at-a-time by a single server. There are m types of jobs, and the probability that any job is of type i is p_i, where $p_1 + \cdots + p_m = 1$. The service time of a type i job has a distribution F_i with mean μ_i. The service times and types of the jobs are independent. Let $N(t)$ denote the number of jobs completed by time t. Show that $N(t)$ is a renewal process and specify its inter-arrival distribution and mean. Let $N_i(t)$ denote the number of type i jobs processed up to time t. Show that $N_i(t)$ is a delayed renewal process and specify $\lim_{t\to\infty} t^{-1}N(t)$.

11. *Continuation.* In the context of Exercise 10, let $X(t)$ denote the type of job being processed at time t. Find the limiting distribution of $X(t)$. Find the portion of time devoted to type i jobs.

12. *Continuation.* Consider the multi-type renewal process with two types of renewals that have exponential distributions with rates λ_i, and type i occurs with probability p_i, , $i = 1, 2$. Show that the renewal function has the density

$$U'(t) = \frac{\lambda_1\lambda_2 + p_1p_2(\lambda_1 - \lambda_2)^2 e^{-(p_1\lambda_2+p_2\lambda_1)t}}{p_1\lambda_2 + p_2\lambda_1}, \qquad t > 0.$$

13. *System Availability.* The status of a system is represented by an alternating renewal process $X(t)$, where the mean sojourn time in a working state 1 is μ_1 and the mean sojourn time in a non-working state 0 is μ_0. The system *availability* is measured by the portion of time it is working, which is $\lim_{t\to\infty} t^{-1}\int_0^t X(s)ds$. Determine this quantity and show that it is equal to the *cycle-availability* measured by $\lim_{n\to\infty} T_n^{-1}\int_0^{T_n} X(s)ds$.

14. *Integrals of Renewal Processes.* Suppose $N(t)$ is a renewal process with renewal times T_n and $\mu = ET_1$. Prove

$$E\Big[\int_0^{T_n} N(s)\,ds\Big] = \mu n(n-1)/2.$$

For any non-random $t > 0$, it follows by Fubini's theorem that

$$E\Big[\int_0^{t} N(s)\,ds\Big] = \int_0^t EN(s)\,ds.$$

Assuming τ is an exponential random variable with rate γ, prove

$$E\Big[\int_0^{\tau} N(s)\,ds\Big] = \int_{\Re_+} e^{-\gamma t}EN(t)\,dt. \tag{20.2}$$

Show that if N is a Poisson process with rate λ, then the expectation in (20.2) is λ/γ^2. (Integrals like these are used to model holding costs; see Section 13 and the next exercise.)

15. *Continuation.* Items arrive to a service station according to a Poisson process $N(t)$ with rate λ. The items are stored until m have accumulated. Then the m items are served in a batch. The service time is exponentially

distributed with rate γ. During the service, items continue to arrive. There is a cost hi per unit time of holding i customers in the system. Assume the station is empty at time 0. Let C_1 denote the expected cost of holding the customers until m have arrived, and let C_2 denote the expected costs for holding the added arrivals in the system during the service. Determine these expected costs.

16. Customers arrive to a service system according to a Poisson process with rate λ. The system can only serve one customer at a time and, while it is busy serving a customer, arriving customers are blocked from getting service (they may seek service elsewhere or simply go unserved). Assume the service times are independent with common distribution G and are independent of the arrival process. For instance, a contractor may only be able to handle one project at a time (or a vehicle can only transport one item at a time).

Determine the following quantities: (a) The portion of time the system is busy. (b) The portion of time the system is not busy. (c) The number of customers per unit time that are served. (d) The portion of customers that are blocked from service.

17. *Delayed Renewals.* A point process $N(t)$ is a *m-step delayed renewal process* if the inter-occurrence times ξ_{m+k}, for $k \geq 1$, are independent with a common distribution F, and no other restrictions are placed on $\xi_1, \ldots, \xi_m$. That is, $N_m(t) \equiv N(t) - N(T_m)$, for $t \geq T_m$ is a renewal process. Show that Corollary 2.5 and Theorem 2.7 hold for such processes. Use the fact that $N(t)$ is asymptotically equivalent to $N_m(t)$ in that

$$N_m(t)/N(t) = 1 - N(T_m)/N(t) \to 1, \quad \text{a.s. as } t \to \infty.$$

18. For a point process $N(t)$ that is not simple, prove $t^{-1}N(t) \to 1/\mu$ as $t \to \infty$ implies $n^{-1}T_n \to \mu$, as $n \to \infty$. Hint: For a fixed positive constant c, note that $N((T_n - c)^+) \leq n \leq N(T_n)$. Divide these terms by T_n and take limits as $n \to \infty$ to obtain (2.3).

19. *Age Replacement Model.* An item (e.g., battery, vehicle, tool, or electronic component) whose use is needed continuously is replaced whenever it fails or reaches age a, whichever comes first. The successive items are independent and have the same lifetime distribution G. The cost of a failure is c_f dollars and the cost of a replacement at age a is c_r. Show that the average cost per unit time is

$$C(a) = [c_f G(a) + c_r(1 - G(a))] / \int_0^a (1 - G(s)] ds.$$

Find the optimal age a that minimizes this average cost.

20. *Point Processes as Jump Processes.* Consider a point process $N(t) = \sum_{k=1}^{\infty} \mathbf{1}(T_k \leq t)$, where $T_1 \leq T_2 \leq \cdots$. It can also be formulated as an integer-valued jump process of the form

$$N(t) = \sum_{n=1}^{\infty} \nu_n \mathbf{1}(\hat{T}_n \leq t),$$

where $\hat{T}_n$ are the "distinct" times at which $N(t)$ takes a jump, and ν_n is the size of the jump. That is, $\hat{T}_n = \min\{T_k : T_k > \hat{T}_{n-1}\}$, where $\hat{T}_0 = 0$, and $\nu_n = \sum_{k=1}^{\infty} \mathbf{1}(T_k = \hat{T}_n)$, $n \geq 1$.

For instance, suppose T_n are times at which customers arrive to a service system. Then $\hat{T}_n$ are the times at which batches of customers arrive to a service system, and at time $\hat{T}_n$, a batch of ν_n customers arrive. Suppose $\hat{T}_n$ are renewal times, and ν_n are i.i.d. and independent of the $\hat{T}_n$. Show that the number of customers that arrive per unit time is $E[\nu_1]/E[\hat{T}_1]$ a.s., provided these expectations are finite. Next, assume $\hat{T}_n$ form a Poisson process with rate λ, and ν_n has a Poisson distribution. Find $EN(t)$ by elementary reasoning, and then show $N(t)$ has a Poisson distribution.

21. *Batch Renewals.* Consider times $T_n = \sum_{k=1}^{n} \xi_k$, where the ξ_k are i.i.d. with distribution F, where $F(0) = P\{\xi_k = 0\} > 0$. The associated point process $N(t)$ is a renewal process with *instantaneous renewals* (or batch renewals). In the notation of Exercise 20, $N(t) = \sum_{n=1}^{\infty} \nu_n \mathbf{1}(\hat{T}_n \leq t)$, where ν_n is the number of renewals exactly at time $\hat{T}_n$. Specify the distribution of ν_n. Are the ν_n i.i.d.? Are they independent of $\hat{T}_n$? Specify the distribution of $\hat{T}_1$ in terms of F.

22. Prove $E[T_{N(t)}] = \mu E[N(t) + 1] - E[\xi_{N(t)}]$. If $N(t)$ is a Poisson process, show that $E[T_{N(t)}] < \mu E[N(t)]$.

23. *Superpositions of Renewal Processes.* Let $N_1(t)$ and $N_2(t)$ be independent renewal processes with the same inter-renewal distribution, and consider the sum $N(t) = N_1(t) + N_2(t)$ (sometimes called a *superposition*). Suppose $N(t)$ is a renewal process. Prove $N(t)$ is a Poisson process if and only $N_1(t)$ and $N_2(t)$ are Poisson processes.

24. *Production-Inventory Model.* Consider a production-inventory system that produces a product at a constant rate of c units per unit time and the items are put in inventory to satisfy demands. The products may be discrete or continuous (e.g., oil, chemical). Demands occur according to a Poisson process $N(t)$ with rate λ, and the demand quantities $X_1, X_2, \ldots$ are independent, identically distributed positive random variables with mean μ, and are independent of the arrival times. Then the inventory level at time t would be

$$Z_x(t) = x + ct - \sum_{n=1}^{N(t)} X_n, \qquad t \geq 0,$$

where x is the initial inventory level. Consider the probability $R(x) \equiv P\{Z_x(t) \geq 0, t \geq 0\}$ of never running out of inventory. Show that if $c < \lambda\mu$, then $R(x) = 0$ no matter how high the initial inventory level x is. Hint: apply a SLLN to show that $Z_x(t) \to -\infty$ as $t \to \infty$ if $c < \lambda\mu$, where x is fixed. Find the limit of $Z_x(t)$ as $t \to \infty$ if $c > \lambda\mu$. (The process $Z_x(t)$ is a classical model of the capital of an insurance company; see Example 17.9.)

25. Derive a renewal equation that $H(t) = E[N(t) - N(t-a)\mathbf{1}(a \le t)]$ satisfies.

26. *Non-homogeneous Renewals.* Suppose $N(t)$ is a point process on $\Re_+$ whose inter-point times $\xi_n = T_n - T_{n-1}$ are independent with distributions F_n. Prove $EN(t) = \sum_{n=1}^{\infty} F_1 \star \cdots \star F_n(t)$.

27. *Subadditivity of Renewal Function.* Prove the subadditivity property

$$U(t+a) \le U(a) + U(t), \qquad a, t \ge 0.$$

Use $a \le T_{N(a)+1}$ in the expression

$$N(t+a) - N(a) = \sum_{k=1}^{\infty} \mathbf{1}(T_{N(a)+k} \le t + a).$$

28. *Elementary Renewal Theorem via Blackwell.* Prove the elementary renewal theorem (Theorem 6.1) by an application of Blackwell's theorem. One approach, for non-arithmetic F, is to use

$$EN(t) = \sum_{k=1}^{\lceil t \rceil} [U(k) - U(k-1)] + EN(\lceil t \rceil) - EN(t).$$

Then use the fact $n^{-1} \sum_{k=1}^{n} c_k \to c$ when $c_k \to c$.

29. *Arithmetic Key Renewal Theorem.* Represent $U \star h(u+nd)$ as a sum like (6.4), and then prove Theorem 7.6 by applying Blackwell's theorem.

30. Prove the following function is Riemann integrable, but not DRI:

$$h(t) = \sum_{n=1}^{\infty} a_n \mathbf{1}(n - \epsilon_n \le t < n + \epsilon_n),$$

where $a_n \to \infty$ and $1/2 > \epsilon_n \downarrow 0$ such that $\sum_{n=1}^{\infty} a_n \epsilon_n < \infty$.

31. Prove that a continuous function $h(t) \ge 0$ is DRI if and only if $I^{\delta}(h) < \infty$ for some $\delta > 0$.

The next eight exercises concern the renewal process trinity: the backward and forward recurrence times $A(t) = t - T_{N(t)}$, $B(t) = T_{N(t)+1} - t$, and the length $L(t) = \xi_{N(t)+1} = A(t) + B(t)$ of the renewal interval containing t. Assume the inter-renewal distribution is non-arithmetic.

32. Draw a typical sample path for each of the processes $A(t)$, $B(t)$, and $L(t)$.

33. Prove $B(t)$ is a Markov process by showing it satisfies the Markov property

$$P\{B(t+u) \le y | B(s) : s < t, B(t) = x\} = P\{B(u) \le y | B(0) = x\}, \quad x, y, t, u \ge 0.$$

34. Formulate a renewal equation that $P\{B(t) > x\}$ satisfies.

35. *Bypassing a renewal equation.* Use Remark 5.6 (without using a renewal equation) to prove

$$P\{B(t) > x\} = \int_{[0,t]} [1 - F(t + x - s)]dU(s).$$

36. Prove $EB(t) = \mu E[N(t) + 1] - t$. Assuming F has a finite variance σ^2, prove

$$\lim_{t\to\infty} EA(t) = \lim_{t\to\infty} EB(t) = \frac{\sigma^2 + \mu^2}{2\mu}.$$

Is this limit also the mean of the limiting distribution $F_e(t) = \frac{1}{\mu}\int_0^t [1 - F(s)]ds$ of $A(t)$ and $B(t)$?

37. *Inspection Paradox.* Consider the length $L(t) = \xi_{N(t)+1}$ of the renewal interval at any time t (this is what an inspector of the process would see at time t). Prove the paradoxical result that $L(t)$ is *stochastically larger* than the length ξ_1 of a typical renewal interval; that is

$$P\{L(t) > x\} \geq P\{\xi_1 > x\}, \qquad t, x \geq 0. \tag{20.3}$$

This inequality is understandable upon observing that the first probability is for the event that a renewal interval bigger than x "covers" t, and this is more likely to happen than a fixed renewal interval being bigger than x. A consequence of this result is $EL(t) \geq E\xi_1$, which is often a strict inequality.

Suppose $\mu = ET_1$ and $\sigma^2 = \mathrm{Var}T_1$ are finite. Recall from (8.14) that the limiting distribution of $L(t)$ is $\frac{1}{\mu}\int_0^x sdF(s)$. Derive the mean of this distribution (as a function of μ and σ^2), and show it is $\geq \mu$.

Show that if $N(t)$ is a Poisson process with rate λ, then

$$EL(t) = \lambda^{-1}(2 - (1 + \lambda t)e^{-\lambda t}).$$

In this case, $EL(t) > E\xi_1$.

38. Prove

$$\lim_{t\to\infty} P\{A(t)/L(t) \leq x\} = x, \qquad 0 \leq x \leq 1.$$

Prove this result with $B(t)$ in place of $A(t)$.

39. Prove

$$\lim_{t\to\infty} E[A(t)^k B(t)^\ell (A(t) + B(t))^m] = \frac{E[T_1^{k+\ell+m}]}{\mu(k + \ell + 1)\binom{k+\ell}{k}}.$$

Find the limiting covariance, $\lim_{t\to\infty} \mathrm{Cov}(A(t), B(t))$.

40. *Delayed Versus Non-delayed Regenerations.* Let $\tilde{X}(t)$ be a real-valued, bounded, delayed regenerative process over T_n. Then $X(t) \equiv \tilde{X}(T_1 + t)$, $t \geq 0$ is a regenerative process. Show that if $\lim_{t\to\infty} EX(t)$ exists (such as

by Theorem 8.7), then $E\tilde{X}(t)$ has the same limit. Take the limit as $t \to \infty$ of

$$E\tilde{X}(t) = \int_{[0,t]} E[X(t-s)]dF(s) + E[\tilde{X}(t)\mathbf{1}(T_1 > t)].$$

41. *Dispatching System.* Items arrive at a depot (warehouse or computer) at times that form a renewal process with finite mean μ between arrivals. Whenever M items accumulate, they are instantaneously removed (dispatched) from the depot. Let $X(t)$ denote the number of items in the depot at time t. Find the limiting probability that there are p_j items in the system ($j = 0, \ldots, M-1$). Find the average number of items in the system over an infinite time horizon.

Suppose the batch size M is to be selected to minimize the average cost of running the system. The relevant costs are a cost C for dispatching the items, and a cost h per unit time for holding an item in the depot. Let $C(M)$ denote the average dispatching plus holding cost for running the system with batch size M. Find an expression for $C(M)$. Show that the value of M that minimizes $C(M)$ is an integer adjacent to the value $M^* = \sqrt{2C/h\mu}$.

42. *Continuation.* In the context of the preceding exercise, find the average time W that a typical item waits in the system before being dispatched. Find the average waiting time $W(i)$ in the system for the ith arrival in the batch.

43. Consider an ergodic Markov chain X_n with limiting distribution π_i. Prove

$$\lim_{n\to\infty} P\{X_n = j, X_{n+1} = \ell\} = \pi_j p_{j\ell}.$$

One can show that $(X_, X_{n+1})$ is a two-dimensional Markov chain that is ergodic with the preceding limiting distribution. However, establish the limit above only with the knowledge that X_n has a limiting distribution.

44. For the Markov process $X(t)$ in Theorem 9.12, show that

$$E_i\Big[\sum_{n=0}^{\infty} g(X(\tau_{n-1}), X(\tau_n))\mathbf{1}(\tau_n \le T_1(i))\Big] = \pi_i^{-1} \sum_{j\in I} \pi_j \sum_{\ell\in I} p_{j\ell}\, g(j,\ell)$$
$$= \frac{1}{p_i q_i} \sum_{j\in I} p_j q_j \sum_{\ell\in I} p_{j\ell}\, g(j,\ell).$$

45. Suppose $X(t)$ is the ergodic Markov-renewal process as in Theorem 10.1. Arguing as in Example 9.8, show that

$$\begin{aligned} E_i\Big[\int_0^{T_1(i)} \mathbf{1}(X(t)=j)\,dt\Big] &= \pi_i^{-1}\pi_j \sum_{\ell\in S} p_{j\ell}\mu_{j\ell}, \\ E_i[T_1(i)] &= \pi_i^{-1}\sum_{j\in S}\pi_j \sum_{\ell\in S} p_{j\ell}\mu_{j\ell} \\ E_i\Big[\int_0^{T_1(i)} f(X(t))\,dt\Big] &= \pi_i^{-1}\sum_{j\in I}\pi_j f(j) \sum_{\ell\in I} p_{j\ell}\mu_{j\ell} \\ &. \end{aligned}$$

Use these formulas to prove Theorem 10.1.

46. For the Markov process $X(t)$ in Example 16.10, show that

$$E\Big[\int_0^{T_1} s\mathbf{1}(X(s)=i)ds\Big] = p_i/q_i.$$

47. Items with volumes $V_1, V_2, \dots$ are loaded on a truck one at a time until the addition of an arriving item would exceed the capacity V of the truck. Then the truck leaves to deliver the items. The number of items that can be loaded in the truck before its volume V is exceeded is $N(V) = \min\{n : \sum_{k=1}^n V_k > V\}$. Assume the V_n are independent with identical distribution F that has a mean μ and variance σ^2. Suppose V is large compared to μ. Specify a single value that would be a good approximation for $N(V)$. What would be a good approximation for $EN(V)$? Specify how to approximate the distribution of $N(V)$ by a normal distribution. Assign specific numerical values for μ, σ^2, and V, and use the normal distribution to approximate the probability $P\{a \le N(V) \le b\}$ for a few values of a and b.

48. *Limiting Distribution of a Cyclic Renewal Process.* Consider a cyclic renewal process $X(t)$ on the states $0, 1, \dots, K-1$ as described in Example 1.9. Its inter-renewal distribution is $F = F_0 \star \cdots \star F_{K-1}$, where F_i is distribution of a sojourn time in state i having a finite mean μ_i. Assume one of the F_i is non-arithmetic. Show that F is non-arithmetic. Prove

$$\lim_{t\to\infty} P\{X(t)=i\} = \frac{\mu_i}{\mu_0 + \ldots + \mu_{K-1}}.$$

Is this limiting distribution the same as $\lim_{t\to\infty} t^{-1}E[\int_0^t \mathbf{1}(X(s)=i)ds]$, the average expected time spent in state i? State any additional assumptions needed for the existence of this limit.

49. Consider a $GI/G/1$ system as in Example 12.4. Let W_n' denote the length of time the nth customer waits in the queue prior to obtaining service. Determine a Little Law for the average wait $W' = \lim_{n\to\infty} n^{-1}\sum_{k=1}^n W_k'$.

50. *System Down Time.* Consider an alternating renewal process that represents the up and down states of a system. Suppose the up times have a distribution G with mean μ and variance σ^2, and the down times have a distribution G_0 with mean μ_0 and variance σ_0^2. Let $D(t)$ denote the length of time the system is down in the time interval $[0, t]$. Find the average expected down time $\beta \equiv \lim_{t\to\infty} t^{-1}ED(t)$. Then show $(D(t) - \beta t)/t^{1/2} \xrightarrow{d} N(0, \gamma^2)$ and specify γ.

51. *Congestion in a Road Race.* The following model was developed by Georgia Tech undergraduate students to assess the congestion in the 10-kilometer Atlanta Road Race, which is held every July 4th. After the pack of elite runners begins the race, the rest of the runners start the race a little later as follows. The runners are partitioned into m groups, with r_k runners assigned to group k, $1 \le k \le m$, depending on their anticipated completion times (the runners in each group being about equal in ability). The groups are released every τ minutes, with group k starting the race at time $k\tau$ (the groups are ordered so that the faster runners go earlier). Although the group sizes r_k are random, assume for simplicity that they are not. Typical numbers are 10 groups of 5000 runners in each group. The basic problem was to design a model for computing the probability that the congestion is above a certain level so that the runners have to slow down to a walk. (The students used this model to determine reasonable group sizes and their start times under which the runners would start as soon as possible, with a low probability of runners being forced to walk.)

The students assumed the velocity of each runner is the same throughout the race, the velocities of all the runners are independent, and the velocity of each runner in group k has the same distribution F_k. The distributions F_k were based on empirical distributions from samples obtained in prior races. Using pictures of past races, it was determined that if the number of runners in an interval of length ℓ in the road was greater than b, then the runners in that interval would be forced to walk.

Under these assumptions, the number of runners in group k that are in an interval $[a, a+\ell]$ on the road at time t is

$$Z_a^k(t) \equiv \sum_{n=1}^{r_k} \mathbf{1}(V_{kn}(t - k\tau) \in [a, a+\ell]),$$

where $V_{k1}, \dots, V_{kr_k}$ are the independent velocities of the runners in group k that have the distribution F_k. Then the total number of runners that are in $[a, a+\ell]$ at time t is

$$Z_a(t) = \sum_{k=1}^{m} Z_a^k(t).$$

Specify how one would use the central limit theorem to compute the probability $P\{Z_a(t) > b\}$ that the runners in $[a, a+\ell]$ at time t would be forced to walk.

52. Consider a delayed renewal process $N(t)$ with initial distribution $F_e(x) = \frac{1}{\mu}\int_0^x [1 - F(s)]ds$. Prove $EN(t) = U \star F_e(t) = t/\mu$ by a direct evaluation of the integral representing the convolution, where $U = \sum_{n=0}^{\infty} F^{n\star}$.

53. Justify expression (14.10), which in expanded form is

$$\sigma^2 = \mu_i^{-1}\mathrm{Var}(Z_{\nu_1} - a\nu_1) = \sum_{j\in S} \pi_j \tilde{f}(j)^2 + 2\sum_{j\in S}\pi_j \tilde{f}(j) \sum_{k\in S}\sum_{n=1}^{\infty} p_{jk}^n \tilde{f}(k).$$

First show that $E[Z_{\nu_1} - a\nu_1)] = 0$, and then and use the expansion

$$\begin{aligned} \mathrm{Var}(Z_{\nu_1} - a\nu_1) &= E_i\Big[[\sum_{n=1}^{\nu_1} \tilde{f}(X_n)]^2\Big] \\ &= E_i\Big[\sum_{n=1}^{\nu_1} \tilde{f}(X_n)^2\Big] + 2E_i\Big[\sum_{n=1}^{\nu_1} V_n\Big], \end{aligned} \tag{20.4}$$

where $V_n = \tilde{f}(X_n)\sum_{\ell=n+1}^{\nu_1} \tilde{f}(X_\ell)$. Apply Proposition 9.4 to the last two expressions in (20.4)(noting that $\sum_{n=1}^{\nu_1} V_n = \sum_{n=0}^{\nu_1 - 1} V_n$). Use the fact that

$$E_i[V_n|X_n = j, \nu_1 \geq n] = \tilde{f}(j)h(j),$$

where $h(j) = E_j[\sum_{n=1}^{\nu_1} \tilde{f}(X_n)]$ satisfies the equation

$$h(j) = \sum_{k\in S} p_{jk}\tilde{f}(k) + \sum_{k\in S} p_{jk}h(k),$$

and hence $h(j) = \sum_{k\in S}\sum_{n=1}^{\infty} p_{jk}^n \tilde{f}(k)$.

54. *CLT for Markov Processes.* Let $X(t)$ denote an ergodic Markov process with exponential sojourn rates q_i and limiting distribution $p_i = \pi_i/q_i/\sum_{j\in S}\pi_j/q_j$, for $i \in S$. Consider the functional $Z(t) = \int_0^t f(X(s))\,ds$, where $f(i)$ denotes a value per unit time when the process $X(t)$ is in state i. Example 11.8 showed that $t^{-1}Z(t) \to a \equiv \sum_{i\in I} p_i f(i)$, a.s., provided the sum exists. For simplicity, fix $i \in S$ and assume $X(0) = i$. Specify conditions, based on Theorem 14.3, under which

$$(Z(t) - at)/t^{1/2} \xrightarrow{d} N(0, \sigma^2), \qquad \text{as } t \to \infty.$$

Give an expression for σ^2 using ideas in Example 14.9 and

$$Z(T_1) - aT_1 = \sum_{k=1}^{\nu_1} [f(X_k)Y_k - aY_k],$$

where X_n is the embedded Markov chain, Y_n is the sojourn time of $X(t)$ in state X_n, and $\nu_1 = \min\{n \geq 1 : X_n = i\}$.

CHAPTER 3

Poisson Processes

Poisson processes are used extensively in applied probability models. Their importance is due to their versatility for representing a variety of physical processes, and due to the central limit phenomenon that a Poisson process is a natural model for a sum of many sparse point processes. The most basic Poisson process is a renewal process on the time axis with exponential inter-renewal times. This type of process is useful for representing times at which an event occurs, such as the times at which items arrive to a network, machine components fail, emergencies occur, a stock price takes a large jump, etc. The limiting behavior of these classical Poisson processes plus a few more properties were discussed in the preceding chapter. This discussion is continued in the first part of the present chapter which covers several characterizations of such processes. A distinguishing feature of a classical Poisson process is that its point locations (i.e., occurrence times) on a finite time interval are equal in distribution to order-statistics from a uniform distribution on the interval.

Applications of classical Poisson processes often involve auxiliary marks or random elements associated with the event occurrence times. For instance, if items arrive to a network at times that form a Poisson process, then a typical mark for an arriving item might be a vector denoting its route in the network and its service times at the nodes on the route. A convenient approach for analyzing such marks is to consider them as part of a larger "space-time" marked Poisson process on a multidimensional space. The properties of these processes are similar to those of "spatial" Poisson processes used for modelling locations of discrete items in the plane or a Euclidean space such as cell phone calls, truck delivery points, disease centers, geological formations, particles in space, fish colonies, etc.

This chapter describes the structure of contemporary Poisson processes on general spaces; space-time and spatial Poisson processes being special cases. The methodology for Poisson processes on general spaces involves technicalities about counting processes on general spaces, such as determining their distributions by Laplace functionals. A Poisson process on a general space is characterized in terms of a mixed binomial process. This is a generalization of the uniform order-statistic characterization of a homogeneous Poisson process on the time axis.

Much of the discussion covers summations, partitions, translations and general transformations of Poisson processes. Included are applications of space-time Poisson processes for analyzing particle systems and stochastic networks. One section shows that many properties of Poisson processes readily extend to several related processes; namely, Cox processes (i.e., Poisson processes with random intensities), compound Poisson processes, and cluster processes. The final results are central limit properties for rare events or points. They justify that a Poisson process is a natural limit for a collection of many sparse families of random points, i.e., a limit of a sum (or superposition) of many sparse point processes.

1. Poisson Processes on $\Re_+$

As in the last chapter, we define a *point process* $N = \{N(t) : t \geq 0\}$ on $\Re_+$ as a counting process $N(t) = \sum_{n=1}^{\infty} \mathbf{1}(T_n \leq t)$, where $0 = T_0 \leq T_1 \leq T_2 \leq \dots$ are random points (or times) such that $T_n \to \infty$ a.s. as $n \to \infty$. The point process N is *simple* when the points are distinct ($T_0 < T_1 < \dots$ a.s.).

We will also refer to the point process as the set of random variables $N = \{N(B) : B \in \mathcal{B}_+\}$, where

$$N(B) = \sum_{n=1}^{\infty} \mathbf{1}(T_n \in B), \qquad B \in \mathcal{B}_+,$$

denotes the number of points in the set B, and $\mathcal{B}_+$ denotes the Borel sets in $\Re_+$ (see Section 1 in the Appendix). Note that $N(B)$ is finite when B is bounded since $T_n \to \infty$ a.s. However, $N(B)$ may be infinite when the set B is not bounded, and $E[N(B)]$ may be infinite even though $N(B)$ is finite. In addition, we write

$$N(a, b] \equiv N((a, b]) = N(b) - N(a), \qquad a \leq b.$$

In the last chapter, a renewal process with exponential inter-renewal times was defined to be a Poisson process. This definition, as we show in the next section, is equivalent to the following one, which is more versatile.

Definition 9. A simple point process $N = \{N(t) : t \geq 0\}$ on $\Re_+$ is a *Poisson process* with rate $\lambda > 0$ if it has *independent increments* in that

$$N(s_1, t_1], \dots, N(s_n, t_n] \text{ are independent, for } s_1 < t_1 \cdots < s_n < t_n,$$

and $N(s, t]$ has a Poisson distribution with mean $\lambda(t - s)$, for any $s < t$.

Suppose N is a Poisson process with rate λ. It is sometimes called a *homogeneous* or *time-stationary* Poisson process, or a *classical* Poisson process. Theorem 2.1 proves, under the preceding definition, that N is also a renewal process whose inter-renewal times are independent exponentially distributed with rate λ. A number of elementary properties of N follow from this renewal characterization. For instance, we saw in Example 1.6 that the time T_n of the nth renewal has the distribution

$$P\{T_n \le t\} = P\{N(t) \ge n\} = 1 - \sum_{k=0}^{n-1} (\lambda t)^k e^{-\lambda t}/k!.$$

The derivative of this expression is $f(t) = \lambda^{n+1} t^n e^{-\lambda t}/n!$, and hence T_n has a gamma distribution with parameters n and λ.

Be mindful that all the properties of renewal processes apply to N. For instance, $t^{-1}N(t) \to \lambda$ a.s. by Corollary 2.5. Another observation is that the Poisson process N is also a Markov jump process.

Some applications of Poisson processes involve only elementary properties of the processes. Here is an example.

EXAMPLE 1.1. *Optimal Dispatching.* Consider a system in which discrete items arrive to a dispatching station according to a Poisson process N with rate λ during a fixed time interval $[0, T]$. The items might represent people to be bussed, ships to be unloaded, computer data or messages to be forwarded, material to be shipped, etc. There is a cost of h dollars per unit time of holding one item in the system. Also, at any time during the period, the items may be dispatched (or processed) at a cost of c dollars, and a dispatch is automatically done at time T. A dispatch is performed instantaneously and all the items in the system at that time are dispatched. Consider a dispatching policy defined by a vector $(n, t_1, \ldots, t_n)$, where n is the number of dispatches to make in the period, and $t_1 < t_2 < \cdots < t_n = T$ are the times of the dispatches. The aim is to find a dispatching policy that minimizes the expected cost.

We will show that the optimal solution is to have n^* dispatches at the times $t_i^* = iT/n^*$, where

$$n^* = \begin{cases} \lfloor x \rfloor & \text{if } \lfloor x \rfloor \lceil x \rceil \ge x^2 \\ \lceil x \rceil & \text{otherwise,} \end{cases} \tag{1.2}$$

and $x = T(h\lambda/2c)^{1/2}$.

This type of policy is a "static" policy in that it is implemented at the beginning of the time period and remains in effect during the period regardless of how the items actually arrive (e.g., there may be 0 items in a dispatch at a predetermined dispatch time t_i). A static policy might be appropriate when it is not feasible or too costly to monitor the system and do real-time dispatching. An alternative is to use a "dynamic" control policy that involves deciding when to make dispatches based on the observed queue of units. Exercise 4 asks if the policy above is optimal for non-Poisson processes.

To solve the problem, we will derive expressions for the total cost and its mean, and then find optimal values of the policy parameters. Under a fixed policy $(n, t_1, \dots, t_n)$, the total cost is

$$Z(n, t_1, \dots, t_n) = cn + h \sum_{i=1}^{n} W_i,$$

where W_i is the the amount of time items wait in the system during the time period $(t_{i-1}, t_i]$. Since $N(a, b]$ is the number of arrivals in a time interval $(a, b]$, it follows that

$$W_i = \int_{t_{i-1}}^{t_i} N(t_{i-1}, s]\, ds.$$

Using Fubini's theorem and $E[N(a, b]] = \lambda(b - a)$,

$$\begin{aligned} EW_i &= \int_{t_{i-1}}^{t_i} E[N(t_{i-1}, s]]\, ds \\ &= \lambda \int_{t_{i-1}}^{t_i} (s - t_{i-1})\, ds = \lambda(t_i - t_{i-1})^2/2. \end{aligned}$$

Then

$$E[Z(n, t_1, \dots, t_n)] = cn + \frac{h\lambda}{2} \sum_{i=1}^{n} (t_i - t_{i-1})^2.$$

Therefore, the aim is to solve the optimization problem

$$\min_{n} \min_{t_1, \dots, t_n} E[Z(n, t_1, \dots, t_n)], \tag{1.3}$$

subject to $t_{i-1} < t_i$ and $\sum_{i=1}^{n} (t_i - t_{i-1}) = T$.

It is well-known that the problem $\min_{x_1 \dots, x_n} \sum_{i=1}^{n} x_i^2$, under the constraint $\sum_{i=1}^{n} x_i = T$, has the solution $x_i^* = T/n$. This result follows by dynamic programming (backward induction), or by the use of Lagrange multipliers.

Applying this result to the problem (1.3), it follows that for fixed n, the subproblem $\min_{t_1, \dots, t_n} E[Z(n, t_1, \dots, t_n)]$ has the solution $t_i^* - t_{i-1}^* = T/n$, so that $t_i^* = iT/n$. Also, note that

$$f(n) \equiv E[Z(n, t_1^*, \dots, t_n^*)] = nc + h\lambda T^2/2n.$$

Then to solve (1.3), it remains to solve $\min_n f(n)$. Viewing n as a continuous variable x, the derivative $f'(x) = c - h\lambda T^2/(2x^2)$ is nondecreasing. Then $f(x)$ is convex and it is minimized at $x^* = T(h\lambda/2c)^{1/2}$. So the integer that minimizes $f(n)$ is either $\lfloor x^* \rfloor$ or $\lceil x^* \rceil$. Now $f(\lfloor x^* \rfloor) \le f(\lceil x^* \rceil)$ if and only if $\lfloor x^* \rfloor \lceil x^* \rceil \ge (x^*)^2$. This yields (1.2).

2. Characterizations of Classical Poisson Processes

The following theorem gives two characterizations of Poisson processes that provide more insight into their structure. Statement (c) has sometimes been used in texts as a definition of a Poisson process. It says that a Poisson process is such that the probability of a point in a small interval is directly proportional to the interval length and the probability of having more than one point in such an interval is essentially 0. These infinitesimal properties are the basis of the Poisson distribution for the quantities $N(a,b]$.

THEOREM 2.1. *For a simple point process $N = \{N(t) : t \geq 0\}$ on $\Re_+$ and $\lambda > 0$, the following statements are equivalent.*
(a) N is a Poisson process with rate λ.
(b) N is a renewal process whose inter-renewal times are exponentially distributed with rate λ.
(c) N has independent increments and, for any t and $h \downarrow 0$,

$$P\{N(t,t+h] = 1\} = \lambda h + o(h), \qquad P\{N(t,t+h] \geq 2\} = o(h).$$

PROOF. (a) $\Rightarrow$ (b). Assertion (b) states that $\xi_n = T_n - T_{n-1}$, $n \geq 1$, are independent and have an exponential distribution with rate λ. That is,

$$P\{A_n\} = e^{-\lambda \sum_{i=1}^n t_i}, \quad n \geq 1, \tag{2.2}$$

where $A_n \equiv \{\xi_1 > t_1, \ldots, \xi_n > t_n\}$ for $t_i > 0$.

Assuming N is a Poisson process with rate λ, we will prove (2.2) by induction. It is true for $n = 1$ since

$$P\{A_1\} = P\{\xi_1 > t_1\} = P\{N(t_1) = 0\} = e^{-\lambda t_1}.$$

Now assume (2.2) is true for some $n-1$. Then

$$P\{A_n\} = P\{A_{n-1}, \xi_n > t_n\} = P\{A_{n-1}\}P\{\xi_n > t_n | A_{n-1}\}. \tag{2.3}$$

Letting $F_n(t) = P\{T_n \leq t | A_{n-1}\}$, and using the Poisson properties of N,

$$\begin{aligned} P\{\xi_n > t_n | A_{n-1}\} &= \int_{\Re_+} P\{\xi_n > t_n | A_{n-1}, T_n = s\} dF_n(s) \\ &= \int_{\Re_+} P\{N(s, s+t_n] = 0 | T_n = s\} dF_n(s) \\ &= e^{-\lambda t_n}. \end{aligned}$$

Substituting this in (2.3), along with (2.2) for $n-1$, yields (2.2) for n.

(b) $\Rightarrow$ (c). Suppose (b) is true. We will first show that

$$N(s,t] \stackrel{d}{=} N(t-s), \qquad s < t. \tag{2.4}$$

Let $B(s) = T_{N(s)+1} - s$ denote the forward renewal recurrence time at time s. Because of the memoryless property of the exponential distribution (Exercise 2), $B(s)$ has an exponential distribution with rate λ (another justification is Example 8.15). Next, since N is a renewal process, $B(s)$ is independent of future renewals of N after time s. Also by the renewal property, $(B(s), N(s,t]) \stackrel{d}{=} (B(0), N(t-s))$. Therefore,

$$N(s,t] = N(s + B(s), t] \stackrel{d}{=} N(B(0), t-s] = N(t-s), \tag{2.5}$$

which proves (2.4).

Using (2.4),the exponential distribution of T_1, and $e^{-\lambda h} = 1-\lambda h + o(h)$, we have

$$\begin{aligned}P\{N(t,t+h] = 1\} &= P\{N(h) = 1\} = P\{T_1 \le h, T_2 > h\}\\ &= \int_0^h e^{-\lambda(h-s)}\lambda e^{-\lambda s} ds = \lambda h e^{-\lambda h} = \lambda h + o(h),\\ P\{N(t,t+h] \ge 2\} &= 1 - P\{N(h) = 0\} - P\{N(h) = 1\} = o(h).\end{aligned}$$

The proof of (c) will be complete upon showing that N has independent increments. We will prove by induction that, for any $s_1 < t_1 < \cdots < s_n < t_n$,

$$(2.6)\quad \zeta_n \equiv (N(s_i,t_i]; i = 1,\ldots,n) \stackrel{d}{=} (N_i(t_i - s_i); i = 1,\ldots,n), \qquad n \ge 1,$$

where $N_1,\ldots,N_n$ are independent renewal processes with $N_i \stackrel{d}{=} N$.

Statement (2.6) is true for $n = 1$ by (2.4). Next, suppose (2.6) is true for some $n-1$, and consider $\zeta_n = (\zeta_{n-1}, N(s_n,t_n])$. Arguing as above, the forward recurrence time $B(s_n)$ at time s_n is independent of ζ_{n-1}, and an analogue of (2.5) holds for $N(s_n,t_n]$. Therefore, (2.4) is true for n.

(c) $\Rightarrow$ (a). Assuming (c) is true, (a) will follow by proving $N(s,t]$ has a Poisson distribution with mean $\lambda(t-s)$, for each $s < t$. We begin with the case $s = 0$ and prove

$$p_n(t) \equiv P\{N(t) = n\} = (\lambda t)^n e^{-\lambda t}/n!, \qquad n \ge 0.$$

To establish this, we derive differential equations for the functions $p_n(t)$ and show their solutions are the preceding Poisson probabilities.

Under the assumptions in (c), for $n \ge 1$,

$$\begin{aligned}p_n(t+h) = P\{&N(t) = n, N(t,t+h] = 0\}\\ &+P\{N(t) = n-1, N(t,t+h] = 1\}\\ &+P\{N(t) = n, N(t,t+h] \ge 2\}.\end{aligned}$$

Since the last probability is $\le P\{N(t,t+h] \ge 2\} = o(h)$, and N has independent increments, the preceding is

$$p_n(t+h) = p_n(t)p_0(h) + p_{n-1}(t)p_1(h) + o(h).$$

In light of this expression, $p_n(t)$ is right-continuous on $\Re_+$. Substituting $p_1(h) = \lambda h + o(h)$, and $p_0(h) = 1 - p_1(h) + o(h)$ in the preceding expression yields

$$(p_n(t+h) - p_n(t))/h = -\lambda p_n(t) + \lambda p_{n-1}(t) + o(h)/h.$$

Similar reasoning shows that $p_n(t)$ is left-continuous on $(0,\infty)$ and

$$(p_n(t) - p_n(t-h))/h = -\lambda p_n(t-h) + \lambda p_{n-1}(t-h) + o(h)/h.$$

Then letting $h \downarrow 0$, yields the differential equations

$$p_n'(t) = -\lambda p_n(t) + \lambda p_{n-1}(t), \qquad n \ge 1.$$

By a similar argument,

$$p_0'(t) = -\lambda p_0(t).$$

To solve this family of differential-difference equations, with boundary conditions $p_n(0) = \mathbf{1}(n = 0)$, first note that $p_0(t) = e^{-\lambda t}$ satisfies the last differential equation. Then using this function and induction on $n \geq 1$ it follows that $p_n(t) = (\lambda t)^n e^{-\lambda t}/n!$. This proves that $N(t)$ has a Poisson distribution with mean λt. Furthermore, by the same argument, one can show that $N(s,t]$ has a Poisson distribution with mean $\lambda(t-s)$ for any $s < t$ (in this case, one uses $p_n(t) \equiv P\{N(s,t] = n\}$). $\square$

3. Location of Points

Many applications of Poisson processes involve knowledge about the locations of their points. Assertion (c) in Theorem 2.1 above suggests that the points of a Poisson process are located independently in a uniform sense on $\Re_+$. This section gives a precise description of this property based on a multinomial characterization of Poisson processes.

We first note that, for a Poisson process N with rate λ, the probability that it has a point at any fixed location t is 0 ($P\{N(\{t\}) = 1\} = 0$). This property follows from Proposition 3.2 for renewal processes or Proposition 7.6 below. Consequently, $N(a,b] \stackrel{d}{=} N(I)$, where I equals (a,b), $[a,b]$, or $[a,b)$, for $a < b$. In light of this observation and Definition 9, it follows that N is a Poisson process with rate λ if and only if $N(I_1), \ldots, N(I_n)$ are independent, for disjoint finite intervals $I_1, \ldots, I_n$, and $N(I)$ has a Poisson distribution with mean $\lambda|I|$ for any finite interval I. Here $|I|$ denotes the length of I.

The next result is a characterization of a Poisson process involving a multinomial distribution (3.2) of the numbers of its points in disjoint intervals. Interestingly, (3.2) is independent of the rate λ. The analogous property for Poisson processes on general spaces is given in Theorem 7.3 and Example 7.1.

THEOREM 3.1. *For a simple point process $N = \{N(t) : t \geq 0\}$ on $\Re_+$ and $\lambda > 0$, the following statements are equivalent.*
(a) N is a Poisson process with rate λ.
(b) (Multinomial Property) *For any $t > 0$, the quantity $N(t)$ has a Poisson distribution with mean λt, and, for any disjoint intervals $I_1, \ldots, I_k$ in $[0,t]$, and nonnegative integers $n_1, \ldots, n_k$,*

$$P\{N(I_1) = n_1, \ldots, N(I_k) = n_k | N(t) = n\} = \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k}, \tag{3.2}$$

where $n = n_1 + \cdots + n_k$ and $p_i = |I_i|/t$.

PROOF. (a) $\Rightarrow$ (b). Suppose (a) holds. By the discussion above, N has independent Poisson increments over any disjoint intervals (they need not be of the form $(a,b]$). Then letting $I_0 = [0,t] \setminus \cup_{i=1}^k I_i$ and $n_0 = 0$, the conditional probability in (3.2) is

$$\frac{P\{N(I_i) = n_i,\ 0 \leq i \leq k\}}{P\{N(t) = n\}} = \frac{\prod_{i=0}^k (\lambda|I_i|)^{n_i} e^{-\lambda|I_i|}/n_i!}{e^{-\lambda t}(\lambda t)^n/n!}.$$

This clearly reduces to the right-hand side of (3.2) since $\sum_{i=0}^k |I_i| = t$.

(b) $\Rightarrow$ (a). Suppose (b) holds. Fix a t and choose any $0 = t_0 < t_1 < \cdots < t_k = t$ and nonnegative integers $n_1, \ldots, n_k$ such that $n = n_1 + \cdots + n_k$. Define $I_i = (t_{i-1}, t_i]$ and $A_i = \{N(I_i) = n_i\}$. Then under the properties in (b),

$$\begin{aligned} P\{\cap_{i=1}^k A_i\} &= P\{N(t) = n\} P\{\cap_{i=1}^k A_i | N(t) = n\} \\ &= \prod_{i=1}^k \frac{[\lambda(t_i - t_{i-1})]^{n_i}}{n_i!} e^{-\lambda(t_i - t_{i-1})} = \prod_{i=1}^k P\{A_i\}. \end{aligned}$$

This proves $N(I_1), \ldots, N(I_k)$ are independent, and $N(I_i)$ has a Poisson distribution with mean $\lambda(t_i - t_{i-1})$.

The proof of (a) will be complete upon showing that the increments $N(s_1, t_1], \ldots, N(s_k, t_k]$ are independent, for any $s_1 < t_1 < \cdots < s_k < t_k = t$. However, this independence follows because these increments are a subset of the increments $N(0, s_1], N(s_1, t_1], N(t_1, s_2], \ldots, N(s_k, t_k]$ over all the adjacent intervals, which are independent as we just proved. $\square$

A special case of (3.2) is the *binomial property*: For $I \subset [0, t]$ and $k \le n$,

$$P\{N(I) = k | N(t) = n\} = \binom{n}{k} (|I|/t)^k (1 - |I|/t)^{n-k}.$$

The multinomial property also yields the joint conditional distribution of point locations in $[0, t]$ given $N(t) = n$, as shown in (3.4) below.

THEOREM 3.3. (Order Statistic Property) *Suppose N is a Poisson process with rate λ. Then, for any disjoint intervals $I_1, \ldots, I_n$ in $[0, t]$*

$$P\Big\{T_1 \in I_1, \ldots, T_n \in I_n \Big| N(t) = n\Big\} = \frac{n!}{t^n} \prod_{i=1}^n |I_i|. \tag{3.4}$$

Hence, the joint conditional density of $T_1, \ldots, T_n$ given $N(t) = n$ is

$$f_{T_1,\ldots,T_n}\Big(t_1, t_2, \ldots, t_n \Big| N(t) = n\Big) = \frac{n!}{t^n}, \qquad 0 < t_1 < \cdots < t_n < t. \tag{3.5}$$

The density (3.5) is the density of the order statistics of n independent uniformly distributed random variables on $[0, t]$ (see Proposition 3.7 below).

PROOF. Expression (3.4) follows since the conditional probability in it equals $P\{N(I_i) = 1, 1 \le i \le n | N(t) = n\}$, which in turn equals the right-hand side of (3.4) by the multinomial property (3.2).

Next, note that (3.4), for $a_1 < b_1 < \cdots < a_n < b_n < t$, is

$$P\Big\{T_i \in (a_i, b_i], 1 \le i \le n \Big| N(t) = n\Big\} = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} \frac{n!}{t^n} dt_1 \cdots dt_n.$$

Then the integrand $n!/t^n$ is the conditional density as asserted in (3.5). $\square$

EXAMPLE 3.6. *Marginal Distributions.* From (3.4), (3.5) and Exercise 18,

$$P\{T_k \le s|N(t) = n\} = \sum_{j=k}^{k} \binom{n}{j} (s/t)^j (1 - s/t)^{n-j},$$

$$f_{T_k}(s|N(t) = n) = \frac{n!}{(k-1)!(n-k)!}(s/t)^{k-1}(1/t)(1-s/t)^{n-k}, \quad 0 \le s \le t.$$

Also, $E[T_k|N(t) = n] = kt/(n+1)$. In particular, for a single point,

$$P\{T_1 \le s|N(t) = 1\} = s/t, \qquad 0 \le s \le t.$$

This is a uniform distribution on $[0, t]$.

We referred to (3.5) as the density of n order statistics of independent uniformly distributed random variables on $[0, t]$. This is justified by the following formula for the density of order statistics of a random sample with a general density.

PROPOSITION 3.7. (Order Statistics) *Suppose $X_1, \dots, X_n$ are independent continuous random variables with density f, and let $X_{(1)} < \cdots < X_{(n)}$ denote the quantities $X_1, \dots, X_n$ in increasing order. These* order statistics *$X_{(1)}, \dots, X_{(n)}$ have the joint density*

$$f(x_1, \dots, x_n) = n! f(x_1) \cdots f(x_n), \qquad x_1 < \cdots < x_n. \tag{3.8}$$

PROOF. Choose any $a_1 < b_1 < \cdots < a_n < b_n$, and let $I_i = (a_i, b_i]$, $1 \le i \le n$. Since $X_{(1)}, \dots, X_{(n)}$ is equally likely to be any one of the $n!$ permutations of $X_1, \dots, X_n$,

$$\begin{aligned} P\{X_{(i)} \in I_i,\ 1 \le i \le n\} &= n! P\{X_i \in I_i,\ 1 \le i \le n\} \\ &= n! \prod_{i=1}^{n} P\{X_i \in I_i\} \\ &= n! \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1) \cdots f(x_n)\, dx_1 \cdots dx_n. \end{aligned}$$

This proves the density formula (3.8). □

4. Functions of Points

Typical quantities of interest for a Poisson process N in a time interval $[0, t]$ are deterministic or random functions of the point locations $T_1, \dots, T_{N(t)}$. A classic example is $\sum_{n=1}^{N(t)} f(T_n)$, where $f : \Re_+ \to \Re$. This section shows how to analyze such functions in terms of random samples.

The following result is an immediate consequence of Theorem 3.3.

COROLLARY 4.1. (Order Statistic Tool) *Let N be a Poisson process with rate λ, and, for each $n \ge 1$, let h_n be a function from $\Re_+^n$ to some Euclidean* *or more general space S, and let $h_0 \in S$. Then, for $t > 0$,*

$$h_{N(t)}(T_1, \dots, T_{N(t)}) \stackrel{d}{=} h_\kappa(X_{(1)}, \dots, X_{(\kappa)}),$$

where $X_{(1)} < \cdots < X_{(n)}$ are the n order statistics associated with independent random variables $X_1, \dots, X_n$ that are uniformly distributed on $[0, t]$ for each n, and κ is a Poisson random variable with mean λt, independent of the X_i's. Furthermore, if each $h_n(x_1, \dots, x_n)$ is symmetric (meaning it is the same for any permutation of $x_1, \dots, x_n$), then

$$h_{N(t)}(T_1, \dots, T_{N(t)}) \stackrel{d}{=} h_\kappa(X_1, \dots, X_\kappa). \tag{4.2}$$

These expressions enable one to analyze a function of the random-length, dependent variables $T_1, \dots, T_{N(t)}$ by the simpler mixed random sample $X_1, \dots, X_\kappa$. The ideas here are related to the characterization of a Poisson process by mixed binomial or sample processes in Theorem 7.3 below.

As an example, for $f : \Re_+ \to R$,

$$\sum_{n=1}^{N(t)} f(T_n) \stackrel{d}{=} \sum_{n=1}^{\kappa} f(X_n).$$

In this case, $h_n(x_1, \dots, x_n) \equiv \sum_{i=1}^n f(x_i)$ is symmetric. Here is another example involving random functions.

Proposition 4.3. (Random Sums) *Suppose N is a Poisson process with rate λ, and define*

$$Z(t) = \sum_{n=1}^{N(t)} f(T_n, Y_n), \tag{4.4}$$

where $Y_1, Y_2 \dots$ are i.i.d. random elements in a space S that are independent of N, and $f : \Re_+ \times S \to \Re$. Assume $\phi(\alpha, t) \equiv E[e^{\alpha f(t, Y_1)}]$ exists for α in a neighborhood of 0. Then the moment generating function of $Z(t)$ is

$$E[e^{\alpha Z(t)}] = e^{-\lambda t(1 - g_t(\alpha))}, \tag{4.5}$$

where $g_t(\alpha) \equiv t^{-1} \int_0^t \phi(\alpha, s)\, ds$. Hence,

$$E[Z(t)] = \lambda \int_0^t E[f(s, Y_1)]\, ds. \tag{4.6}$$

Proof. First note that

$$\begin{aligned} E[e^{\alpha Z(t)}] &= E\Big[E[e^{\alpha Z(t)} | N(s), s \le t]\Big] \\ &= E\Big[\prod_{n=1}^{N(t)} E[e^{\alpha f(T_n, Y_n)} | N(s), s \le t]\Big] = E\Big[\prod_{n=1}^{N(t)} \phi(\alpha, T_n)\Big]. \end{aligned}$$

Applying (4.2) with $h_n(x_1, \dots, x_n) \equiv \prod_{i=1}^n \phi(\alpha, x_i)$ to the last expression,

$$E[e^{\alpha Z(t)}] = E\Big[\prod_{n=1}^{\kappa} \phi(\alpha, X_n)\Big].$$

Then conditioning on κ, which is independent of the X_n's, we have

$$E[e^{\alpha Z(t)}] = E\Big[\prod_{n=1}^{\kappa} E[\phi(\alpha, X_n)]\Big].$$

Clearly $E[\phi(\alpha, X_n)] = g_t(\alpha)$, since X_n is uniformly distributed on $[0, t]$. Using this along with the independence of the X_n's and the well-known Poisson generating function $E[\beta^\kappa] = e^{-\lambda t(1-\beta)}$, we obtain

$$E[e^{\alpha Z(t)}] = E[g_t(\alpha)^\kappa] = e^{-\lambda t(1-g_t(\alpha))}.$$

This proves (4.5). In addition, (4.6) follows, since the derivative of (4.5) with respect to α at $\alpha = 0$ is $E[Z(t)] = t\lambda g_t'(0)e^{-\lambda t(1-g_t(0))}$, where $g_t(0) = 1$ and $g_t'(0) = \phi'(0) = t^{-1}\int_0^t E[f(x, Y_1)]\,dx$. □

EXAMPLE 4.7. *Discounted Cash Flows.* A special case of the random sum (4.4) is

$$Z(t) = \sum_{n=1}^{N(t)} Y_n e^{-\gamma T_n},$$

where $\gamma > 0$. This is a standard model for discounted costs or revenues, where γ is a deterministic discount rate. For instance, suppose that sales of a product occur at times that form a Poisson process N with rate λ. The amount of revenue from the nth sale is a random variable Y_n. Then the total discounted revenue up to time t is given by $Z(t)$. By (4.5) and (4.6), the moment generating function and mean of $Z(t)$ are

$$\begin{aligned} E[e^{\alpha Z(t)}] &= \exp\{-\lambda t(1 - t^{-1}\int_0^t E[e^{\alpha e^{-\gamma x} Y_1}]\,dx)\} \\ E[Z(t)] &= \lambda E Y_1(1 - e^{-\gamma t})/\gamma. \end{aligned}$$

EXAMPLE 4.8. *Compound Poisson Process.* Another example of (4.4) is $Z(t) = \sum_{n=1}^{N(t)} Y_n$. This is like the preceding discounted cash flow, but without discounting. Letting $\gamma = 0$ in the preceding example, it follows that

$$\begin{aligned} E[e^{\alpha Z(t)}] &= e^{-\lambda t(1-E[e^{\alpha Y_1}])} \\ E[Z(t)] &= \lambda t E Y_1. \end{aligned}$$

This generating function of $Z(t)$ is that of a compound Poisson distribution with rate λt and distribution $F(y) = P\{Y_1 \le y\}$, which is

$$P\{Z(t) \le z\} = \sum_{n=0}^{\infty} e^{-\lambda t}(\lambda t)^n F^{n\star}(z)/n!, \qquad z \in \Re.$$

The process $\{Z(t) : t \ge t\}$ is called a *compound Poisson process*; further properties of it are in Example 13.6.

5. General Poisson Processes

This section introduces the terminology for point processes and Poisson processes on general spaces that is needed for the rest of the chapter.

A point process is a counting process that represents a random set of points in a space. Typical spaces are the real line, the plane, the multi-dimensional Euclidean space $\Re^d$, or, more generally, a complete, separable metric space (a Polish space). Following the standard convention, we will discuss point processes on a polish space S. The exposition will be understandable by thinking of S as an Euclidean space. We let $\mathcal{S}$ denote the family of Borel sets of S (see section 3), and let $\hat{\mathcal{S}}$ denote the family of *bounded* or *locally compact* Borel sets (a set is bounded if it is contained in a compact set). We refer to S simply as a *space*, and denote other spaces of this type by S', $\tilde{S}$, etc.

We will represent a set of points in S by a counting measure. Specifically, suppose that $x_1, \ldots, x_k$ are (deterministic) locations of points (or unit masses) in S, where $k \leq \infty$. There may be more than one point at a location, and the order of the subscripts on the locations is invariant under permutations. These points are represented by the *counting measure* ν on S defined by

$$\nu(B) = \sum_{n=1}^{k} \mathbf{1}(x_n \in B), \quad B \in \mathcal{S},$$

which denotes the number of point in B. Here $\nu(S) = k$ and $\sum_{n=1}^{k}(\cdot) = 0$ when $k = 0$. For simplicity, we write such sums as

$$\nu(B) = \sum_{n} \delta_{x_n}(B), \quad B \in \mathcal{S}, \tag{5.1}$$

where $\delta_x(B) = \mathbf{1}(x \in B)$ is a Dirac measure with unit mass at x.

We will only consider counting measures ν on S that are *locally finite*, meaning that they are finite on bounded sets ($\nu(B) < \infty$, $B \in \hat{\mathcal{S}}$). Let $\mathbb{M}$ denote the set of all such locally finite counting measures on $(S, \mathcal{S})$. Endow $\mathbb{M}$ with the σ-field $\mathcal{M}$ on $\mathbb{M}$ generated by the sets $\{\nu \in \mathbb{M} : \nu(B) = n\}$, for $B \in \mathcal{S}$ and $n \in \mathbb{N}_+$.

DEFINITION 10. A *point process* N on a space S is a measurable map from a probability space $(\Omega, \mathcal{F}, P)$ to the space $(\mathbb{M}, \mathcal{M})$. The quantity $N(B)$ is the number of points in the set $B \in \mathcal{S}$. From the representation (5.1),

$$N(B) = \sum_{n} \delta_{X_n}(B), \quad B \in \mathcal{S}, \tag{5.2}$$

where the X_n denote the *locations* of the points of N.

For the following discussion, assume that N is a point process on the space S. Technical properties of the space S are not used explicitly in the sequel. One can simply think of N as a counting process on $S = \Re^d$ that is locally finite ($N(B) < \infty$ a.s. for $B \in \hat{\mathcal{S}}$, the bounded Borel sets in S).

The *probability distribution* of the point process N (i.e., $P\{N \in \cdot\}$) is determined by its finite-dimensional distributions

$$P\{N(B_1) = n_1, \ldots, N(B_k) = n_k\}, \qquad B_1, \ldots, B_k \in \hat{\mathcal{S}}. \tag{5.3}$$

In other words, two point processes N and N' on S are equal in distribution, denoted by $N \stackrel{d}{=} N'$, if their finite-dimensional distributions are equal:

$$(N(B_1) \ldots, N(B_k)) \stackrel{d}{=} (N'(B_1) \ldots, N'(B_k)), \qquad B_1, \ldots, B_k \in \hat{\mathcal{S}}.$$

In constructing a point process, it suffices to define the probabilities (5.3) on sets B_i that generate $\mathcal{S}$. For instance, when $S = \Re^d$, "rectangles" of the form $(a, b]$ generate $\mathcal{S}$.

The *intensity measure* (or mean measure) of the point process N is

$$\mu(B) \equiv E[N(B)], \qquad B \in \mathcal{S}.$$

Note that $\mu(B)$ may be infinite, even if B is bounded. When $S = \Re^d$, the intensity is sometimes of the form $\mu(B) = \int_B \lambda(x)dx$, where $\lambda(x)$ is the rate of N at the location x and dx denotes the Lebesgue measure. We call $\lambda(x)$ the location-dependent *rate function* of N.

We are now ready to define Poisson processes.

DEFINITION 11. A point process N on a space S is a *Poisson process with intensity measure* μ that is locally finite if the following conditions are satisfied.
• N has *independent increments*: The quantities $N(B_1), \ldots, N(B_n)$ are independent for disjoint sets $B_1, \ldots, B_n$ in $\hat{\mathcal{S}}$.
• For each $B \in \hat{\mathcal{S}}$, the quantity $N(B)$ is a Poisson random variable with mean $\mu(B)$.

This definition uses the convention that $N(B) = 0$ a.s. when $\mu(B) = 0$. Note that if $\mu(\{x\}) > 0$, then the number of points $N(\{x\})$ exactly at x has a Poisson distribution with mean $\mu(\{x\})$. On the other hand, if $\mu(\{x\}) = 0$, then $N(\{x\}) = 0$ a.s. From the definition it follows that the finite-dimensional distributions of a Poisson process are uniquely determined by its intensity measure, and vice versa. That is, if N and N' are Poisson processes on S with respective intensities μ and μ', then $N \stackrel{d}{=} N'$ if and only if $\mu = \mu'$.

Do Poisson processes exit? In other words, does there exist a point process on a probability space that satisfies the properties in Definition 11? We will establish the existence after we show in Theorem 7.5 below that a Poisson process can be characterized by independent random elements, which do exist. We end this section with a few comments on the earlier definition of Poisson processes on $\Re_+$.

EXAMPLE 5.4. *Poisson Processes on* $\Re_+$. A Poisson process N on $\Re_+$ (or on $\Re$) with intensity measure μ is sometimes called a *nonhomogeneous* Poisson process. We denote its point locations (as we have been doing) by

$0 < T_1 \le T_2 \le \ldots$ instead of X_n, and call them "times" when appropriate. In particular, $N(B) = \sum_n \delta_{T_n}(B)$ has a Poisson distribution with mean $\mu(B)$. We also write $N(t) = N(0,t]$, for $t > 0$, and $N(a,b] = N((a,b])$ and $\mu(a,b] = \mu((a,b])$, for $a < b$. When $\mu(B) = \int_B \lambda(t)dt$, we say N is Poisson with *rate function* $\lambda(t)$. In case $\mu(t) = E[N(t)] = \lambda t$, for some $\lambda > 0$, then N a Poisson process with rate λ (consistent with Definition 9); it is sometimes called a *homogeneous* Poisson process with rate λ. The results in the preceding sections for homogeneous Poisson processes have obvious analogues for nonhomogeneous processes. For instance, N satisfies the multinomial property (3.2) with $p_i = \mu(B_i)/\mu(0,t]$, and Exercise 19 describes its order statistic property.

6. Integrals and Laplace Functionals of Poisson Processes

Laplace transforms are useful for identifying distributions of nonnegative random variables and for establishing convergence in distribution of random variables. The analogous tool for point processes is a Laplace functional. This section covers a few properties of Laplace functionals and related integrals of point processes. These are preliminaries needed to establish the existence of Poisson processes, the topic of the next section. The use of Laplace transforms and functionals for establishing the convergence of random variables and point processes are covered later in Sections 14 and 15.

We begin with a little review. Recall that the Laplace transform of a nonnegative random variable X with distribution F is

$$\hat{F}_X(\alpha) \equiv E[e^{-\alpha X}] = \int_{\Re_+} e^{-\alpha x} dF(x), \qquad \alpha \ge 0.$$

This function uniquely determines the distribution of X in that $X \stackrel{d}{=} Y$ if and only if $\hat{F}_X(\cdot) = \hat{F}_Y(\cdot)$. For instance, the Laplace transform of a Poisson random variable X with mean λ is

$$\hat{F}_X(\alpha) = \sum_{n=0}^{\infty} e^{-\alpha n} e^{-\lambda} \lambda^n / n! = e^{-\lambda(1-e^{-\alpha})}.$$

Now, if Y is a nonnegative integer-valued random variable with $E[e^{-\alpha Y}] = e^{-\lambda(1-e^{-\alpha})}$, then Y has a Poisson distribution with mean λ. Here is an example for sums.

EXAMPLE 6.1. *Sums of Independent Poisson Random Variables.* Suppose $Y_1, \ldots, Y_n$ are independent Poisson random variables with respective means $\mu_1, \ldots, \mu_n$. Then $\sum_{i=1}^n Y_i$ has a Poisson distribution with mean $\mu = \sum_{i=1}^n \mu_i$ (which we assume is finite when $n = \infty$). To see this result, note that by the independence of the Y_i and the form of their Laplace transforms,

$$E[e^{-\alpha \sum_{i=1}^n Y_i}] = \prod_{i=1}^n E[e^{-\alpha Y_i}] = e^{-\mu(1-e^{-\alpha})}.$$

We recognize this as being the Laplace transform of a Poisson distribution with mean μ, and so $\sum_{i=1}^{n} Y_i$ has this distribution.

The rest of this section covers analogous properties for Laplace functionals of point processes. Consider a point process $N = \sum_n \delta_{X_n}$ on a space S. The "integral" of a function $f : S \to \Re_+$ with respect to N is the sum

$$Nf \equiv \int_S f(x)N(dx) \equiv \sum_n f(X_n),$$

provided it is finite. It is finite when f has a compact support (i.e., $\{x : f(x) > 0\}$ is contained in a compact set). Similarly, the integral of $f : S \to \Re_+$ with respect to a measure μ will be denoted by

$$\mu f \equiv \int_S f(x)\mu(dx).$$

We will often use integrals of functions f in the set $C_K^+(S)$ of all continuous functions $f : S \to \Re_+$ with compact support.

Definition 12. The *Laplace functional* of the point process N is

$$E[e^{-Nf}] = E\Big[\exp\{-\int_{\Re_+} f(x)N(dx)\}\Big], \qquad f : S \to \Re_+.$$

The function f is a "variable" of this expectation (just as the parameter α is a variable in a Laplace transform $E[e^{-\alpha X}]$).

The following result contains the basic property that the Laplace functional of a point process uniquely determines its distribution (the proof is in [**21**]). It also justifies that a Laplace functional is uniquely defined on the set $C_K^+(S)$ viewed as "test" functions.

Theorem 6.2. *For point processes N and N' on S, each one of the following statements is equivalent to $N \stackrel{d}{=} N'$.*
(a) $Nf \stackrel{d}{=} N'f, \qquad f \in C_K^+(S)$.
(b) $E[e^{-Nf}] = E[e^{-N'f}], \quad f \in C_K^+(S)$.

Laplace functionals are often more convenient to use than finite-dimensional distributions in deriving the distribution of a point process constructed as a function of random variables or point processes. A standard approach for establishing that a point process is Poisson is to verify that its Laplace functional has the following form; this also yields its intensity measure.

Proposition 6.3. (Poisson Laplace Functional) *For a Poisson process N on S with intensity measure μ, and $f : S \to \Re_+$,*

$$E[e^{-Nf}] = \exp[-\int_S (1 - e^{-f(x)})\mu(dx)].$$

PROOF. First consider the simple function $f(x) = \sum_{i=1}^k a_i \mathbf{1}(x \in B_i)$, for some nonnegative $a_1, \ldots, a_k$ and disjoint $B_1, \ldots, B_k$ in $\hat{\mathcal{S}}$. Then

$$Nf = \sum_{i=1}^k \int_S a_i \mathbf{1}(x \in B_i) N(dx) = \sum_{i=1}^k a_i N(B_i).$$

Using this and the independence of the $N(B_i)$, we have

$$\begin{aligned} E[e^{-Nf}] &= \prod_{i=1}^k E[e^{-a_i N(B_i)}] \;=\; \exp[-\sum_{i=1}^k \mu(B_i)(1 - e^{-a_i})] \\ &= \exp[-\int_S (1 - e^{-f(x)})\mu(dx)]. \end{aligned}$$

Next, for any $f : S \to \Re_+$, there exist simple functions $f_n \uparrow f$ (e.g., $f_n(x) = n \wedge (\lfloor 2^n f(x) \rfloor / 2^n)$). Then by the monotone convergence theorem (see the Appendix, Theorem 8.6) and the first part of this proof,

$$\begin{aligned} E[e^{-Nf}] &= \lim_{n\to\infty} E[e^{-Nf_n}] = \lim_{n\to\infty} \exp[-\int_S (1 - e^{-f_n(x)})\mu(dx)] \\ &= \exp[-\int_S (1 - e^{-f(x)})\mu(dx)]. \end{aligned}$$

This completes the proof. □

Recall that Example 6.1 uses Laplace transforms to prove that a sum of independent Poisson random variables is Poisson. Here is an analogous result for a sum of Poisson processes.

THEOREM 6.4. (Sums of Independent Poisson Processes) *Suppose that $N_1, \ldots, N_n$ are independent Poisson processes on S with respective intensity measures $\mu_1, \ldots, \mu_n$. Then their* sum (or superposition) $N = \sum_{i=1}^n N_i$ *is a Poisson process with intensity measure $\mu = \sum_{i=1}^n \mu_i$. This is also true for $n = \infty$ provided μ is locally finite.*

PROOF. One can prove this, as suggested in Exercise 6, by verifying that N satisfies the defining properties of a Poisson process. Another approach, using Laplace functionals and Proposition 6.3, is to show that

$$E[e^{-Nf}] = e^{-\mu h}, \qquad f \in C_K^+(S),$$

where $h(x) = 1 - e^{-f(x)}$. But this follows since by the independence of the N_i and the form of their Laplace functionals in Proposition 6.3,

$$E[e^{-Nf}] = \prod_{i=1}^n E[e^{-N_i f}] = \prod_{i=1}^n e^{-\mu_i h} = e^{-\mu h}.$$

□

EXAMPLE 6.5. A company that produces a household cleaning fluid has a bottle-filling production line that occasionally has to stop for repair due to imperfections in the bottles or due to worker errors. There are four types of line stoppages: (1) minor stop (under 30 minutes) due to bottle imperfection, (2) major stop (over 30 minutes) due to bottle imperfection, (3) minor stop due to worker error, and (4) major stop due to worker error. These four types of stoppages occur according to independent Poisson processes with respective rates $\lambda_1, \dots, \lambda_4$. Then by Theorem 6.4, line stops due to any of these causes occur according to a Poisson process with rate $\lambda_1 + \cdots + \lambda_4$. Similarly, minor stops occur according to a Poisson process with rate $\lambda_1 + \lambda_3$, and major stops occur according to a Poisson process with rate $\lambda_2 + \lambda_4$.

We end this section with more insight on integrals $Nf = \sum_n f(X_n)$ with respect to a point process N. Expressions (6.7) below says the mean of such an integral equals the corresponding integral with respect to the intensity measure. Theorem 3.5 for renewal processes is a special case. The variance of Nf has the nice form (6.8), when N is Poisson.

THEOREM 6.6. *Let $N = \sum_n \delta_{X_n}$ be a point process on S with intensity measure μ. For any $f : S \to \Re$,*

$$E\Big[\sum_n f(X_n)\Big] = \int_S f(x)\mu(dx), \tag{6.7}$$

provided the integral exists. That is, $E[Nf] = \mu f$. If in addition, N is a Poisson process, then

$$Var\Big(\sum_n f(X_n)\Big) = \int_S f(x)^2\mu(dx), \tag{6.8}$$

provided the integral exists. That is, $Var(Nf) = \mu f^2$.

PROOF. The proof of (6.7) is similar to that of Theorem 3.5 or Proposition 6.3. Namely, first one shows $E[Nf] = \mu f$ is true when f is a simple function, and then monotone convergence yields the equality for general f, which is a monotone limit of simple functions.

To prove (6.8), note that by Proposition 6.3, we have

$$E[e^{-\alpha Nf}] = e^{-h(\alpha)}, \tag{6.9}$$

where $h(\alpha) = \int_S (1 - e^{-\alpha f(x)})\mu(dx)$. The derivative of this expression at $\alpha = 0$, yields

$$E[Nf] = h'(0)e^{-h(0)} = \mu f.$$

Furthermore, taking the second derivative of (6.9) at $\alpha = 0$, and using $h(0) = 1$ and $h'(0) = E[Nf]$, we obtain

$$\begin{aligned} E[(Nf)^2] &= \lim_{\alpha \downarrow 0} \Big[(h'(\alpha))^2 e^{-h(\alpha)} - h''(\alpha)e^{-h(\alpha)}\Big] \\ &= \lim_{\alpha \downarrow 0} \Big[\int_S f(x)^2 e^{-\alpha f(x)}\mu(dx) + (h'(\alpha))^2\Big] \\ &= \int_S f(x)^2\mu(dx) + (E[Nf])^2. \end{aligned}$$

This proves (6.8). $\square$

7. Poisson Processes as Mixed Binomial Processes

This section contains a characterization of Poisson processes in terms of mixed binomial (or sample) processes. The characterization is used to establish the existence of Poisson processes. It is also used in applications dealing with point locations of a process like those above in Section 4 involving the order statistic property of points on $\Re_+$.

Our goal is to give a constructive description of Poisson in terms of elementary binomial processes defined as follows.

DEFINITION 13. Suppose $X_1, X_2, \dots$ are i.i.d. random elements in the space S with distribution $\mu(B) \equiv P\{X_1 \in B\}$. For a fixed n, the point process $N \equiv \sum_{i=1}^{n} \delta_{X_i}$ is a *binomial* (or sample) process on S based on μ. Each $N(B)$ has a binomial distribution with parameters n and $\mu(B)$. If, in addition, κ is a nonnegative integer-valued random variable independent of the X_i, then the point process $M \equiv \sum_{i=1}^{\kappa} \delta_{X_i}$ is a *mixed binomial* (or mixed sample) process based on μ and κ.

EXAMPLE 7.1. Fires occur in a region S of a city at locations $X_1, X_2, \dots$ that are independent with distribution F. Then the spatial locations of n fires in S is given by the binomial process $N = \sum_{i=1}^{n} \delta_{X_i}$. In particular, for $B \in \mathcal{S}$ and $m \leq n$,

$$P\{N(B) = m\} = \binom{n}{m} F(B)^m (1 - F(B))^{n-m}.$$

Also, for $B_1, \dots, B_k$ in $\mathcal{S}$ that form a partition of S, and $n = n_1 + \cdots + n_k$,

$$P\{N(B_1) = n_1, \dots, N(B_k) = n_k\} = \frac{n!}{n_1! \cdots n_k!} F(B_1)^{n_1} \cdots F(B_k)^{n_k}.$$

Suppose the number of fires in a year is a random variable κ independent of the locations. Then the spatial locations of these κ fires in S is given by the mixed binomial process $M = \sum_{i=1}^{\kappa} \delta_{X_i}$. If κ has a Poisson distribution with mean λ, then as shown next in Lemma 7.2, M is Poisson with intensity $\lambda F(\cdot)$.

Our first observation is that a Poisson process with a finite intensity measure is a special kind of mixed binomial process.

LEMMA 7.2. *Suppose N is a Poisson process on S with intensity measure μ such that $0 < \mu(S) < \infty$. Then N is equal in distribution to a mixed binomial process M on S based on F and κ, where $F(\cdot) \equiv \mu(\cdot)/\mu(S)$ and κ has a Poisson distribution with mean $\mu(S)$.*

PROOF. Consider the representation $M = \sum_{i=1}^{\kappa} \delta_{X_i}$ where X_i are independent with distribution F and are independent of κ. Then using the generating function $E[z^\kappa] = e^{-\mu(S)(1-z)}$, it follows that, for any $f \in C_K^+(S)$,

$$\begin{aligned} E[e^{-Mf}] &= E\Big[E[e^{-\sum_{i=1}^{\kappa} f(X_i)}|\kappa]\Big] \\ &= E\Big[E[e^{-f(X_1)}]^\kappa\Big] \\ &= \exp\{-\mu(S)(1 - E[e^{-f(X_1)}])\} \\ &= \exp\{-\int_S (1 - e^{-f(x)})\mu(dx)\}. \end{aligned}$$

We recognize this from Proposition 6.3 as the Laplace functional of the Poisson process N with intensity μ, and hence $M \stackrel{d}{=} N$. □

The preceding result extends to the following characterization, which says that a Poisson process is "locally" a mixed binomial process.

THEOREM 7.3. *Let N be a point process on S, and let μ be a locally finite measure on S. The following statements are equivalent.*
(i) *N is a Poisson process with intensity measure μ.*
(ii) *N on each $B \in \hat{\mathcal{S}}$ with $\mu(B) > 0$ is equal in distribution to a mixed binomial process M on B based on F and κ, where $F(\cdot) = \mu(\cdot \cap B)/\mu(B)$ and κ has a Poisson distribution with mean $\mu(B)$.*

PROOF. (i) $\Rightarrow$ (ii). If N is Poisson with intensity μ, then it is Poisson with intensity μ on any $B \in \hat{\mathcal{S}}$ with $\mu(B) > 0$. Thus, by Lemma 7.2, N on B is equal in distribution to M specified in (ii).

(ii) $\Rightarrow$ (i) To prove N is Poisson with intensity μ, it suffices by Proposition 6.3 to show

$$E[e^{-Nf}] = \exp\{-\int_S (1 - e^{-f(x)})d\mu(x)\}, \qquad f \in C_K^+(S). \tag{7.4}$$

Choose $B \in \hat{\mathcal{S}}$ that contains $\{x : f(x) > 0\}$ (the support of f) and $\mu(B) > 0$. Under the conditions in (ii), N on B is equal in distribution to a mixed binomial process M on B. Then $E[e^{-Nf}] = E[e^{-Mf}]$ by Proposition 6.2. Furthermore, by Lemma 7.2, M is a Poisson process on B with intensity measure $\mu(B)F(\cdot) = \mu(\cdot \cap B)$, and so $E[e^{-Mf}]$ equals the right-hand side of (7.4) by Proposition 6.3. These observations prove (7.4). □

We are now ready to establish that Poisson processes exist.

THEOREM 7.5. (Existence of Poisson Processes) *There exists a Poisson process N on S with intensity measure μ.*

PROOF. First note that a mixed binomial process exists since it is a function of an infinite collection of random variables, which exist by Theorem 6.1 in the Appendix (i.e., one can construct a probability space and the independent random variables on it). Then a Poisson process N with a "finite" intensity measure μ exists since it is a mixed binomial process by Lemma 7.2.

Next, consider the case when μ is infinite. Choose bounded sets $B_1, B_2, \ldots$ in $\mathcal{S}$ that partition S such that $\mu(B_n) > 0$. By the preceding part of the proof, there exists a Poisson process N_n on S, for each n with intensity $\mu_n(\cdot) \equiv \mu(\cdot \cap B_n)$. By Theorem 6.1, we can define these N_n on a common probability space so that they are independent. Then define $N = \sum_n N_n$. By Theorem 6.4, N is a Poisson process on S with intensity $\sum_n \mu_n = \mu$. $\square$

We end this section with a criterion for a Poisson process to be simple. A point process N on S is *simple* if $P\{N(\{x\}) \leq 1, x \in S\} = 1$ (i.e., its points are distinct).

PROPOSITION 7.6. *A Poisson process N with intensity measure μ is simple if and only if $\mu(\{x\}) = 0$, $x \in S$. Hence any Poisson process on an Euclidean space is simple its its intensity has the form $\mu(B) = \int_B \lambda(x)dx$, for some rate function $\lambda(x)$.*

PROOF. In light of Theorem 7.3, it suffices to prove this when μ is finite. In this case, N is equal is distribution to a mixed binomial process $M = \sum_{i=1}^{\kappa} \delta_{X_i}$ as in Lemma 7.2, where each X_i has distribution $F(\cdot) = \mu(\cdot)/\mu(S)$ and κ has a Poisson distribution with mean $\mu(S)$. Now

$$P\{N \text{ is simple}\} = P\{M \text{ is simple}\} = \sum_{n=2}^{\infty} P\{D_n\}P\{\kappa = n\},$$

where $D_n \equiv \{X_1, \ldots, X_n \text{ are distinct}\}$. Then N is simple if and only if $P\{D_n\} = 1$, for each $n \geq 2$. The latter statement is true, by Exercise 28, if and only if $F(\{x\}) = \mu(\{x\})/\mu(S) = 0$, $x \in S$. Hence N is simple if and only if $\mu(\{x\}) = 0$, $x \in S$. $\square$

8. Deterministic Transformations of Poisson Processes

This section addresses the following issue: If the points of a Poisson process are mapped to some space by a deterministic map, then do these points also form a Poisson process? The answer is yes, provided only that the intensity measure for the new process is locally finite. The next section proves a similar result for more general random transformations of Poisson processes.

For the following result, suppose N is a Poisson process on S with intensity measure μ. Consider a transformation of N in which its points in S are mapped to a space S' (possibly S). Specifically, suppose that a point of N located at $x \in S$ is mapped to a new location $g(x) \in S'$, where $g : S \to S'$. Then the number of points mapped into $B \in \mathcal{S}'$ is

$$N'(B) \equiv \sum_n \delta_{g(X_n)}(B), \qquad B \in \mathcal{S}'.$$

This N' is a point process on S', provided it is locally finite. This is true, of course, when the mean measure $E[N'(B)]$ is locally finite. We say that N' is a *transformation of N under the map g*.

It is convenient to represent the transformed process N' by the inverse of g, which is

$$g^{-1}(B) \equiv \{x \in S : g(x) \in B\}, \qquad B \in \mathcal{S}'.$$

Noting that $\delta_{g(X_n)}(B) = \delta_{X_n}(g^{-1}(B))$, we can write

$$N'(B) = N(g^{-1}(B)), \qquad B \in \mathcal{S}'.$$

Keep in mind that there may be multiple points at a single location and g need not be one-to-one.

THEOREM 8.1. *In the preceding setting, the transformed process N' is a Poisson process on S' with intensity $E[N'(B)] = \mu(g^{-1}(B))$, $B \in \mathcal{S}'$, provided this measure is locally finite.*

PROOF. For any $f : S' \to \Re_+$, using Proposition 6.3 and a change of variable in the integral, we have

$$\begin{aligned} E[e^{-N'f}] &= E[e^{-\sum_n f(g(X_n))}] = E[e^{-\int_S f(g(x))N(dx)}] \\ &= \exp\{-\int_S (1 - e^{-f(g(x))})\mu(dx)\} = \exp\{-\int_{S'} (1 - e^{-f(x')})\mu g^{-1}(dx)\}. \end{aligned}$$

By Proposition 6.3, the last expression is the Laplace functional of a Poisson process with intensity $\mu g^{-1}(\cdot)$ and hence N' is such a process. □

EXAMPLE 8.2. Suppose N is a Poisson process in the nonnegative quadrant $S = \Re_+^2$ of the plane with intensity measure μ. Let $N'(r)$ denote the number of points of N within a distance r from the origin (i.e., in the set $D_r \equiv \{(x, y) \in S : \sqrt{x^2 + y^2} \le r\}$). We can represent N' as a mapping of N in which a point (x, y) of N is mapped to its distance $g(x, y) = \sqrt{x^2 + y^2}$ from the origin. Then

$$N'(r) = N(g^{-1}([0, r])) = N(D_r).$$

By Theorem 8.1, N' is a Poisson on $\Re_+$ with intensity $E[N'(r)] = \mu(D_r)$. This intensity is clearly finite for each r. For instance, if the Poisson process N is homogeneous with a constant rate λ, then $E[N'(r)] = \lambda\pi r^2$.

Further information about a transformation of a Poisson process can be obtained by considering the points in the domain as well as the range of the transformation. Specifically, consider the transformation of the Poisson process N in Theorem 8.1, where each point x of N is mapped to $g(x) \in S'$ and $N' = \sum_n \delta_{g(X_n)}$ is the resulting point process on S'. This transformation can also be represented by the point process M on the product space $S \times S'$ defined by

$$M(A \times B) = \sum_n \delta_{(X_n, g(X_n))}(A \times B). \tag{8.3}$$

This is the number of points of N in $A \in \mathcal{S}$ that are mapped into $B \in \mathcal{S}'$. Note that M contains the original process $N(\cdot) = M(\cdot \times S')$ as well as the transformed process $N'(\cdot) = M(S \times \cdot)$. The M is an example of a "marked" point process, where $g(X_n)$ is a "mark" associated with X_n. We use marked point processes in the next sections to model more general random transformations of point processes and other phenomena.

COROLLARY 8.4. *The marked point process M defined by (8.3) is a Poisson process with intensity*

$$E[M(A \times B)] = \mu(A \cap g^{-1}(B)), \quad A \in \mathcal{S}, \ B \in \mathcal{S}'. \tag{8.5}$$

PROOF. We can write (8.3) as $M = \sum_n \delta_{h(X_n)}$, where $h(x) = (x, g(x))$. So M is a transformation of N by h. Thus, by Theorem 8.1, M is a Poisson process with intensity (8.5) since $M(A \times B) = N(A \cap g^{-1}(B))$. The intensity (8.5) is locally finite since μ is. $\square$

A basic property of a Poisson process N on a product space $S_1 \times S_2$ is that the projection $N(S_1 \times \cdot)$ on S_2 is also a Poisson process provided its intensity $E[N(S_1 \times \cdot)]$ is locally finite. This follows immediately from the definition of a Poisson process. Here is an extension of this fact.

EXAMPLE 8.6. *Projections of a Poisson Process.* Consider a Poisson process $N = \sum_n \delta_{\mathbf{X}_n}$ on $S \subset S_1 \times \cdots S_m$, with intensity μ, where $\mathbf{X}_n = (X_n^1, \ldots, X_n^m)$. Let $N^i = \sum_n \delta_{X_n^i}$ denote the projection of N on the subspace $S^i = \{x^i : \mathbf{x} \in S\}$, and let $M^i = \sum_n \delta_{(\mathbf{X}_n, X_n^i)}$ be the marked point process that describes the points in the domain and range of the mapping of N by the projection map $g_i(\mathbf{x}) = x^i$. By Corollary 8.4, M^i is Poisson with intensity $\mu^i(A \times B) = \mu\{\mathbf{x} \in A : x^i \in B\}$. Hence $N^i(\cdot) = M^i(S \times \cdot)$ is a Poisson process with intensity $E[N^i(B)] = \mu\{\mathbf{x} \in S : x^i \in B\}$, provided this intensity is locally finite. For instance, suppose N is a homogeneous Poisson process on $\Re_+^m$ with rate λ. Then $N^i(0, b] = \infty$ a.s. However, M^i still gives insights on the projection, since each $M^i(A \times \cdot)$, for A fixed, is a Poisson process on S^i describing the projection coming from points in A.

Next, consider the more general projection $N_I = \sum_n \delta_{g_I(\mathbf{X}_n)}$ on the space $S_I = \{g_I(\mathbf{x}) : \mathbf{x} \in S\}$, where $g_I(\mathbf{x}) \equiv (x^i : i \in I)$, for $I \subset \{1, \ldots, m\}$. The related process $M_I = \sum_n \delta_{(\mathbf{X}_n, g_I(\mathbf{X}_n))}$ on $S \times S_I$ is Poisson by Corollary 8.4. Hence $N_I(\cdot) = M(S \times \cdot)$ is Poisson with $E[N_I(B)] = \mu\{\mathbf{x} \in S : g_I(\mathbf{x}) \in B\}$, provided this is locally finite.

EXAMPLE 8.7. Let $N = \sum_n \delta_{(X_n, Y_n)}$ denote a Poisson process on the unit disc S in $\Re^2$ with rate function $\lambda(x, y)$. Consider the projection of N on the interval $S' = [-1, 1]$, which is described by the process $N' \equiv \sum_n \delta_{X_n}$ on S'. By Example 8.6, N' is Poisson with

$$E[N'(a, b]] = \int_a^b \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \lambda(x, y)\, dy\, dx.$$

More generally, projections on $S' = [-1, 1]$ from points located in sets like $A_u = \{(x, y) \in S : y \geq u\}$, $u \in (0, 1]$ are described by $M \equiv \sum_n \delta_{((X_n, Y_n), X_n)}$

on $S \times S'$, which is Poisson with

$$E[M(A_u \times (a,b])] = \int_a^b \int_u^1 \lambda(x,y)\,dy\,dx.$$

Next, consider the transformation $\overline{N}$ of N under which a point in the unit disc S is mapped to the closest point on the unit circle C. To define $\overline{N}$, it is convenient to represent a point in S by its polar coordinates (r,θ), where $(x,y) = (r\cos\theta, r\sin\theta)$, and to view $N = \sum_n \delta_{(R_n,\Theta_n)}$ as a Poisson process on $S = \{(r,\theta) \in [0,1] \times [0,2\pi)\}$ with rate function $\lambda(r\cos\theta, r\sin\theta)$. The unit circle can be expressed as $C = [0,2\pi)$, since each point on the circle has the form $(1,\theta)$. The transformation under consideration maps a point at (r,θ) to $(1,\theta)$ (i.e., to $\theta \in C$), and so the transformed process is $\overline{N} = \sum_n \delta_{\Theta_n}$, which is simply the projection of N on the coordinate set C. Therefore $\overline{N}$ is Poisson with

$$E[\overline{N}(B)] = \int_B \int_0^1 \lambda(r\cos\theta, r\sin\theta)dr\,d\theta.$$

See Exercise 25 for more details on these processes.

9. Marked Poisson Processes

In the preceding section, we showed how a deterministic transformation of a Poisson process can be modelled by a marked Poisson process. We now describe more general marked point process models.

We will use the following notation throughout this section. Let $N = \sum_n \delta_{X_n}$ be a point process on S. On the same underlying probability space, suppose $M = \sum_n \delta_{(X_n,Y_n)}$ is a point process on a product space $S \times S'$. We call M a *marked point process* associated with N, and refer to Y_n as a *mark* of X_n. The mark Y_n may be an attribute of X_n or any random element whose dependency on X_n or N would be determined by the application.

Here are two examples:

- M represents a “random” transformation of N in which X_n is transformed to Y_n (e.g., $Y_n = g(X_n)$ is the deterministic transformation studied in the preceding section).
- M is a spatial location-demand model in which $X_n \in \Re^2$ is the location of a demand and $Y_n = (D_n, Q_n)$, where D_n is the type of demand and Q_n is the size of the demand (quantity of a product or length of a service being demanded).

The marked point process M contains N because $N(\cdot) = M(\cdot \times S')$ (the projection of M on the subspace S). Note that $M(S \times \cdot) = \sum_{n=1}^{N(S)} \delta_{Y_n}(\cdot)$ may not be a point process. For instance, if $N(S) = \infty$ and Y_n are i.i.d. with $P\{Y_1 \in B\} > 0$, then $M(S \times B) = \infty$ by the strong law of large numbers. We call M a marked Poisson process if it is Poisson in the usual sense. In this case, N would also necessarily be Poisson since it is part of M.

We will now describe the marked point process M associated with the point process N under the condition that each mark Y_n has a distribution that depends only on its associated X_n.

EXAMPLE 9.1. *p-Marked Point Processes.* Suppose the marked point process M is such that a mark associated with a location $x \in S$ takes a value in a set $B \in \mathcal{S}'$ with probability $p(x, B)$, independent of everything else. The $p(x, B)$ is a *probability kernel from S to S'*, which means $p(\cdot\,, B)$ is a measurable map for each $B \in \mathcal{S}'$, and $p(x, \cdot)$ is a probability on S' for each $x \in S$.

The $p(x, \cdot)$ is a conditional distribution of a mark associated with a location x as follows. Consider a fixed bounded set $A \in \hat{\mathcal{S}}$, and denote the points of M on the subspace $A \times S'$ by $(X_{n_k} Y_{n_k})$, $1 \le k \le N(A)$. Then the assumption on the marks is that, for any $B_1, B_2 \ldots \in \hat{\mathcal{S}}'$,

$$\begin{aligned} P\{Y_{n_k} \in B_k,\, 1 \le k \le N(A) \Big| N\} &= \prod_{k=1}^{N(A)} P\{Y_{n_k} \in B_k | X_{n_k}\} \\ &= \prod_{k=1}^{N(A)} p(X_{n_k}, B_k). \end{aligned} \tag{9.2}$$

That is, the Y_{n_k}'s are conditionally independent given N, and $p(X_{n_k}, B)$ is the conditional distribution of Y_{n_k} given X_{n_k}.

DEFINITION 14. In the preceding context, the point process M that satisfies (9.2) is a *p-marked point process* associated with N. Also, Y_n are *location-dependent marks of X_n*, and a mark associated with a point of N at x has the distribution $p(x, \cdot)$.

The Laplace functional of M and its intensity are as follows.

PROPOSITION 9.3. *Suppose M is a p-marked point process associated with the point process N. Then, for any $f \in C_K^+(S \times S')$,*

$$E[e^{-Mf}] = E[e^{\int_S \log h(x) N(dx)}], \tag{9.4}$$

where $h(x) = \int_{S'} e^{-f(x,y)} p(x, dy)$. If N has an intensity measure μ, then the intensity of M is given by

$$E[M(A \times B)] = \int_A p(x, B)\mu(dx), \quad A \in \mathcal{S},\ B \in \mathcal{S}'. \tag{9.5}$$

PROOF. For a fixed $f \in C_K^+(S \times S')$, let $A \in \hat{\mathcal{S}}$ be such that $f(x, y) = 0$ for $(x, y) \notin A \times S'$. Then by property (9.2) and $Mf = \sum_{k=1}^{N(A)} f(X_{n_k}, Y_{n_k})$,

$$E[e^{-Mf}] = E\Big[\prod_{k=1}^{N(A)} E[e^{-f(X_{n_k}, Y_{n_k})} | N]\Big] = E\Big[\prod_{k=1}^{N(A)} h(X_{n_k})\Big].$$

By the selection of A,

$$\prod_{k=1}^{N(A)} h(X_{n_k}) = \prod_n h(X_n) = e^{\sum_n \log h(X_n)} = e^{\int_S \log h(x) N(dx)}.$$

Using this in the preceding display yields (9.4).

To prove (9.5), it suffices to show, for $f \in C_K^+(S \times S')$, that $E[Mf] = \mu \bar{f}$, where $\bar{f}(x) = \int_{S'} f(x,y)p(x,dy)$. Using property (9.2), we have

$$\begin{aligned} E[Mf] &= E\Big[E\Big[\sum_{k=1}^{N(A)} f(X_{n_k}, Y_{n_k})\Big|N\Big]\Big] \\ &= E\Big[\sum_{k=1}^{N(A)} \bar{f}(X_{n_k})\Big] = E\Big[\sum_n \bar{f}(X_n)\Big]. \end{aligned}$$

Then $E[Mf] = \mu \bar{f}$ follows by Theorem 6.6. $\square$

The next result establishes that a p-marked point process associated with a Poisson process is Poisson.

THEOREM 9.6. *Suppose M is a p-marked point process associated with a Poisson process N with intensity μ. Then M is a Poisson process on $S \times S'$ with intensity measure given by (9.5). In addition, the process of marks $N' = M(S \times \cdot) = \sum_n \delta_{Y_n}$ is Poisson with intensity $E[N'(B)] = \int_S p(x,B)\mu(dx)$, $B \in \mathcal{S}'$, provided this measure is locally finite.*

PROOF. By Proposition 9.3, we know that, for any $f \in C_K^+(S \times S')$, $E[e^{-Mf}] = E[e^{Ng}]$, where $g(x) = \log[\int_{S'} e^{-f(x,y)} p(x,dy)]$. Also, because N is Poisson with intensity μ and $-g \in C_K^+(S)$, it follows by Proposition 6.3 that

$$\begin{aligned} E[e^{-Mf}] &= E[e^{Ng}] = \exp\{-\int_S (1 - e^{g(x)})\mu(dx)\} \\ &= \exp\{-\int_{S \times S'} (1 - e^{-f(x,y)}) p(x,dy)\mu(dx)\}. \end{aligned}$$

Thus M is Poisson with intensity (9.5) by Proposition 6.3. The process N' is also Poisson since it is the projection of the Poisson process M on S'. $\square$

The preceding result establishes that a p-marked point process of a Poisson process is a marked Poisson process. The converse is also true as follows.

REMARK 9.7. If M is a marked Poisson process on $S \times S'$, then M is a p-marked point process associated with $N(\cdot) \equiv M(\cdot \times S')$, where the probability kernel $p(x,B)$ is defined by (9.5), with μ being the intensity of N. This result is based on the fact that the intensity of M can be decomposed as in (9.5), where, for each fixed B, the $p(x,B)$ as a function of x is the Radon–Nikodym derivative of $E[M(\cdot \times B)]$ with respect to μ. Since this intensity is equal to that of the marked Poisson process in Theorem 9.6, it follows that M is equal in distribution to that marked Poisson process.

A basic property of Markov chains, related to transformations of Poisson processes, is as follows.

EXAMPLE 9.8. *Markov/Poisson Particle System.* Consider a system of particles that reside in a space S. At time 0, the particles are located in S according to a Poisson process $N_0(\cdot)$ with intensity measure μ. Thereafter, each particle moves independently in S at discrete times such that a particle located at x at time n moves to a set B at time $n+1$ with probability $p(x, B)$. In other words, each particle moves independently according to a discrete-time Markov chain on S with probability kernel $p(x, B)$. A particle in state x at time 0 will be in a set B at time n with probability $p^n(x, B)$, which is the n-step probability defined by

$$p^n(x, B) = \int_S p^{n-1}(y, B)p(x, dy), \qquad n \geq 1.$$

When S is countable, the matrix $(p^n(x, y))$ is the nth product of the matrix $(p(x, y))$.

Our interest is in the sequence of point processes $\{N_n : n \geq 0\}$ on S, where $N_n(B)$ denotes the number of particles in $B \in \mathcal{S}$ at time n. We will also consider the more general marked point process M_n on S^2, where $M_n(A \times B)$ denotes the number of particles in A at time 0 that are in B at time n.

PROPOSITION 9.9. *Suppose the intensity measure μ for the Poisson process N_0 is an* invariant measure *of $p(x, B)$ in that*

$$\int_S p(x, B)\mu(dx) = \mu(B), \qquad B \in \mathcal{S}. \tag{9.10}$$

Then M_n is a Poisson process on S^2 with

$$E[M_n(A \times B)] = \int_A p^n(x, B)\mu(dx), \qquad A, B \in \mathcal{S}.$$

In addition, $\{N_n : n \geq 0\}$ is a stationary Markov chain whose stationary distribution is that of a Poisson process on S with intensity measure μ.

PROOF. Clearly, M_n is a marked transformation of the Poisson process N_0 under the probability kernel $p^n(x, B)$, and so the first assertion follows by Theorem 9.6. Next, note that $\{N_n : n \geq 0\}$ is a Markov chain because the particles move independently by the Markov chain probability $p(x, B)$. To prove this Markov chain N_n is stationary and its stationary distribution is that of N_0, it suffices to show that N_n is Poisson with intensity μ. But this follows since $N_n(\cdot) = M_n(S \times \cdot)$, where M_n is Poisson, and

$$E[N_n(B)] = \int_S p^n(x, B)\mu(dx) = \mu(B).$$

The last equality follows by an induction argument using (9.10). □

Marked point processes are frequently used to model random phenomena in space as well as time as follows.

Example 9.11. *Space-Time Poisson Processes.* Suppose $N = \sum_n \delta_{T_n}$ is a point process on $\Re_+$ representing the times $0 \le T_1 \le T_2 \le \dots$ at which an event occurs. Suppose the nth event at time T_n involves auxiliary information denoted by a random element Y_n in a space S. For instance, T_n might be the arrival time of the nth item to a service system, and Y_n is its required service time. Then a convenient way of modelling these random elements over time is by the marked point process $M = \sum_n \delta_{(T_n,Y_n)}$ on $\Re_+ \times S$. This is called a *space-time* process. If M is a Poisson process in the usual sense, we call it a space-time Poisson process.

For instance suppose N is Poisson with rate λ, and the marks Y_n have the location-dependent distribution $p(t,\cdot)$. Then by Theorem 9.6, M is a space-time Poisson process with intensity $E[M((a,b]\times B)] = \lambda \int_a^b p(t,B)dt$.

10. Partitions and Translations of Poisson Processes

This section describes several types of transformations of Poisson processes that arise in a variety of applications. These transformations are further illustrations of marked Poisson process models.

Example 10.1. *Thinning of a Poisson Process.* Let N be a Poisson process on S with intensity μ. Suppose the points of N are deleted according to the rule that a point at x is retained with probability $p(x)$, and the point is deleted with probability $1 - p(x)$. Let N_1 and N_2 denote the resulting processes of retained and deleted points, respectively. Note that $N = N_1 + N_2$. By Corollary 10.2 below, N_1 and N_2 are independent Poisson processes with respective mean measures

$$E[N_1(A)] = \int_A p(x)\mu(dx), \quad E[N_2(A)] = \int_A (1-p(x))\mu(dx), \quad A \in \mathcal{S}.$$

Interestingly, N_1 and N_2 are independent even though $N = N_1 + N_2$.

As an example, suppose a web site that sells products has visitors arriving to it according to a Poisson process N with rate λ. Suppose p percent of these visitors buy a product, which means that each visitor independently buys a product with probability p. Then from the preceding result, the times of sales form a Poisson process with rate $p\lambda$, and the visits without sales occur according to a Poisson process with rate $(1-p)\lambda$.

Thinning of a point process is a special case of the following partitioning procedure for decomposing a point process into several subprocesses. Consider a Poisson process N on S with intensity μ. Suppose N is partitioned into a countable family of processes N_i, $i \in I$, on S by the following rule.
Partitioning Rule: A point of N at x is assigned to subprocess N_i with probability $p(x,i)$, where $\sum_{i\in I} p(x,i) = 1$.

The processes N_i form a partition of N in that $N = \sum_{i\in I} N_i$.

Corollary 10.2. (Partitioning of a Poisson Process) *The subprocesses N_i, $i \in I$, of the Poisson process N are independent Poisson processes with* *intensities*

$$E[N_i(B)] = \int_B p(x,i)\mu(dx), \quad B \in \mathcal{S},\ i \in I.$$

Proof. Let $M(B \times \{i\})$ denote the number of points of N in B that are assigned to N_i. That is, $M(B \times \{i\}) = N_i(B)$. Clearly, M is a $p(x,i)$-marked point process on $S \times I$ associated with N, and so M is Poisson by Theorem 9.6. Since M has independent increments and the subprocesses N_i represent M on the disjoint subsets $S \times \{i\}$, for $i \in I$, they are independent Poisson processes. Furthermore,

$$E[N_i(B)] = E[M(B \times \{i\})] = \int_B p(x,i)\mu(dx).$$

□

The preceding result for partitions is the opposite of the result that a sum of independent Poisson processes is also Poisson (recall Theorem 6.4).

Example 10.3. Telephone calls in a region S of the USA are assumed to occur according to a space-time Poisson process M on $\Re_+ \times S$, where $M((0,t] \times B)$ denotes the number of calls connected in the subregion $B \subset S$ in the time interval $(0,t]$, and $E[M((0,t] \times B)] = \lambda t \mu(B)$. There are three types of calls: (1) Long distance calls outside the USA. (2) Long distance calls within the USA. (3) Local calls. The calls are independent, and a call at time t and location x is a type i call with probability $p(t,x;i)$, $i = 1,2,3$. Then by Corollary 10.2, the number of type i calls occur according to a space-time Poisson process

$$E[M_i((0,t] \times B)] = \lambda \int_0^t \int_B p(s,x;i)\mu(dx)ds.$$

Furthermore, M_1, M_2, M_3 are independent and $M = M_1 + M_2 + M_3$.

Splitting and merging of flows in a network, as we now show, are typical examples of partitioning and summing of point processes.

Example 10.4. *Routing in A Graph.* Consider the directed graph shown in Figure 1 in which units are routed in the directions of the arrows. Let $N_{jk}(t)$ denote the number of units that are routed on the arc from node j to node k in the time interval $(0,t]$. Assume that items enter the graph by independent Poisson processes N_{0j}, $j = 1,2,3$ on $\Re_+$ with respective rates λ_{0j}, $j = 1,2,3$. Upon entering the graph, each item is routed independently through the graph according to the probabilities on the arcs, and there are no delays at the nodes (travel through the graph is instantaneous). For instance, an item entering node 3 is routed to node 5 or node 6 with respective probabilities p_{35} and p_{36}, where $p_{35} + p_{36} = 1$.

Our results on partitioning and sums of Poisson processes yield the following properties. First note that flows N_{12} and N_{13} are independent Poisson processes with rates $\lambda_{12} = p_{12}\lambda_{01}$ and $\lambda_{13} = p_{13}\lambda_{01}$, since they are partitions of N_{01}. Next, the flow into node 2 is the sum $N_{02} + N_{12}$ (of independent

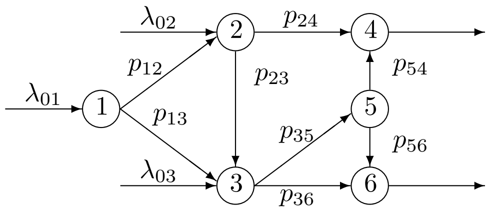

FIGURE 1. Partitioning and Merging of Flows

flows) and hence is Poisson with rate $\lambda_{02}+p_{12}\lambda_{01}$. Similar properties extend to the other flows in the graph. Specifically, each flow N_{jk} from j to k is a Poisson process, and one can evaluate their rates λ_{jk} in the obvious way. For instance, knowing λ_{12} and λ_{13} as mentioned above,

$$\begin{aligned} \lambda_{23} &= p_{23}(\lambda_{02}+\lambda_{12}), & \lambda_{36} &= p_{36}(\lambda_{03}+\lambda_{13}+\lambda_{23}), \\ \lambda_{35} &= p_{35}(\lambda_{03}+\lambda_{13}+\lambda_{23}), & \lambda_{60} &= \lambda_{36}+p_{56}\lambda_{35}. \end{aligned}$$

Also, some of the flows are independent (denoted by $\perp$). Examples are $N_{12} \perp N_{13}$, $N_{36} \perp N_{56}$, $N_{36} \perp N_{24}$, and $N_{13} \perp N_{24}$. On the other hand, many flows are not independent (denoted by $\not\perp$). Examples are $N_{12} \not\perp N_{24}$, $N_{35} \not\perp N_{40}$, $N_{13} \not\perp N_{60}$, and $N_{23} \not\perp N_{40}$.

In addition, the flow $N_k = \sum_i N_{jk}$ through each node k is a Poisson process with intensity $\sum_j \lambda_{jk}$. Clearly all the N_k's are dependent. If the arc between 5 and 4 did not exist, however, then N_4 would be independent of N_3, N_5, and N_6.

Another example of a transformation of a Poisson process is as follows. Suppose N is a Poisson process on $S = \Re^d$ with intensity measure μ. Assume that a point of N at x is independently translated to another location $x+Y$ by a random vector Y in S that has a distribution $G_x(\cdot)$. That is, x is mapped into a set $B \subset S$ by a probability kernel $p(x,B) = G_x(B-x)$, where $B - x = \{y - x : y \in B\}$. Let $M(A \times B)$ denote the number of points of N in A that are translated into B. This M is a p-marked point process associated with the Poisson process N. Thus, Theorem 9.6 yields the following result.

COROLLARY 10.5. (Translation of a Poisson process) *The marked translation process M defined above is Poisson with*

$$E[M(A\times B)] = \int_A G_x(B-x)\mu(dx), \qquad A, B \subset \Re^d_+.$$

In particular, the process $N'(B) \equiv M(A\times B)$ denoting the number of points of N translated into B is Poisson with $E[N'(B)] = \int_S G_x(B-x)\mu(dx)$.

EXAMPLE 10.6. *Trees in a Forest.* The locations (X_n, Y_n) of a certain type of tree in a forest form a Poisson process with intensity measure μ. Suppose the height of a tree at a location (x,y) has a distribution $G_{x,y}(\cdot)$.

That is, the heights Z_n of the trees at the respective locations (X_n, Y_n) are marks, and $M \equiv \sum_n \delta_{(X_n,Y_n,Z_n)}$ forms a Poisson process on $\Re^2 \times \Re_+$ with

$$E[M(A \times B \times (0,b])] = \int_A \int_B G_{x,y}(b)\mu(dx\,dy).$$

After several years of growth, it is anticipated that the increase in height for a tree has a distribution $H_{(x,y,z)}(\cdot)$, where (x,y) is the location and z is the original height. In other words the increases Z'_n are p-marks of (X_n, Y_n, Z_n) with $p((x,y,z),\cdot) = H_{(x,y,z)}(\cdot)$. Then the collection of trees is depicted by the point process $M' \equiv \sum_n \delta_{(X_n,Y_n,Z_n+Z'_n)}$. By Corollary 10.5, M' is a Poisson process with

$$E[M'(A \times B \times (0,b])] = \int_A \int_B \int_{\Re_+} H_{(x,y,z)}(b-z)G_{x,y}(dz)\mu(dx\,dy).$$

In the preceding example, a little more realism could be added by considering the possibility that while some trees grow as indicated, other trees may die according to a location-dependent thinning. Then one would have a combined translation–thinning transformation. Similarly, complicated systems might involve transformations involving a combination of translations, thinnings, partitions, deterministic maps and random transformations.

11. Space-Time Poisson Models

This section consists of three examples that illustrate how one can analyze complex systems by transformations of space-time Poisson processes. Related examples are in the next section and in Exercises 37, 39 and 38.

EXAMPLE 11.1. *Maxima of Marks for a Poisson Process.* Suppose $N = \sum_n \delta_{T_n}$ is a Poisson process on $\Re_+$ with intensity measure μ, and Y_n are real-valued p-marks of T_n where $p(t,\cdot)$ is the distribution of a mark at time t. Consider the stochastic process

$$Y(t) = \max_{n \le N(t)} Y_n, \qquad t \ge 0,$$

where $Y(t) = 0$ when $N(t) = 0$. One can obtain information about this maxima process in terms of the space-time Poisson process $M = \sum_n \delta_{T_n,Y_n}$ on $\Re_+ \times \Re$ with $E[M((0,t] \times B)] = \int_{(0,t]} p(s,B)\mu(ds)$. For instance, the event $\{Y(t) \le y\}$ equals $\{M((0,t] \times (y,\infty)) = 0\}$, and so

$$P\{Y(t) \le y\} = e^{-\int_{(0,t]} p(s,(y,\infty))\mu(ds)}.$$

Also, if $\mu(t) = \lambda t$ and $p(t,\cdot) = G(\cdot)$, independent of t, then $Y(t)$ is a Markov chain subordinated to the Poisson process N, and

$$\begin{aligned} P\{Y(s+t) \le y | Y(s) = x\} &= \mathbf{1}(x \le y)P\{M((s,s+t] \times (y,\infty)) = 0\} \\ &= \mathbf{1}(x \le y)e^{-\lambda t(1-G(y))}. \end{aligned}$$

EXAMPLE 11.2. *$M_t/G_t/\infty$ Service System.* Items arrive to a service system at times that form a Poisson process on $\Re_+$ with intensity measure μ. An item that arrives at time t spends a random amount of time in the system that has a distribution $G_t(\cdot)$ and then departs. This sojourn time is independent of the other items in the system and everything else. Keep in mind that items may arrive in batches if there are times t for which $\mu(\{t\}) > 0$. The number of arrivals at such a time has a Poisson distribution with mean $\mu(\{t\})$, but the items in this batch may not depart at the same time since their sojourn times are independent. Let $Q(t)$ denote the quantity of items in the system at time t that arrived after time 0. We are not considering items that may be in the system at time 0. The process $\{Q(t) : t \geq 0\}$ is a $M_t/G_t/\infty$ process with time-dependent arrivals and services.

The process $Q(t)$ is a typical model for the quantity of items in a service system with a large number of parallel servers (envisioned as infinite servers) in which there is essentially no queueing prior to service. For instance, $Q(t)$ could be the number of: (1) Computers being used in a wireless network with a high capacity. (2) Groups of people dining in a cafeteria. (3) Vehicles in a parking lot. (4) Patients in a hospital. (5) Calls being processed in a call center.

To analyze the process $Q(t)$ the first step is to define it by the system data. The data is represented by the marked point process $M \equiv \sum_n \delta_{(T_n,V_n)}$ on $\Re_+^2$, where T_n is the arrival time of the nth item and V_n is its sojourn or service time. The V_n are location-dependent marks of T_n with the distribution $p(t,B) = G_t(B)$. Then by Theorem 9.6, M is a space-time Poisson process with $E[M(A \times B)] = \int_{(0,t]} G_s(B)\mu(ds)$.

Since the quantity $Q(t)$ is a function of the arrival and departure times of the items, let us consider the marked point process

$$N \equiv \sum_n \delta_{(T_n,T_n+V_n)}, \qquad \text{on } S = \{(t,y) \in \Re_+^2 : y \geq t\}.$$

This process N is a transformation of M under the map $g(t,v) = (t, t+v)$. Then N is a space-time Poisson process by Theorem 8.1. In particular, $N((a,b] \times (c,d])$, where $b \leq c$, is the number of items that arrive in $(a,b]$ and depart in $(c,d]$, and its mean is

$$E[N((a,b] \times (c,d])] = \int_{(a,b]} [G_s(d-s) - G_s(c-s)]\mu(ds).$$

Using the preceding notation, the quantity of items in the system at time t is defined by

$$Q(t) = \sum_n \mathbf{1}(,T_n \leq t, T_n + V_n > t) = N((0,t] \times (t,\infty]).$$

Since N is a Poisson process, it follows that $Q(t)$ has a Poisson distribution with

$$E[Q(t)] = \int_{(0,t]} [1 - G_s(t-s)]\mu(ds). \tag{11.3}$$

Although the distribution of each $Q(t)$ is Poisson, the entire process, of course, is not Poisson.

In addition to analyzing the number of items in the system, one may want information about the departure process. This is useful when the departures form an arrival process into another service system. Now, the total number of departures in $(0,t]$ is $D(t) = \sum_n \delta_{T_n+V_n}((0,t])$. That is, D is the projection of N on its second coordinate, and so D is a Poisson process with

$$E[D(t)] = E[N(t) - Q(t)] = \int_{(0,t]} G_s(t-s)\mu(ds).$$

Finally, consider the special case in which the Poisson arrival process is homogeneous with rate λ and the service distribution $G_t(\cdot) = G(\cdot)$ is independent of t. The system is called an $M/G/\infty$ system with arrival rate λ and service distribution G. Then $Q(t)$ has a Poisson distribution and, using a change-of-variable $u = t - s$ in (11.3),

$$E[Q(t)] = \lambda \int_0^t [1 - G(u)]du.$$

Exercise 38 shows that when G has a mean α, the limiting distribution of $Q(t)$ is Poisson with rate $\lambda\alpha$, as $t \to \infty$. Also, the total number of departures $D(t)$ in the time interval $(0,t]$ is a Poisson process with

$$E[D(t)] = \lambda \int_0^t G(s)ds.$$

An abstraction of the preceding model is as follows.

Example 11.4. *Poisson Input-Output-Mobility Model.* Consider a system in which items enter a space S at times $T_1 \le T_2 \le \dots$ that form a Poisson process with intensity measure μ. The nth item that arrives at time T_n moves in the space S for a while and then exists the system (by entering the outside state 0). The movement is determined by a stochastic process $Y_n \equiv \{Y_n(t) : t \ge 0\}$ with state space $S \cup \{0\}$, where the outside 0 is an absorbing state ($Y_n(t) = 0$ for all $t > \inf\{s : Y_n(s) = 0\}$). Specifically, the nth item enters S at the location $Y_n(0)$, and, at time $t > T_n$ its location is $Y_n(t - T_n)$. Let $\mathcal{Y}$ denote a function space that contains the sample paths of Y_n (e.g., $\mathcal{Y}$ could be a space of real-valued functions that are continuous, or piece-wise constant).

Assume the Y_n are location-dependent marks of T_n with distribution $p(t,\cdot)$ which is the conditional distribution of the process Y_n starting at time t. For simplicity, assume Y_n depends on t only through its initial value

$Y_n(0)$ (the entry point in S of the nth point), whose distribution is denoted by $F_t(\cdot)$. Then conditioning on $Y_n(0)$,

$$p(t,\cdot) = \int_S P\{Y_n \in \cdot | Y_n(0) = x\} F_t(dx). \tag{11.5}$$

In other words, the system data consists of the point process $M \equiv \sum_n \delta_{(T_n, Y_n)}$ on $\Re_+ \times \mathcal{Y}$, which is a space-time Poisson process by Theorem 9.6.

Now, the number of items in the set $B \in \mathcal{S}$ at time t is given by

$$N_t(B) = \sum_n \mathbf{1}(T_n \le t, Y_n(t - T_n) \in B) = \sum_n \delta_{g_t(T_n, Y_n)}(B),$$

where $g_t(s, y) = y(t - s)$ and $y(\cdot) \in \mathcal{Y}$. Since N_t is a transformation of the Poisson process M by the map g_t, it follows by Theorem 8.1 that N_t is a Poisson process on S for each fixed t, and from (11.5),

$$E[N_t(B)] = \int_{(0,t]} \int_S P^{t-s}(x, B) F_s(dx) \mu(ds).$$

where $P^{t-s}(x, B) \equiv P\{Y_n(t - s) \in B | Y_n(0) = x\}$.

Next, note that the number of departures from the set B in the time interval $(a, b]$ is $D((a, b] \times B) = \sum_n \mathbf{1}(h(T_n, Y_n) \in (a, b] \times B)$, where $h(s, y) = (s, y(t - s))$. This D is a transformation of the Poisson process M by the map h, and so by Theorem 8.1, D is a space-time Poisson process on $\Re_+ \times S$ with

$$E[D((0, t] \times B)] = \int_{(0,t]} \int_B P^{t-s}(x, \{0\}) F_s(dx) \mu(ds)$$

12. Network of $M_t/G_t/\infty$ Stations

In this section, we show how the ideas in the preceding section extend to the analysis of flows in a stochastic network of $M_t/G_t/\infty$ stations. The network dynamics are determined by marks of Poisson processes, and the analysis amounts to formulating appropriate Poisson processes that represent parameters of interest, and then specifying their intensity measures.

Consider a network of m service stations (or nodes) that operate as follows. Items enter the network at times $T_1 \le T_2 \le \ldots$ that form a Poisson process with intensity measure μ. The nth item entering the network at time T_n selects, or is assigned, a random route $\mathbf{R}_n = (R_{n1}, \ldots, R_{nL_n})$ through the network, where $R_{nk} \in \{1, \ldots, m\}$ denotes the kth node the item visits, and the length $1 \le L_n \le \infty$ may be random and depend on $\mathbf{S}_n$. After visiting node the last node R_{nL_n} on its route, the item exits the network and enters node 0 ("outside" the network) and stays there forever. In addition, the item selects, or is assigned, a vector of nonnegative sojourn (or visit) times $\mathbf{V}_n = (V_{n1}, \ldots, V_{nL_n})$, where V_{nk} is the item's sojourn time at node R_{nk}. The time at which the item departs from node R_{nk} is

$$\tau_{nk} \equiv T_n + \sum_{j=1}^{k} V_{nj}, \qquad k \le L_n,$$

where τ_{nL_n} is the time at which the item exits the network.

The main assumption is that the route and waiting time vectors $Y_n \equiv (\mathbf{R}_n, \mathbf{V}_n)$ are marks of the arrival times T_n. This implies there are no interactions among the items that affect their waiting times, so each node operates like an $M_t/G_t/\infty$ system. As above, we consider only those items that enter the network "after" time 0. In summary, the system dynamics are represented by the space-time process $M = \sum_n \delta_{(T_n,Y_n)}$, which is Poisson by Theorem 9.6.

Many features of the network are expressible by space-time Poisson processes of the form

$$N_t = \sum_n \delta_{(T_n, g_t(T_n,Y_n))}, \tag{12.1}$$

$$E[N_t((a,b] \times B)] = \int_{(a,b]} P\{g_t(T_n,Y_n)) \in B | T_n = s\} \mu(ds). \tag{12.2}$$

The N_t is Poisson since it is a deterministic transformation of the Poisson process M. The function g_t would depend on the application at hand; in some cases, g_t and N_t do not depend on t. Typical uses of these space-time Poisson processes are as follows.

Locations of Items at Time t. The space-time Poisson process describing where the items are located is

$$N_t((a,b] \times B) = \#\text{ of items that arrive in the time interval } (a,b] \text{ that are in } B \subset S \equiv \{0,1,\dots,m\} \text{ at time } t.$$

The location of the item at time t that arrives at time T_n is

$$g_t(T_n,Y_n) = \begin{cases} 0 & \text{if } \tau_{nL_n} \le t \\ R_{nk} & \text{if } \tau_{n(k-1)} \le t < \tau_{nk}, \text{ for some } k \le L_n. \end{cases} \tag{12.3}$$

In particular, the quantities $Q_i(t) = N_t((0,t] \times \{i\})$, $1 \le i \le m$, at the nodes at time t are independent Poisson random variables with

$$E[Q_i(t)] = \int_{(0,t]} P\{g_t(T_n,Y_n) = i | T_n = s\} \mu(ds). \tag{12.4}$$

Departure Process. The space-time Poisson process describing the times at which items exit the network is

$$N((a,b] \times B) = \#\text{ of items arriving in } (a,b] \text{ whose exit time from the network is in } B \subset \Re_+.$$

The item arriving at T_n exits the network a time $g(T_n,Y_n) = \tau_{nL_n}$.

For typical applications, the mean values (12.2) of the space-time Poisson processes would be determined by the distributions of the routes and sojourn times of the items. The routes depend on the structure of the network and the nature of the items and services. A standard assumption is that the routes are independent and Markovian, where p_{jk} denotes the probability of an item moving to node k upon departing from node j. Then the probability of a particular route $(r_1, \dots, r_\ell)$ of nonrandom length ℓ is $p_{0r_1} \cdots p_{r_\ell 0}$. Another convention is that there are several types of items and all items of the same type take the same route. In this case, the probability of a route is the probability that the item entering the network is the type that takes that route. The simplest sojourn times at a node are those that are i.i.d., depending on the node and independent of everything else. Then the sums of sojourn times are characterized by convolutions of the distributions. The next level of generality is that the service times are independent at the nodes, but their distributions may depend on the route as well as the node. An example of dependent service times is that an item entering a certain subset of routes is initially assigned a service time according to some distribution and then that time is its service time at "each" node on its route.

Here is an example of a particular network.

Example 12.5. *An Acyclic Network.* Consider the stochastic network shown in Figure 2, which operates as described above with the following additional properties. Items arrive at the nodes 1, 2 and 3 from outside according to independent Poisson processes with respective rates λ_1, λ_2, λ_3. The sojourn or service times at the nodes are independent random variables, and the sojourn times at node i have the distribution $G_i(\cdot)$. When an item ends its sojourn at node 1, it departs and enters node 2 with probability p_{12}, or it enters node 3 with probability $p_{13} = 1 - p_{12}$. Analogously, departures from node 2 enter node 3 with probability p_{23}, or enter node 4 with probability $p_{24} = 1 - p_{23}$. Also, departures from node 3 enter node 5 ($p_{35} = 1$), and departures from nodes 4 and 5 exit the network.

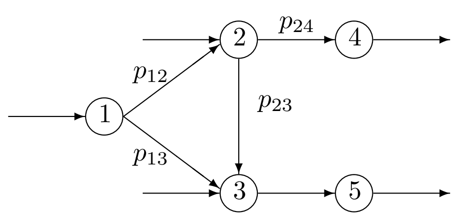

Figure 2. Acyclic Network

The times $T_1 < T_2 < \dots$ at which items enter the network from outside form a Poisson process with rate $\lambda \equiv \lambda_1 + \lambda_2 + \lambda_3$, since this process is the sum of the three independent Poisson processes flowing into nodes 1, 2 and 3. The probability that an arrival at any time T_n enters node i is λ_i/λ. This is the probability that the exponential time of an arrival at i is smaller than those exponential arrival times at the other nodes; see Exercise 1. The item that arrives at time T_n traverses a route $\mathbf{R}_n = (R_{n1}, \dots, R_{n\ell_n})$ in $\mathcal{R}$ (the set of all routes), and its sojourn times at the ℓ_n nodes on the route are $\mathbf{V}_n = (V_{n1}, \dots, V_{n\ell_n})$. The joint distribution of these marks $Y_n = (\mathbf{R}_n, \mathbf{V}_n)$

as functions of the network data λ_i, G_i and p_{jk} is

$$P\{\mathbf{R}_n = \mathbf{r}, \mathbf{V}_n \le \mathbf{v} | T_n\} = p(\mathbf{r}) \prod_{k=1}^{\ell} G_{r_k}(v_k),$$

where $p(\mathbf{r}) = (\lambda_{r_1}/\lambda) p_{r_1 r_2} \cdots p_{r_{\ell-1} r_\ell}$ is the probability of route $\mathbf{r} = (r_1, \dots, r_\ell)$.

To analyze the quantity of items on the routes as well as at the nodes, let us consider the space-time point process

$$N_t((a,b] \times B) = \#\text{ of items arriving in } (a,b] \text{ whose route and node location } (\mathbf{r}, i) \text{ is in } B \subset \mathcal{R} \times \{0,1,\dots,m\} \text{ at time } t.$$

As in (12.1), N_t is a Poisson process on $\Re_+ \times \mathcal{R} \times \{0,1,\dots,m\}$, for fixed t, where $g_t(T_n, Y_n) = (\mathbf{R}_n, f_t(T_n, Y_n))$ and $f_t(T_n, Y_n)$ is defined by (12.3). The item that enters at T_n is at node $f_t(T_n, Y_n)$ at time t. In particular, the quantity of items

$$Q_i(t) = \sum_{\mathbf{r} \in \mathcal{R}_i} N_t((a,b] \times \{\mathbf{r}\} \times \{i\})$$

at node i at time t has a Poisson distribution. Here $\mathcal{R}_i$ is the set of all routes that contain node i. Also, since N_t has independent increments, it follows that $Q_i(t)$ is independent of $Q_j(t)$ if $\mathcal{R}_i$ and $\mathcal{R}_j$ are disjoint. For instance, $Q_4(t)$ is independent of $Q_3(t)$ and $Q_5(t)$.

The next step is to evaluate the intensity of N_t. Let $P_{\mathbf{r}}^u(i)$ denote the conditional probability that an item is at node i given that is on route $\mathbf{r}$ for a time u since it entered the network. By the independence of the sojourn times,

$$P_{\mathbf{r}}^u(i) = \begin{cases} G_{r_1} \star \cdots \star G_{r_\ell}(u) & \text{if } i = 0 \\ G_{r_1} \star \cdots \star G_{r_{k-1}} \star \overline{G}_{r_k}(u) & \text{if } r_k = i \ne 0, \text{ for some } k \le \ell. \end{cases}$$

Here $\overline{G}(t) = 1 - G(t)$. For instance, the conditional probability that an item is at node 3 at time t, given that it enters route $\mathbf{r} = (1,2,3,5)$ at time s, is

$$\begin{aligned} G_1 \star G_2 \star \overline{G}_3(t-s) &= P\{\tau_{n2} \le t < \tau_{n3} | \mathbf{R}_n = \mathbf{r}, T_n = s\} \\ &= P\{f_t(T_n, Y_n) = 3 | \mathbf{R}_n = \mathbf{r}, T_n = s\}. \end{aligned}$$

Then from (12.2), it follows that

$$E[N_t((a,b] \times \{\mathbf{r}\} \times \{i\})] = \lambda p(\mathbf{r}) \int_a^b P_{\mathbf{r}}^{t-s}(i) ds.$$

The last integral equals $\int_a^b P_{\mathbf{r}}^u(i) du$, under the change-of-variable $u = t - s$.

In particular, the number of items arriving in $(0,t]$ that are on route $\mathbf{r} = (1,2,3,5)$ and in node 3 at time t has a Poisson distribution with mean

$$E[N_t((0,t] \times \{\mathbf{r}\} \times \{3\})] = \lambda_1 p_{12} p_{23} \int_0^t G_1 \star G_2 \star \overline{G}_3(u) du.$$

The process N_t yields considerable information about numbers of items at nodes and on routes as well. For instance, the quantity $Q_3(t)$ of items at node 3 at time t is the sum of the quantities of items on the routes in $\mathcal{R}_3 \equiv \{(3,5),(2,3,5),(1,3,5),(1,2,3,5)\}$, all the routes containing node 3. Then $Q_3(t)$ has a Poisson distribution and

$$\begin{aligned} E[Q_3(t)] &= \int_0^t \Big[\lambda_3 \overline{G}_3(u) + \lambda_2 p_{23} G_2 \star \overline{G}_3(u) + \lambda_1 p_{13} G_1 \star \overline{G}_3(u) \\ &\quad + \lambda_1 p_{12} p_{23} G_1 \star G_2 \star \overline{G}_3(u)\Big] du. \end{aligned}$$

The term in brackets is $\lambda \sum_{\mathbf{r} \in \mathcal{R}_3} p(\mathbf{r}) P^u_{\mathbf{r}}(3)$.

Similarly, the quantity of items on route $\mathbf{r}$ at time t is

$$Q_{\mathbf{r}}(t) = \sum_{k=1}^{\ell} Q_{r_k}(t) = N_t((0,t] \times \{\mathbf{r}\} \times \{r_1, \ldots, r_\ell\}).$$

This quantity, being part of the Poisson process N_t, has a Poisson distribution with $E[Q_{\mathbf{r}}(t)] = \sum_{k=1}^{\ell} E[Q_{r_k}(t)]$. For instance,

$$E[Q_{(2,4)}(t)] = \int_0^t \lambda_2 [\overline{G}_2(u) + G_2 \star \overline{G}_4(u)] du.$$

Let us now consider the departure times of the items from the nodes, which are depicted by the process

$$\begin{aligned} N((a,b] \times B) = \#\ &\text{of items arriving in } (a,b] \text{ whose departure times} \\ &\text{from the 5 nodes are in } B \subset \Re^5_+. \end{aligned}$$

The departure times are well-defined since an item cannot visit a node more than once. Now, N is a space-time Poisson process as in (12.1) and (12.2), where the departure times are given by

$$g(T_n, Y_n) = (g_1(T_n, Y_n), \ldots, g_5(T_n, Y_n)),$$

and $g_i(T_n, Y_n) = \sum_{k=1}^{\ell_n} \tau_{nk} \mathbf{1}(R_{nk} = i)$, the departure time from node i of the item that enters at T_n.

In particular, the departure process at node i is

$$D_i(t) = N((0,t] \times \{(t_1, \ldots, t_5) \in \Re^5_+ : t_i \le t\}), \qquad t \ge 0.$$

Now, D_i is a Poisson process since it is the projection on the ith departure-time coordinate of the Poisson process N. Its mean is

$$E[D_i(t)] = \lambda \int_0^t P\{g_i(T_n, Y_n) \le t | T_n = s\} ds. \tag{12.6}$$

For instance,

$$\begin{aligned} E[D_3(t)] &= \int_0^t \Big[\lambda_3 G_3(u) + \lambda_2 p_{23} G_2 \star G_3(u) + \lambda_1 p_{13} G_1 \star G_3(u) \\ &\quad + \lambda_1 p_{12} p_{23} G_1 \star G_2 \star G_3(u)\Big] du. \end{aligned}$$

Similarly to the independence of quantities at the nodes, processes D_i and D_j are independent if $\mathcal{R}_i$ and $\mathcal{R}_j$ are disjoint. For instance, D_4 is independent of D_3 and D_5.

13. Relatives of Poisson Processes

This section describes several variations of Poisson processes that arise naturally in applications.

EXAMPLE 13.1. *Cox Processes.* Loosely speaking, a Poisson process with a random intensity measure is a Cox Process. Such a process on $\Re_+$ has the following constructive definition. Let $N_1(t)$ be a homogeneous Poisson process on $\Re_+$ with rate 1, and let $\{\eta(t) : t \geq 0\}$ be an nondecreasing right-continuous real-valued stochastic process that is independent of N_1. Then the composition

$$N(t) = N_1(\eta(t)), \qquad t \geq 0.$$

is a Cox process on $\Re_+$ directed by η. The conditional distribution of $N(s, t] = N_1(\eta(s), \eta(t)]$ given η is

$$P\{N(s,t] = n|\eta\} = e^{-\eta(s,t]}\eta(s,t]^n/n!,$$

where $\eta(s,t] \equiv \eta(t) - \eta(s)$ is viewed as a random measure. Furthermore, the conditional distribution of N on disjoint subsets given η is that of a Poisson process. Equivalently, its conditional Laplace functional (obtained in Exercise 41) has the form (13.2) below, which is the Laplace functional for a Poisson process with intensity η (recall Proposition 6.3).

In general, a point process N on a space S is a *Cox process* directed by a locally-finite random measure η on S if N and η are defined on the same probability space and

$$E[e^{-Nf}|\eta] = \exp\{-\int_S (1 - e^{-f(x)})\eta(dx)\}, \qquad f \in C_K^+(S). \tag{13.2}$$

A Cox process is sometimes called a conditional Poisson process, a doubly stochastic Poisson process, or a Poisson process in a randomly changing environment. Several characterizations of Cox processes are established in [**21**]. Because Cox processes are essentially Poisson processes with an extra expectation, most results for Poisson processes have counterparts for Cox processes. For instance, if $N_1, \ldots, N_m$ are Cox processes on S directed by $\eta_1, \ldots, \eta_m$, respectively, and $(N_1, \eta_1), \ldots, (N_1, \eta_1)$ are independent, then $N = N_1 + \cdots N_m$ is a Cox process directed by $\eta = \eta_1 + \cdots \eta_m$.

Cox processes arise as a transformation of a Poisson process that has one more layer of randomness than those above. For instance, consider a Poisson process N on $\Re_+$ with intensity measure μ. First consider a transformation of N in which a point of N at t is mapped to a location $\gamma(t)$, where $\gamma(t)$ is a stochastic process on $\Re_+$ that is independent of N. Then the transformed process $N'(B) = N(\gamma^{-1}(B))$ is a Cox process directed by

$\eta(B) = \int_{\Re_+} \mathbf{1}(\gamma(t) \in B)\mu(dt)$, provided this is a.s. locally finite. This follows since by Theorem 8.1 N' is Poisson when $\gamma(t)$ is deterministic.

As a second example, suppose N is partitioned into m subprocesses $N_1, \dots, N_m$ by the rule that a point of N at t is assigned to subprocess $\alpha(t)$, where $\alpha(t)$ is a stochastic process on $\{1, \dots, \}$ that is independent of N. Then as in Corollary 10.2, $N_1, \dots, N_m$ are conditionally independent Poisson processes given $\alpha(\cdot)$. Hence each N_i is a Cox process directed by $\eta(B) = \int_{\Re_+} \mathbf{1}(\alpha(t) = i)\mu(dt)$.

EXAMPLE 13.3. *Markov-Modulated Poisson Process.* In computer and telecommunications systems, a useful model for the occurrences of an event in time is a Cox process N directed by $\eta(t) = \int_0^t Y(s)\, ds$, where $Y(t)$ is an ergodic Markov jump process on a countable state space I that models a changing environment in which events occur. For instance, a flow of data may be Poisson, but dependent on an environment (type or source of the data, congestion in a network, etc.) that is changing according to a Markov process.

Since the Cox process has the form $N(t) = N_1(\eta(t))$, its behavior far out in time is related to the limiting behavior of $Y(t)$. In particular,

$$t^{-1}N(t) \to \lambda \equiv \sum_{i\in I} ip_i, \qquad \text{a.s. as } t \to \infty,$$

where p_i is the stationary distribution of Y. This strong law of large numbers follows since by the strong laws of large numbers for Poisson and Markov processes, $N_1(t)/t \to 1$ and $\eta(t)/t \to \lambda$ a.s., and so

$$t^{-1}N(t) = (\eta(t)/t)N_1(\eta(t))/\eta(t) \to \lambda, \quad \text{a.s. as } t \to \infty.$$

Next, consider a variation of the $M_t/G/\infty$ system in which items arrive for service according to the preceding Cox process N, and G is the distribution of the independent service times. The system data is $M \equiv \sum_n \delta_{(T_n,V_n)}$ on $\Re_+^2$, where T_n is the arrival time of the nth item and V_n is its sojourn or service time. Analogously to Theorem 9.6, M is a space-time Cox process directed by η with $E[M([0,t) \times [0,v))|\eta] = \int_{(0,t]} G_s(v)Y(s)ds$.

Consider the quantity of items $Q(t) = \sum_n \mathbf{1}(T_n + V_n > t)$ in the system at time t. Arguing as in Example 11.2, the "conditional" distribution of $Q(t)$ given η is Poisson with

$$E[Q(t)|\eta] = \int_0^t [1 - G(t-s)]Y(s)ds.$$

Furthermore, under the additional assumptions that $Y(t)$ is stationary and G has a mean α, we have

$$\lim_{t\to\infty} P\{Q(t) = n\} = \sum_{i\in I} \pi_i (i\alpha)^n e^{-i\alpha}/n!. \tag{13.4}$$

This is a conditional Poisson distribution with random mean $\alpha Y(0)$.

To prove (13.4), note that

$$\begin{aligned} P\{Q(t)=n\} &= E\Big[(E[Q(t)|\eta])^n e^{-E[Q(t)|\eta]}/n!\Big], \\ E[Q(t)|\eta] &\stackrel{d}{=} Y(0)\int_0^t [1-G(u)]du \xrightarrow{d} \alpha Y(0). \end{aligned}$$

In light of these properties, (13.4) follows from the dominated convergence theorem for convergence in distribution (Theorem 8.10 in the Appendix).

Example 13.5. *Serial Marking of a Poisson Process.* The results above for a single marking of a Poisson process extend to a series of markings as follows. Starting with a Poisson process $N=\sum_n \delta_{X_n}$, if Y_n are p-marks of X_n, then $M=\sum_n \delta_{(X_n,Y_n)}$ is a Poisson process. Similarly, if Y_n' are p'-marks of (X_n,Y_n), then $M'=\sum_n (X_n,Y_n,Y_n')$ is again a Poisson process. These marking steps can be continued several times, with the end result being a Poisson process from which one can "read" off many results. In addition to serial markings in applications, they are useful for proving results for compound Poisson processes as discussed below.

Example 13.6. *Compound Poisson Process: Independent Increments.* Let $N(t)$ be a homogeneous Poisson process on $\Re_+$ with rate λ. Consider the stochastic process

$$Z(t)=\sum_{n=0}^{N(t)} Y_n, \qquad t\ge 0,$$

where Y_n are real-valued random variables that are i.i.d. with distribution F and are independent of N. As mentioned in Example 4.8, $\{Z(t): t\ge 0\}$ is a *compound Poisson process* with rate λ and distribution F.

The name comes from the fact that $Z(t)$ has a compound Poisson distribution with rate λt and distribution F:

$$P\{Z(t)\le z\}=\sum_{n=0}^{\infty} e^{-\lambda t}(\lambda t)^n F^{n\star}(z)/n!, \qquad z\in\Re. \tag{13.7}$$

This follows by conditioning $Z(t)$ on $N(t)$ and using $P\{Z(t)\le z|N(t)=n\}=F^{n\star}(z)$. In addition, observe that, by conditioning $Z(t)$ on $N(t)$,

$$E[Z(t)]=\lambda t E[Y_1], \qquad \mathrm{Var}Z(t)=\lambda t(\mathrm{Var}Y_1)^2,$$

provided these moments exist (conditioned on $N(t)$ the variance of $Z(t)$ is $N(t)\mathrm{Var}Y_1$). Here is more insight into the structure of the process $Z(t)$.

Theorem 13.8. *The process $\{Z(t): t\ge 0\}$ has stationary, independent increments: $Z(t_1)-Z(s_1),\dots,Z(t_n)-Z(s_n)$, for $s_1<t_1<\cdots s_n<t_n$, are independent, and $Z(s+t)-Z(s)\stackrel{d}{=}Z(t)$ for $s,t\ge 0$.*

Proof. Using the process $M=\sum_n \delta_{(T_n,Y_n)}$, we can write

$$Z(t)=\sum_n Y_n\mathbf{1}(T_n\le t)=\int_{\Re} yM((0,t]\times dy).$$

Under the assumptions, M is a space-time Poisson process with

$$E[M((s,s+t]\times B)] = \lambda t F(B).$$

Now, for any $s_1 < t_1 < \dots s_n < t_n$, consider the increments

$$Z(t_i) - Z(s_i) = \int_{\Re} yM((s_i,t_i]\times dy), \qquad 1 \le i \le n.$$

They are independent since the point processes $M((s_i,t_i]\times\cdot)$, for $1 \le i \le n$, on $\Re$ are independent, because M has independent increments.

Next, note that $E[M((s,s+t]\times B)] = E[M((0,t]\times B)]$. Since M is Poisson and its distribution is uniquely determined by its intensity, it follows that $M((s,s+t]\times\cdot)] \stackrel{d}{=} M((0,t]\times\cdot)$ Consequently,

$$Z(s+t) - Z(s) = \int_{\Re} yM((s,s+t]\times dy) \stackrel{d}{=} \int_{\Re} yM((0,t]\times dy) = Z(t).$$

Hence $Z(t)$ has stationary increments. □

Instead of the Y_n being independent of N suppose Y_n are p-marks of T_n. Then the process $Z(t) = \sum_{n=0}^{N(t)} Y_n$, for $t \ge 0$, is a *location-dependent compound Poisson process* with intensity measure μ and distribution $p(t,\cdot)$. Many of its properties follow directly from the fact that $M = \sum_n \delta_{(T_n,Y_n)}$ is a Poisson process. For instance, see Exercises 44 and 45. Also, since M is Poisson, the results above for Poisson processes extend to compound Poisson processes by using a p-marking of M, which would be a "second" marking of N as mentioned in Example 13.5. Exercise 44 illustrates this ideas for partitions of compound Poisson processes.

There are other relatives of compound Poisson processes of the form $M(A) = \sum_n Y_n \delta_{X_n}(A)$, where $N = \sum_n \delta_{X_n}$ is a Poisson process on a general space, and the marks Y_n are random vectors, matrices, or elements of a group with an addition operation. Here is an example when Y_n are point processes.

EXAMPLE 13.9. *Poisson Cluster Processes.* Let $N = \sum_n \delta_{X_n}$ denote a Poisson process on a general space S with intensity measure μ. Suppose that each point X_n generates a cluster of points in a space S' that are represented by a point process N'_n. Assume N'_n are point processes on a space S' that are i.i.d. and independent of N. Then the number of points from the processes N'_n in a set B that are generated by points of N in the set A is

$$M(A\times B) = \sum_n N'_n(B)\delta_{X_n}(A).$$

This defines a point process M on $S \times S'$ called a *marked cluster process* generated by the Poisson process N; the $M(S\times\cdot)$, provided it is locally finite, is simply the cluster process on S'.

Since $M(A \times B) = \sum_{n=0}^{N(A)} N'_n(B)$, it follows by conditioning on N that $M(A \times B)$ has the compound Poisson distribution

$$P\{M(A \times B) \le n\} = \sum_{k=0}^{\infty} e^{-\mu(A)} \mu(A)^k F^{k\star}(n; B)/k!,$$

where $F(n; B) = P\{N'_1(B) \le n\}$. Also, $E[M(A \times B)] = E[N(A)]E[N'_1(B)]$ and

$$\mathrm{Var} M(A \times B) = E[N(A)](\mathrm{Var} N'_1(B))^2.$$

More general cluster processes, where the N'_n are marks of X_n, are analyzed in Exercise 47.

14. Poisson Law of Rare Events

Poisson processes are natural models for rare events in time, or rare points in a space. This is partly due to central-limit phenomena in which sums of thin or rarefied point processes converge in distribution to Poisson processes. A classical case for random variables is a Binomial random variable converging to a Poisson random variable as in Example 14.1 below. In this section, we present a generalization of this result that gives conditions under which a sum of many rare indicator random variables converges to a Poisson random variable. Analogous results under which a sum of point processes converges to a Poisson process are in the next section.

Here is a classical example of the Poisson law of rare events below.

Example 14.1. *Binomial Convergence to Poisson.* Suppose $Y_{n1}, \dots, Y_{nn}$ are independent Bernoulli random variables with $P\{Y_{ni} = 1\} = p_n$. Then $Z_n \equiv \sum_{i=0}^{n} Y_{ni}$ has a binomial distribution with parameters n and p_n. If $np_n \to \mu > 0$ as $n \to \infty$, then $Z_n \xrightarrow{d} Z$, where Z has a Poisson distribution with mean μ. This is a special case of the following result.

Theorem 14.2. (Poisson Law of Rare Events) *Suppose* $Y_{n1}, Y_{n2}, \dots$, *for* $n \ge 1$, *are a countable number of independent random variables that take values* 0 *or* 1, *and satisfy the* uniformly null *property*

$$\sup_i P\{Y_{ni} = 1\} \to 0, \quad \textit{as } n \to \infty. \tag{14.3}$$

Let Z be a Poisson random variable with mean μ. Then as $n \to \infty$,

$$Z_n \equiv \textstyle\sum_i Y_{ni} \xrightarrow{d} Z \quad \textit{if and only if} \quad \textstyle\sum_i P\{Y_{ni} = 1\} \to \mu.$$

Proof. We will use the property of Laplace transforms that $Z_n \xrightarrow{d} Z$ if and only if $E[e^{-\alpha Z_n}] \to E[e^{-\alpha Z}]$. By the independence of the Y_{ni},

$$E[e^{-\alpha Z_n}] = \prod_i E[e^{-\alpha Y_{ni}}] = \prod_i (1 - c_{ni}), \qquad \alpha \ge 0,$$

where $c_{ni} = E[1-e^{-\alpha Y_{ni}}]$. Also, $E[e^{-\alpha Z}] = e^{-c}$, where $c = \mu(1-e^{-\alpha})$, since Z has a Poisson distribution with mean μ. From these observations,

$$Z_n \xrightarrow{d} Z \;\Leftrightarrow\; E[e^{-\alpha Z_n}] \to e^{-c} \Leftrightarrow \prod_i (1-c_{ni}) \to e^{-c}. \tag{14.4}$$

Moreover, under the assumption (14.3) and $c_{ni} \le (1-e^{-\alpha}) < 1$, it follows by Lemma 14.5 for $c_{ni} = c\mu^{-1}P\{Y_{ni}=1\}$, that

$$\prod_i (1-c_{ni}) \to e^{-c} \;\Leftrightarrow\; \sum_i c_{ni} \to c \;\Leftrightarrow\; \sum_i P\{Y_{ni}=1\} \to \mu.$$

Combining this string of equivalences with (14.4) proves the assertion. □

The preceding proof of the Poisson convergence boils down to the following result on the convergence of real numbers.

Lemma 14.5. *Suppose $c_{n1}, c_{n2}, \ldots$, for $n \ge 1$, are a countable (possibly finite) number of real numbers in $(0,a]$, where $a < 1$, that satisfy the* uniformly null *property*

$$\sup_i c_{ni} \to 0 \qquad \text{as } n \to \infty. \tag{14.6}$$

Then, for any $c > 0$,

$$\lim_{n\to\infty} \prod_i (1-c_{ni}) = e^{-c} \quad \textit{if and only if} \quad \lim_{n\to\infty} \sum_i c_{ni} = c.$$

Proof. The assertion is equivalent to

$$\bar{s}_n \equiv -\sum_i \log(1-c_{ni}) \to c \quad \text{if and only if} \quad s_n \equiv \sum_i c_{ni} \to c. \tag{14.7}$$

Since $\log(1-c_{ni}) = -\sum_{m=1}^{\infty} c_{ni}^m/m$, the difference in these sums is

$$d_n \equiv \bar{s}_n - s_n = \sum_i c_{ni}^2 \sum_{m=2}^{\infty} c_{ni}^{m-2}/m.$$

Using $c_{ni} \le a$ and $\alpha_n \equiv \sup_i c_{ni}$, we have

$$d_n \le \frac{1}{1-a}\sum_i c_{ni}^2 \le \frac{\alpha_n s_n}{1-a} \le \frac{\alpha_n \bar{s}_n}{1-a}. \tag{14.8}$$

Now, if $\bar{s}_n \to c$, then from (14.8) and (14.6) we have $d_n \to 0$, and hence $s_n = \bar{s}_n - d_n \to c$. Similarly, if $s_n \to c$, then $\bar{s}_n = s_n + d_n \to c$. These observations prove (14.7). □

15. Poisson Convergence Theorems*

This section contains Poisson convergence theorems for sequences of point processes. These results are extensions of the Poisson law of rare events in Theorem 14.2 above. The main theorem is that a sum of many independent sparse point processes converges to a Poisson process. Consequently, certain sums of renewal processes and rare transformations of a point process converge to a Poisson process. Also included are examples justifying that Poisson processes are reasonable approximations for thinnings and partitions of a point process.

We will use the following notion of weak convergence, which is reviewed in Section 8 of the Appendix. Suppose $\mu, \mu_1, \mu_2, \dots$ are probability measures on S. The probabilities μ_n *converge weakly* to μ as $n \to \infty$, denoted by $\mu_n \xrightarrow{w} \mu$, if $\mu_n f \to \mu f$, as $n \to \infty$, for each bounded continuous function $f : S \to \Re$ (recall $\mu f = \int_S f(x)\mu(dx)$). This is equivalent to

$$\lim_{n\to\infty} \mu_n(B) = \mu(B), \qquad B \in \hat{\mathcal{S}}_\mu, \tag{15.1}$$

where $\hat{\mathcal{S}}_\mu \equiv \{B \in \hat{\mathcal{S}} : \mu(\partial B) = 0\}$, the set of all bounded sets whose boundary has μ-measure 0.

A sequence of random elements in a metric space converges in distribution to a random element if their distributions converge weakly. In particular, a sequence of point processes N_n on S *converges in distribution to* N as $n \to \infty$, denoted by $N_n \xrightarrow{d} N$, if $P\{N_n \in \cdot\} \xrightarrow{w} P\{N \in \cdot\}$. This weak convergence is equivalent to the convergence of the finite-dimensional distributions (condition (ii) in the next theorem).

A few points in our analysis use the slightly more general notion of vague convergence of measures. Suppose $\mu, \mu_1, \mu_2, \dots$ are locally finite measures on S. The measures μ_n *converge vaguely* to μ, denoted by $\mu_n \xrightarrow{v} \mu$, if

$$\mu_n f \to \mu f, \quad \text{as } n \to \infty, \text{ for each } f \in C_K^+(S).$$

This is equivalent to (15.1), and vague convergence is the same as weak convergence when all the measures are probability measures.

The following are several equivalent conditions for point processes to converge in distribution. Here $\hat{\mathcal{S}}_N = \{B \in \hat{\mathcal{S}} : N(\partial B) = 0 \text{ a.s. }\}$.

THEOREM 15.2. *For point processes $N, N_1, N_2, \dots$ on S, the following statements are equivalent as $n \to \infty$.*

(i) $N_n \xrightarrow{d} N$.

(ii) $(N_n(B_1), \dots, N_n(B_k)) \xrightarrow{d} (N(B_1), \dots, N(B_k)), \quad B_1, \dots, B_k \in \hat{\mathcal{S}}_N$.

(iii) $N_n f \xrightarrow{d} N f, \qquad f \in C_K^+(S)$.

(iv) $E[e^{-N_n f}] \to E[e^{-N f}], \qquad f \in C_K^+(S)$.

A proof of this result is in [**20**]. Condition (ii) says the finite-dimensional distributions of N_n converge to those of N. When S is an Euclidean space, the sets B_i can be replaced by bounded rectangles. Condition (iii) relates the convergence in distribution of integrals with respect to point processes to the convergence of the processes. The convergence (iv) of Laplace functionals is a convenient tool for proving $N_n \xrightarrow{d} N$, when the functionals can be factored conveniently (as in the proof of Theorem 15.8 below).

Here is an elementary but useful fact. It justifies the convergence of a Poisson process when its intensity converges; e.g., see Exercise 49.

PROPOSITION 15.3. (Convergence of Poisson Processes) *For each* $n \geq 1$, *suppose* N_n *is a Poisson process on a space* S *with intensity measure* μ_n. *If* $\mu_n \xrightarrow{v} \mu$ *and* μ *is locally finite, then* $N_n \xrightarrow{d} N$, *where* N *is a Poisson process with intensity* μ.

PROOF. From Proposition 6.3, we know that $E[e^{-N_n f}] = e^{-\mu_n h}$, for $f \in C_K^+(S)$, where $h(x) \equiv 1 - e^{-f(x)}$. By Theorem 15.2, we have $\mu_n h \xrightarrow{v} \mu h$, and so

$$E[e^{-N_n f}] = e^{-\mu_n h} \to e^{-\mu h} = E[e^{-Nf}].$$

Thus, $N_n \xrightarrow{d} N$ by Theorem 15.2. □

We are now ready to consider the convergence of non-Poisson processes. We begin with a motivating example.

EXAMPLE 15.4. Consider a sum $N(t) = \sum_{i=1}^n N_i(t)$, for $t \geq 0$, where $N_1, \ldots, N_n$ are independent renewal processes. Of course, N is generally not a renewal process. However, suppose the times between renewals for each process N_i tend to be large (i.e., $F_i(t)$ is small, where F_i is the inter-renewal distribution). Consequently, the contribution $N_i(a, b]$ to $N(a, b]$ would tend to be 0. In other words, each N_i rarely contributes a point to N on bounded intervals. However, if the number n of these contributions is large, it might be reasonable to approximate N by a Poisson process with intensity $E[N(t)] = \sum_{i=1}^n E[N_i(t)]$. A Poisson convergence theorem justifying such an approximation is as follows. The opposite situation in which $N_i(a, b]$ tends to be large is addressed in Exercise 52.

THEOREM 15.5. (Sums of Renewal Processes) *Suppose, for* $n \geq 1$, *that* $N_n(t) = \sum_i N_{ni}(t)$ *is a point process on* $\Re_+$, *where* N_{ni} *is a countable set of independent renewal processes with inter-renewal distributions* F_{ni}. *Assume the inter-renewal times are uniformly rare in that*

$$\lim_{n\to\infty} \sup_i F_{ni}(t) = 0, \qquad t \geq 0.$$

Let N *be a Poisson process on* $\Re_+$ *with intensity measure* μ. *Then* $N_n \xrightarrow{d} N$, *as* $n \to \infty$ *if and only if, for each* t *with* $\mu(\{t\}) = 0$,

$$\lim_{n\to\infty} \sum_i F_{ni}(t) = \mu(t). \tag{15.6}$$

PROOF. This result follows by Theorem 15.8 below, since

$$\begin{aligned}
\sum_i P\{N_{ni}(t) \geq 2\} &= \sum_i \int_{(0,t]} F_{ni}(t-s) F_{ni}(ds) \\
&\leq \sup_i F_{ni}(t) \sum_i F_{ni}(t),
\end{aligned}$$

and (15.6) is the same as (15.10) because $P\{N_{ni}(t) \geq 1\} = F_{ni}(t)$. □

The next result is a general Poisson convergence theorem for sums of uniformly rare point processes. Suppose that

$$N_n \equiv \sum_i N_{ni}, \qquad n \geq 1,$$

is a point process on a space S, where $N_{n1}, N_{n2}, \ldots$ is a countable number of independent point processes on S. Assume the point processes N_{ni} are *uniformly null*, meaning that

$$\lim_{n\to\infty} \sup_i P\{N_{ni}(B) \geq 1\} = 0, \qquad B \in \hat{\mathcal{S}}. \tag{15.7}$$

Let N be a Poisson process on S with intensity measure μ.

THEOREM 15.8. (Grigelionis) *For the processes defined above, $N_n \xrightarrow{d} N$, as $n \to \infty$ if and only if*

$$\lim_{n\to\infty} \sum_i P\{N_{ni}(B) \geq 2\} = 0, \qquad B \in \hat{\mathcal{S}}, \tag{15.9}$$

$$\lim_{n\to\infty} \sum_i P\{N_{ni}(B) \geq 1\} = \mu(B), \qquad B \in \hat{\mathcal{S}}_\mu. \tag{15.10}$$

PROOF. The convergence $N_n \xrightarrow{d} N$ is equivalent, by Theorem 15.2, to

$$E[e^{-N_n f}] \to E[e^{-Nf}], \qquad f \in C_K^+(S). \tag{15.11}$$

Using the independence of the N_{ni} and letting $c_{ni} \equiv E[1 - e^{-N_{ni}f}]$, we have

$$E[e^{-N_n f}] = \prod_i E[e^{-N_{ni}f}] = \prod_i (1 - c_{ni}).$$

Also, by Proposition 6.3, $E[e^{-Nf}] = e^{-\mu h}$, where $h(x) \equiv 1 - e^{-f(x)}$. Combining these observations, it follows that (15.11) is equivalent to

$$\prod_i (1 - c_{ni}) \to e^{-\mu h}, \qquad f \in C_K^+(S). \tag{15.12}$$

Keep in mind that c_{ni} is a function of f.

We will complete the proof by applying Lemma 14.5 to establish that (15.9) and (15.10) are necessary and sufficient for (15.12). Clearly c_{ni} are in $(0, a]$, where $a = 1 - \exp\{-\max_{x\in S} f(x)\}$. Next, note that

$$c_{ni} = E[1 - e^{-N_{ni}f}] \leq P\{N_{ni}(S_f) \geq 1\},$$

where S_f is the support of f. Then (15.7) implies

$$\sup_i c_{ni} \leq \sup_i P\{N_{ni}(S_f) \geq 1\} \to 0.$$

In light of this property, Lemma 14.5 says that (15.12) is equivalent to

$$\sum_i E[1 - e^{-N_{ni}f}] = \sum_i c_{ni} \to \mu h, \qquad f \in C_K^+(S). \tag{15.13}$$

Therefore, it remains to show that (15.9) and (15.10) are necessary and sufficient for (15.13).

We can write

$$\text{(15.14)}\quad \sum_i c_{ni} = \sum_i E[(1-e^{-N_{ni}f})\mathbf{1}(N_{ni}(S_f)=1)] + \sum_i E[(1-e^{-N_{ni}f})\mathbf{1}(N_{ni}(S_f)\geq 2)].$$

The last sum is bounded by $\sum_i P\{N_{ni}(S_f)\geq 2)\}$ which converges to 0 by assumption (15.9). The first sum on the right-hand side in (15.14) equals $\eta_n h$, where

$$\eta_n(B) \equiv \sum_i E[N_{ni}(B\cap S_f)\mathbf{1}(N_{ni}(S_f)=1)] = \sum_i P\{N_{ni}(B\cap S_f)=1\}.$$

The last sum has the same form as the sum in (15.9) minus the one in (15.10), and so these assumptions imply $\eta_n \xrightarrow{v} \mu$, which yields $\eta_n h \to \mu h$. Using the preceding observations in (15.14) proves that (15.9) and (15.10) are sufficient for (15.13).

Conversely, suppose (15.13) is true. This property for the function $f(x)\equiv -\mathbf{1}(x\in B)\log s$, where $B\in\hat{\mathcal{S}}$ and $s\in[0,1]$, says

$$\text{(15.15)}\qquad H_n(s)\equiv \sum_i E[1-s^{N_{ni}(B)}] \to (1-s)\mu(B),$$

since $h(x)=1-e^{-f(x)}=(1-s)\mathbf{1}(x\in B)$. Then (15.10) follows since

$$\text{(15.16)}\qquad \sum_i P\{N_{ni}(B)\geq 1\} = H_n(0)\to\mu(B).$$

Next, consider the factorization

$$\begin{aligned} H_n(s) &= \sum_i [1-\sum_{m=0}^{\infty} s^m P\{N_{ni}(B)=m\}] \\ &= (1-s)H_n(0) + \sum_i\sum_{m=2}^{\infty}(s-s^m)P\{N_{ni}(B)=m\}. \end{aligned}$$

Then using this expression along with (15.15) and (15.16), we have

$$(s-s^2)\sum_i P\{N_{ni}(B)\geq 2\} \leq H_n(s)-(1-s)H_n(0)\to 0.$$

Thus (15.13) is true. These observations prove that (15.13) implies (15.9) and (15.10), which completes the proof. □

Theorem 15.8 justifies Poisson limits for sums of independent renewal processes (Theorem 15.5 and Exercise 52). Although Theorem 15.8 is for sums of independent point processes, it also applies to certain sums of conditionally independent point processes, including those associated with random transformations and thinning of point processes, which we now discuss.

EXAMPLE 15.17. *A Thinned Process.* Let N be a point process on $\Re_+$ (e.g., a renewal process) that satisfies $t^{-1}N(t) \xrightarrow{d} \lambda$ as $t \to \infty$, where λ is a positive constant. Suppose each point of N is independently retained with probability p and deleted with probability $1-p$. Let $N_p(t)$ denote the number of retained points in $(0,t]$. When p is very small, the retained points are rare and so it appears that it would be appropriate to approximate the p-thinning N_p of N by a Poisson process. Based on Corollary 15.24 below, it is reasonable to approximate N_p by a Poisson process with rate $p\lambda$ when p is small.

EXAMPLE 15.18. *Poisson Limit of Rare Transformations.* Consider a sequence of point processes $N_n = \sum_j \delta_{X_{nj}}$ on a space S with intensity measures μ_n. Let M_n be a marked p_n-transformation of N_n on $S \times S'$.

A natural prerequisite for M_n to converge is that the transformations should be uniformly null. Accordingly, we will use the condition

$$\lim_{n\to\infty} \sup_{x\in A} p_n(x,B) = 0, \quad A \in \hat{\mathcal{S}},\ B \in \hat{\mathcal{S}}'. \tag{15.19}$$

Next, observe that the conditional mean measure of M_n given N_n is

$$\begin{aligned} \eta_n(A\times B) &\equiv E[M_n(A\times B)|N_n] = \sum_j p_n(X_{nj},B)\mathbf{1}(X_{nj}\in A) \\ &= \int_A p_n(x,B)N_n(dx), \quad A\in\mathcal{S},\ B\in\mathcal{S}'. \end{aligned}$$

The convergence in distribution of these random mean measures η_n is another prerequisite for M_n to converge. For such random measures, the convergence $\eta_n \xrightarrow{d} \eta$ is analogous to convergence in distribution of point processes, and equivalent statements for this are given in Theorem 15.2 (with η in place of N).

THEOREM 15.20. *Suppose the sequence M_n of marked p_n-transformations of N_n satisfies (15.19). Also, assume $\eta_n \xrightarrow{d} \mu$ as $n \to \infty$, where μ is a (non-random) locally finite measure on $S\times S'$. Then $M_n \xrightarrow{d} M$ as $n\to\infty$, where M is a Poisson process on $S\times S'$ with intensity measure μ.*

PROOF. We can write $M_n = \sum_i M_{ni}$, where $M_{ni} \equiv \delta_{X_{ni},Y_{ni}}$ and Y_{ni} are p_n-marks of the X_{ni}, for $i \geq 1$. Although the point processes M_{ni}, $i\geq 1$, are not independent, they are conditionally independent given N_n. Clearly $P\{M_{ni}(B)\geq 2|N_n\} = 0$ and, under assumption (15.19),

$$\sup_i P\{M_{ni}(A\times B)\geq 1|N_n\} \leq \sup_{x\in A} p_n(x,B) \to 0, \quad A\in\hat{\mathcal{S}},\ B\in\hat{\mathcal{S}}'.$$

Also, $\eta_n \xrightarrow{d} \mu$ implies, for $A\times B \in \widehat{\mathcal{S}\times\mathcal{S}'}_\mu$,

$$\begin{aligned} \sum_i P\{M_{ni}(A\times B)\geq 1|N_n\} &= E[M_n(A\times B)|N_n] \\ &= \eta_n(A\times B) \xrightarrow{d} \mu(B). \end{aligned}$$

Applying Theorem 15.8 to the conditional distribution of M_n given N_n, and using Theorem 15.2, it follows that

$$E[e^{-M_n f}|N_n] \xrightarrow{d} E[e^{-Mf}], \qquad f \in C_K^+(S).$$

Taking expectations of this and using the dominated convergence theorem for convergence in distribution (Theorem 8.10 in the Appendix), we have $E[e^{-M_n f}] \to E[e^{-Mf}]$. Thus, $M_n \xrightarrow{d} M$ by Theorem 15.2. □

EXAMPLE 15.21. *Poisson Limits of Partitions.* Let $N(t)$ be a point process on $\Re_+$. Suppose N is partitioned as in Corollary 10.2 by the following rule: Each point of N is assigned to subprocess $i \in I$ (a countable set) with probability $p(i)$, independent of everything else, where $\sum_{i\in I} p(i) = 1$. Then $N = \sum_{i\in I} N_i$, where N_i denotes the ith subprocesses in the partition. We address the issue of finding conditions under which the subprocesses $\{N_i : i \in I_0\}$, for a subset $I_0 \subset I$, are approximately independent Poisson processes. The thinning in Example 15.17 is such a partition consisting of two subprocesses, where $I_0 = \{0\}$ and $I = \{0, 1\}$.

To justify an approximation of the subprocesses by a limit theorem, assume the partitioning probabilities are functions of n such that $p_n(i) \to 0$, for $i \in I_0$ (I is necessarily infinite when $I_0 = I$). Denote the ith subprocess by $N_n^i(t)$. Its conditional mean given N is $E[N_n^i(t)|N] = p_n(i)N(t)$. This mean converges to 0, which would not lead to a non-zero limit of N_n^i.

To obtain a non-zero limit, a normalization of the processes N_n^i is in order. Accordingly, assume there is a positive constant λ such that

$$t^{-1}N(t) \xrightarrow{d} \lambda. \tag{15.22}$$

This ensures that $N(t) \to \infty$ and that the points of N appear at a positive rate out to infinity. Next, assume the partitioning is uniformly rare on I_0: there exist positive constants $a_n \to \infty$ and $r(i)$, such that

$$\lim_{n\to\infty} a_n p_n(i) = r(i), \qquad i \in I_0. \tag{15.23}$$

Under the preceding assumptions, it is natural to consider the convergence of the processes

$$N_{ni}(t) \equiv N_n^i(a_n t), \qquad i \in I_0.$$

These are normalizations of the processes N_n^i under a rescaling of time so that a_n is the new unit of time. The N_{ni} on a "fixed" interval $(0, t]$ represents subprocess i on the interval $(0, a_n t]$, which becomes larger as $n \to \infty$. The choice of a_n for the time unit is because, as $n \to \infty$,

$$E[N_{ni}(t)|N] = a_n p_n(i)(N(a_n t)/a_n) \xrightarrow{d} r(i)\lambda, \qquad i \in I_0.$$

The following result describes the Poisson limits of the subprocesses. Interestingly, the processes N_{ni} for $i \in I_0$ are dependent for each n but in the limit they are independent.

COROLLARY 15.24. *Under assumptions (15.22) and (15.23),*

$$(N_{ni} : i \in I_0) \xrightarrow{d} (N_i : i \in I_0), \quad \text{as } n \to \infty, \tag{15.25}$$

where the limiting processes are independent homogeneous Poisson processes with respective rates $r(i)\lambda$, $i \in I_0$.

PROOF. The partition of N we are studying is a special p_n-transformation of the process $N(a_n\cdot)$ with $p_n(t,B) = \sum_{i\in B} p_n(i)$. Specifically, the number of the $N(a_n t)$ points assigned to subprocess $i \in I_0$ is

$$N_{ni}(t) = M_n((0,t] \times \{i\}),$$

where M_n is a marked p_n-transformation on $\Re_+ \times I_0$ of $N(a_n\cdot)$ as in Example 15.18. For each $t \geq 0$ and $B \subset I_0$,

$$\sup_{s\leq t} p_n(s,B) = \sum_{i\in B} p_n(i) \to 0.$$

Furthermore, under assumptions (15.22) and (15.23),

$$\begin{aligned} \eta_n((0,t]\times B) \quad &\equiv \quad E[M_n((0,t]\times B)|N(a_n\cdot)] \;=\; a_n \sum_{i\in B} p_n(i)(N(a_n t)/a_n) \\ &\xrightarrow{d} \quad \eta((0,t]\times B) \;\equiv\; \sum_{i\in B} r(i)\lambda t. \end{aligned}$$

Thus, the assumptions of Theorem 15.20 are satisfied, and so $M_n \xrightarrow{d} M$, where M is a Poisson process with $E[M((0,t]\times B)] = \sum_{i\in B} r(i)\lambda t$. Hence, assertion (15.25) follows since $N_{ni}(t) = M_n((0,t]\times\{i\})$. □

16. Exercises

1. Let $X_1, \ldots, X_m$ be independent exponentially distributed random variables with respective rates $\lambda_1, \ldots, \lambda_m$. Define $Y = \min\{X_1, \ldots, X_m\}$ and $\lambda = \sum_{i=1}^m \lambda_i$. Show that

$$P\{X_j = Y\} = \lambda_j/\lambda, \qquad P\{Y > t\} = e^{-\lambda t},$$
$$P\{X_j = Y, Y > t\} = P\{X_j = Y\}P\{Y > t\}.$$

Also, find simple expressions for

$$P\{X_1 < X_2\}, \quad P\{\max\{X_1, \ldots, X_m\} \leq x\}.$$

2. A space station requires the continual use of two systems whose lifetimes are independent exponentially distributed random variables X_1 and X_2 with respective rates λ_1 and λ_2. In addition, when system 2 fails, it is replaced by a spare system whose lifetime X_3 is exponentially distributed with rate λ_3, independent of the other systems. Find the distribution of the time $Y = \min\{X_1, X_2 + X_3\}$ at which one of the systems becomes inoperative. Find the probability that system 1 will fail before system 2 (with its spare) fails.

3. *Dispatching.* In Example 1.1, what properties of the Poisson process are not needed to obtain the optimal dispatching policy? What is the optimal policy when the arrival process $N(t)$ is a simple, stationary point process with $N(t) = \lambda t$, such as a stationary renewal process?

4. *Waiting to be Dispatched.* Items arrive to a dispatching station according to a Poisson process with rate λ, and all items in the system will be dispatched at a time t. Example 1.1 shows that the expected time items wait before being dispatched at time t is $E[\int_0^t N(s)ds] = \lambda^2/2$. Suppose there is a cost hw^2 for holding an item in the system for a time w. Then the total holding cost in $(0, t]$ is $C = \sum_{n \ge 1} h(t - T_n)^2$. Find $E[\alpha C]$ and $E[C]$.

5. Requests for a product arrive to a storage facility according to a Poisson process with rate λ per hour. Given that n requests are made in a t-hour time interval, find the probability that at least k requests were made in the first hour. Is this conditional probability different if the beginning of the one-hour period is chosen according to a probability density $f(s)$ on the interval $[0, t-1]$?

6. From Theorem 6.4, we know that the sum $N = N_1 + \cdots N_n$ of independent Poisson processes is Poisson. Prove this statement by verifying that N satisfies the defining properties of a Poisson process.

7. Prove the following statement. A simple point process N on $\Re_+$ is a Poisson process with rate λ if and only if N has independent increments, $EN(1) = \lambda$, and, for any t and n, the conditional joint density of $T_1, \ldots, T_n$ given $N(t) = n$ is the same as the density of the order statistics of n independent uniformly distributed random variables on $[0, t]$.

8. *Shot Noise Process.* Suppose that shocks (or pulses) to a system occur at times that form a Poisson process with rate λ. The shock at time T_n has a magnitude Y_n and this decays exponentially over time with rate γ. Then the cumulative effect of the shocks at time t is

$$Z(t) = \sum_{k=1}^{N(t)} Y_n e^{-\gamma(t-T_n)}.$$

Assume $Y_1, Y_2, \ldots$ are i.i.d. with mean μ and variance σ^2, and are independent of N. Find expressions for the mean and variance of $Z(t)$ in terms of μ and σ^2.

9. *Randomly Discounted Cash Flows.* In the context of Example 4.7, consider the generalization

$$Z(t) = \sum_{n=1}^{N(t)} Y_n e^{-\gamma_n T_n},$$

where $\gamma_1, \gamma_2, \ldots$ are independent nonnegative discount rates with distribution G that are independent of N and the Y_n's. Show that

$$E[Z(t)] = \lambda EY_1 \int_0^t \int_{\Re_+} e^{-\gamma x} G(d\gamma)dx.$$

10. Calls arrive to an operator at a call center at times that form a Poisson process $N(t)$ with rate λ. The time τ devoted to a typical call has an exponential distribution with rate μ, and it is independent of N. Then $N(\tau)$ is the number of calls that arrive while the operator is busy answering a call. Find the Laplace transform and mean and variance of $N(\tau)$. Find $P\{N(\tau) \le 1\}$.

11. For a Poisson process N with rate λ, show that

$$E[T_\ell - T_k | N(t) = n\} = (\ell - k)/(n+1), \qquad k < \ell \le n, \ t > 0.$$

Find an expression for $E[t - T_k | N(t) = n\}$.

12. Consider a set of N jobs that are assigned to m workers for processing. Each job is randomly assigned to worker i with probability p_i, for $i = 1, \ldots, m$. Let N_i denote the number of items assigned to category i, so that $N = N_1 + \cdots + N_m$. Suppose N has a Poisson distribution with mean λ. Describe the joint distribution of $N_1, \ldots, N_m$.

13. Patients at an emergency room are categorized into m types. Assume the arrivals of the m types of patients occur at times that form independent homogeneous Poisson processes with respective rates $\lambda_1, \ldots, \lambda_m$.
(a) Find the probability that a type 1 patient arrives before a type 2 patient.
(b) What is the probability that the next patient to arrive after a specified time is of type 1?
(c) Find the probability that in the next 5 arrivals, there are exactly 3 type 1 patients.
(d) Find the probability that 3 type 1 patients arrive before the first type 2 patient.
(e) Find the probability that the next patient to arrive is of type 1, 2 or 3.

14. *Dynamic Servicing.* Customers randomly request service at a manufacturing facility during an eight-hour day according to a Poisson process with intensity λ_t. The requested orders are satisfied as soon as possible, but may be delayed due to machine workloads, worker schedules, machine availability, etc. Past history shows that a request at time t will be satisfied either: (1) That day. (2) The next day. (3) Some time later. The request at time t is satisfied under scenario i with probability $p_i(t)$, and the expected revenue for such an order is r_i, where $i = 1, 2, 3$.
(a) Find the distribution of the number of requests in a day that are satisfied under each scenario i, where $i = 1, 2, 3$.
(b) Find the daily expected revenue for satisfying the customers, and find

the variance of this revenue.
(This is an actual model of customer requests for orders of paper labels produced by a company.)

15. Requests for a product (information or service) arrive from m cities at times that form independent Poisson processes with rates $\lambda_1, \dots, \lambda_m$. Given that there are n requests from the cities in the time interval $(0, t]$, find the conditional probability that n_1 are from city 1 and n_2 are from city 2.

16. At the end of a production shift, it is anticipated that there will be N jobs left to be processed, where N has a Poisson distribution with mean μ. Suppose the jobs are processed in parallel and the times to complete them are independent with a distribution G. Let $Q(t)$ denote the number of jobs in the system at time t, and let $D(t)$ denote the number of jobs completed in $(0, t]$. Find the distributions of $Q(t)$ and $D(t)$. Is D a Poisson process?

Answer this question under an alternative scenario in which the jobs are processed serially (one at a time) and G is an exponential distribution with rate λ.

17. Let $X_{(1)} \le \cdots \le X_{(n)}$ denote the order statistics from a random sample of size n from an exponential distribution with rate λ. Consider the distances between points $D_1 = X_{(1)}$, and $D_k = X_{(k)} - X_{(k-1)}$, $2 \le k \le n$. Show that these distances are independent, and that D_k has an exponential distribution with rate $(n - k + 1)\lambda$.

18. Let $X_{(1)} \le \cdots \le X_{(n)}$ denote the order statistics of a random sample from a continuous distribution $F(x)$ with density $f(x)$. Show that the distribution and density of $X_{(k)}$ are

$$P\{X_{(k)} \le x\} = \sum_{j=1}^{k} \binom{n}{j} F(x)^j (1 - F(x))^{n-j},$$

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} F(x)^{k-1} f(x) (1 - F(x))^{n-k}, \qquad x \in \Re.$$

19. *Point Locations for Non-Homogeneous Poisson Processes.* Consider a Poisson process N on R_+ with rate function $r(x)$, where $r(x) > 0$ for each x. Prove the following *Order Statistic Property*:
The conditional density of $T_1, \dots, T_n$ given $N(t) = n$ is

$$f_{T_1,\dots,T_n}\Big(t_1, t_2, \dots, t_n \Big| N(t) = n\Big) = n! f(t_1) \cdots f(t_n),$$

for $0 < t_1 < \cdots < t_n < t$, where $f(s) = r(s)/\mu(0, t]$. This is equal to the joint density of the order statistics of n independent random variables on $[0, t]$ with density $f(s)$.

20. Consider a Poisson process N on R_+ with rate function $r(t) = 3t^2$. Let $T_1 < T_2 < \dots$ denote the point locations.

(a) Show that $W_n = T_n - T_{n-1}$, $n \geq 1$, are dependent.
(b) Find the distributions of T_1 and T_n.
(c) Find the distribution of W_n.

21. Suppose N is a Poisson process on R^2 with rate function $\lambda(x, y)$ at the location (x, y). For instance, N could represent locations of certain types of animal nests, diseased trees, land mines, auto accidents, houses of certain types of people, flaws on a surface, potholes, ... Let D_n denote the distance from the origin to the n-th nearest point of N.
(a) Find an expression for the distribution and mean of D_1.
(b) Find an expression for the distribution of D_n when $\lambda(x, y) \equiv \lambda$.
(c) Are the differences $D_n - D_{n-1}$ independent (as they are for Poisson interpoint distances on R)?
(d) Suppose there is a point located at (x^*, y^*). What is the distribution of the distance to the nearest point?
(e) Is $\lambda(x, y) = 1/(x^2 + y^2)^{1/2}$ a valid rate function for N to be a Poisson process under our definition?
(f) Specify a rate function $\lambda(x, y)$ under which $P\{N(R^2) < \infty\} = 1$.

22. Let M denote a Poisson process on $\Re_+^d$ with intensity μ. Show that $N(t) = M((0, t]^d)$, for $t \geq 0$, is a Poisson process on $\Re_+$ with $E[N(t)] = \mu((0, t]^d)$. This is an alternate approach to proving the departure process in a $M_t/G_t/\infty$ system is Poisson; see Example 11.2.

23. *Highway Model.* Vehicles enter an infinite highway denoted by $\Re$ at times that form a Poisson process N on the time axis $\Re_+$ with intensity measure μ. For simplicity, assume the highway is empty at time 0. The vehicle arriving at time T_n enters at a location X_n on the highway $\Re$ and moves on it with a velocity V_n for a time τ_n and then exits the highway. The velocity may be negative, denoting a movement in the negative direction and vehicles may automatically pass one another on the highway with no change in velocity. The X_n are i.i.d. with distribution F and are independent of N. The pairs (V_n, τ_n) are independent of N and, they are conditionally independent given the X_n's with

$$G_x(v, t) = P\{V_n \leq v, \tau_n \leq t | X_k, k \geq 1, X_n = x\},$$

a non-random distribution independent of n.
(a) Justify that $M \equiv \delta_{(T_n, X_n, V_n, \tau_n)}$, $n \geq 1$, is a Poisson process on $\Re_+ \times R^2 \times R_+$ and describe its intensity.
(b) Consider the departure process D on $\Re \times \Re_+$ where $D(A \times (a, b])$ is the number of departures from A in the time interval $(a, b]$. Justify that D is a Poisson process and specify its intensity. Find the expected number of departures in $(0, t]$.
(c) For a fixed t let $N_t(A)$ denote the number of vehicles in $A \subset \Re$ at time t. Justify that N_t a Poisson process on $\Re$ and specify its intensity.
(d) Suppose a vehicle is at the location at $x \in \Re$ at time t and let $X(t)$ denote the distance to the nearest vehicle. Specify assumptions on μ, F and

$G_x(v,t)$ that would guarantee that there is at most one point at any location on the highway. Under these assumptions, find the distribution of $X(t)$.

24. *Continuation.* In the preceding highway model, assume there are vehicles on the highway at time 0 at locations that form a Poisson process $N_0(\cdot)$ with rate λ, and this process is independent of the other vehicles. The sojourn times of these vehicles on the highway are like those of the other vehicles vehicles, and they all operate under the distribution $G_x(v,t) = G(v)(1-e^{-\mu t})$. Solve parts (b)–(d) of the preceding exercise.

25. In the context of Example 8.7, suppose N is a homogeneous Poisson process on the unit disc S in $\Re^2$ with rate λ. For the Poisson processes N' and M related to the projection of N on the line $S' = [-1,1,]$, show that the rate function of N' is $\lambda = 2\lambda\sqrt{1-x^2}$, and that

$$E[M(A_u \times (0,b])] = \lambda \int_0^b (\sqrt{1-x^2} - u)\, dx$$

for $A_u = \{(x,y) \in S : y \geq u\}$ and $b \leq \sqrt{1-u^2}$.

Next, consider the transformation of N where a point in the unit disc S is mapped to the closest point on the unit circle C. Under this map, using polar coordinates, $\overline{M}(A \times B) = \sum_n \delta_{(R_n,\Theta_n)}(A)\delta_{\Theta_n}(B)$ denotes the number of points of N in A that are mapped into B. Justify that $\overline{M}$ is a Poisson process on $S \times C$, and give an expression for $E[\overline{M}(A \times (0,b])]$, where $A = \{(r,\theta) \in S : \theta \in (0,b], r \in [\frac{1}{\sin\theta+\cos\theta}, 1]\}$ and $b \leq \pi/2$ (note that $x + y = r(\sin\theta + \cos\theta) \geq 1$ when $(x,y) \in A$).

For a fixed B, consider the process $N(r) = \overline{M}(A_r \times B)$, $r \in [0,1]$, where A_r is a unit disc in $\Re^2$ with radius r. Show that N is a Poisson process and specify its rate function $\lambda(r)$.

26. Suppose $N_1,\ldots,N_m$ are independent Poisson processes on a space S with respective intensities $\mu_1,\ldots,\mu_m$, and let $N = \sum_{i=1}^m N_i$, which is a Poisson process with intensity $\mu = \mu_1 + \cdots \mu_m$. For instance, $N_i(B)$ might be the number of crimes of type i in a region B of a city and $N(B)$ is the total number of crimes. Show that the conditional distribution of $N_1(B),\ldots,N_m(B)$ given $N(B) = n$ is a multinomial distribution.

27. Let $N_1,\ldots,N_m$ be independent Poisson processes on $\Re_+$ with location-dependent rates $\lambda_1(t),\ldots,\lambda_m(t)$, respectively. Let τ_i denote the time of the first occurrence in process N_i. Find the distribution of $\tau \equiv \min_{1\leq i\leq m}\tau_i$. Find $P\{\tau_i = \tau\}$.

28. Let $X_1, X_2, \ldots$ be i.i.d. random variables with a distribution F. Show that $X_1,\ldots,X_n$ are distinct with probability one for any $n \geq 2$ if and only if F is continuous. (This statement is also true if these are random elements in a space S and "continuous F" is replaced by $F\{x\} = 0$, $x \in S$). Use induction and $P\{X_1 \neq X_2\} = \int_{\Re}(1 - F(\{x\})F(dx)$.

29. Justify that the space–time process M in the last paragraph of Example 9.11 is a Poisson process by verifying the defining conditions for a Poisson process.

30. As in Example 8.7, suppose $N = \sum_n \delta_{(X_n,Y_n)}$ is a Poisson process on the unit disc S in $\Re^2$ with location-dependent rate $\lambda(x,y)$. The Poisson process $M = \sum_n \delta_{((X_n,Y_n),X_n)}$ on $S \times S'$ represents the number of points of N that are projected onto the x-axis $S' = [-1,1]$. Give an expression for the location-dependent intensity of M. Now, switching to polar coordinates as in Example 8.7, the process $\overline{M} \equiv \sum_n \delta_{((R_n,\Theta_n),\Theta_n)}$ represents the number of points of N that are mapped onto the unit circle C.

31. Suppose M is a Poisson process on $\Re_+^2$ with intensity μ. Show that the process $N(t) \equiv M((0,t] \times (0,t])$, is a Poisson process on $\Re_+$ with $E[N(t)] = \mu((0,t] \times (0,t])$. This result can be used to prove that the departure process $D(t)$ for the $M_t/G_t/\infty$ model in Example 11.2 is a Poisson process. (More generally, $N(t) \equiv M(B_t)$ is a Poisson process for $B_t \uparrow \Re_+^2$.)

32. Deposits to a bank account occur at times that form a Poisson process with rate λ and the amounts deposited are independent random variables with distribution F (independent of the times). Also, withdrawals occur at times that form a Poisson process with rate μ and the amounts deposited are independent random variables with distribution G (independent of the times). The deposits and withdrawals are independent. Let $X(t)$ denote the balance of the bank account at time t, where $X(0) = 0$, and the balance may be negative. Finds the mean, variance and distribution of $X(t)$.

33. Suppose $X(t) = \min_{n \le N(t)} Y_n$, for $t \ge 0$, where $N = \sum_n \delta_{T_n}$ is a Poisson process on $\Re_+$ and Y_n are independent random variables with distribution F, independent of N. For instance, Y_n could be bids on a property and $X(t)$ is the smallest bid up to time t. Find the distribution and mean of $X(t)$. Answer this question for the more general setting in which Y_n are p-marks of T_n, where $p(t,(0,y])$ is the distribution of typical mark at time t.

34. *E-mail Broadcasting.* An official of an organization sends e-mail messages to various subgroups of the organization at times that form a Poisson process with intensity μ. Each message is sent to all the intended recipients simultaneously, and it is sent to individual i with probability p_i, $i = 1,\dots,m$, where m is the number of individuals in the organization. Let $N_i(t)$ denote the number of messages individual i receives from the official in the time interval $(0,t]$. Justify that N_i is a Poisson process and specify its intensity.

Consider a subset of individuals $I \subset \{1,\dots,m\}$. Specify whether or not the processes N_i, $i \in I$, are independent. Is the sum $\sum_{i\in I} N_i$ Poisson? If so, specify its intensity.

35. *Continuation.* In the setting of the preceding exercise, suppose there is a probability r_i that individual i will reply to an e-mail from the official, and then send the reply within time that has a distribution $G_i(t)$. Let $R_i(t)$ denote the number of replies the official receives in $(0, t]$ from all the messages sent to individual i. Is R_i Poisson? If so, specify its intensity. Is the sum $\sum_{i \in I} R_i$ Poisson? If so, specify its intensity.

Suppose individual i receives a message from the official, and is planning to reply to it. Find the probability that before the reply is sent, another message from the official will arrive?

36. A satellite circles a body in outer space and records a special feature of the body (e.g. rocks, water, low elevations) along a path it monitors. As an idealized model, assume the feature occurs on the polar-angle space $S \equiv ([0, 2\pi]$ at angles $\Theta_1 \leq \Theta_2 \ldots \leq 2\pi$ that form a Poisson process with intensity μ. We will only consider one orbit of the satellite. Suppose the satellite is moving at a (deterministic) velocity of γ radians per unit time. Upon observing an occurrence at Θ_n the satellite sends a message to a station that receives it after a time τ_n. Suppose the transmission times τ_n are independent with distribution G and are independent of the positions of the occurrences. Consider the point process M on $S \times \Re_+^2$, where $M((\alpha, \beta] \times (a, b] \times (c, d])$ is the number of occurrences in the radian set $(\alpha, \beta]$ that are observed in the time set $(a, b]$ and received at the station in the time set $(c, d]$. Describe the process M and its intensity measure in terms of the system data.

Next, let $N(t)$ denote the number of messages received at the station in $(0, t]$ whose transmission time exceeds a certain limit L, where $G(L) < 1$. Describe the process N and specify its intensity measure.

37. *Multiclass $M/G/\infty$ System.* Consider an $M/G/\infty$ system in which item arrive at times that form a Poisson process with rate λ. There are m classes or types of items and p_i is the probability that an item is of class i. The processing time of a class i item has a distribution $G_i(\cdot)$. Assume the system is empty at time 0. Let $Q_i(t)$ denote the quantity of class i items in the system at time t. Specify its distribution. Determine whether or not $Q_1(t), \ldots, Q_m(t)$ are independent. Let $D_i(t)$ denote the number of departures of class i items in $(0, t]$. Describe these processes including their independence.

38. *Limiting Behavior of $M/G/\infty$ System.* Consider the $M/G/\infty$ system in Example 11.2 with arrival rate λ and service distribution G, which has a mean α. Show that the limiting distribution of the quantity of items in the system $Q(t)$ is Poisson with mean $\lambda\alpha$ as $t \to \infty$. Use the fact that $\alpha = \int_0^\infty [1 - G(u)] du$. Turning to the departures, consider the point process $D_t(A)$ on $\Re_+$ that records the numbers of departures in a time set A after time t. In particular, show that the number of departures $D_t(0, b]$ in the

interval $(t, t+b]$ has a Poisson distribution with

$$E[D_t(0,b]] = \lambda\Big[\int_t^{t+b} G(u+b)du - \int_t^{t+b} G(u)du\Big].$$

Show that the limiting distribution of $D_t(0,b]$ is Poisson with mean λb. More generally, show that the finite-dimensional distributions of D_t converge to those of a homogeneous Poisson process D with rate λ. This proves, in light of Theorem 15.2, that $D_t \xrightarrow{d} D$.

39. *Spatial $M/G/\infty$ System.* Consider a system in which items enter a space S at times $T_1 \le T_2 \le \dots$ that form a Poisson process with intensity measure μ. The nth item that arrives at time T_n enters S at the location X_n and remains there for a time V_n and then exits the system. Suppose $F_t(\cdot)$ is the distribution of the location in S of an item arriving at time t, and $G_{(t,x)}(\cdot)$ is the distribution of an item's sojourn time at a location x. More precisely, assume (X_n, V_n) are location-dependent marks of T_n with distribution

$$p(t, A \times (0,v]) = \int_A G_{(t,x)}(v) F_t(dx).$$

Let $N_t(B)$ denote the number of items in the set $B \in \mathcal{S}$ at time t. Show that N_t is a Poisson process on S with

$$E[N_t(B)] = \int_{(0,t]} \int_B [1 - G_{(s,x)}(t-s)] F_s(dx) \mu(ds).$$

Next, let $D((a,b] \times B)$ denote the number of departures from the set B in the time interval $(a,b]$. Show that D is a space-time Poisson process on $\Re_+ \times S$ and specify $E[D((0,t] \times B)]$.

40. For the network in Example 12.5, justify that the following processes are Poisson and specify their intensity measures.

$D(t) = \#$ of items that depart from the network in $(0,t]$.
$D_1(t) = \#$ of items that enter node 1 and
depart from the network in $(0,t]$.
$D_{(2,3,5)}(t) = \#$ of items that complete the route $(2,3,5)$ in $(0,t]$.

Justify that the following random variables, for a fixed t, have a Poisson distribution and specify their means.

$Q(t) = \#$ of items that are in the network at time t.
$Q_1(t) = \#$ of items that are beyond their first node at time t.
$Q_{3|2}(t) = \#$ of items in node 3 at time t that came from node 2.

41. *Time Transformations and Cox Processes.* Let $N_1(t) = \sum_n \mathbf{1}(T_n \le t)$ denote a homogeneous Poisson process on $\Re_+$ with rate λ, and η is a locally finite measure on $\Re_+$ with $\eta(\Re_+) = \infty$. Consider the process $N(t) =$

$N_1(\eta(t))$, $t \geq 0$. Show that N is a transformation of N_1 under a map $g : \Re_+ \to \Re_+$; that is find g such that

$$N(t) = \sum_n \mathbf{1}(T_n \leq \eta(t)) = \sum_n \mathbf{1}(g(T_n) \leq t).$$

Use this to show that N is a Poisson process on $\Re_+$ with intensity η.

Next, suppose that η is a locally-finite random measure. Then N is a Cox process as in Example 13.1. Show that

$$E[e^{-Nf}|\eta] = \exp\{-\int_{\Re_+} (1 - e^{-f(t)})\eta(dt)\}\}, \qquad f \in C_K^+(\Re_+).$$

42. Suppose that N is a *Cox process* on S directed by a locally-finite random measure η. Show that $E[N(B)] = \mathrm{Var}N(B)$, for $B \in \mathcal{S}$.

43. *Poisson Process Directed by a Cyclic Renewal Process.* The state of a system is represented by a continuous-time cyclic renewal process $X(t)$ on states $0, 1, \ldots, K-1$ as in Example 1.9. The sojourn times in the states are independent, and the sojourn time in state i has a continuous distribution F_i with mean μ_i. Exercise 48 shows that

$$\lim_{t\to\infty} P\{X(t) = i\} = \frac{\mu_i}{\mu_0 + \ldots + \mu_{K-1}}.$$

Suppose the system fails occasionally such that, while it is in state i, failures occur according to a Poisson process with rate λ_i, independent of everything else. Let $N(t)$ denote the number of failures in $(0, t]$. Show that

$$t^{-1}N(t) \to \frac{\sum_{k=0}^{K-1} \lambda_i \mu_i}{\sum_{k=0}^{K-1} \mu_i}, \quad a.s. \text{ as } t \to \infty.$$

Assume the system begins in state 0 and let τ denote the first time it returns to state 0 (the time to complete a cycle). Show that

$$E[N(\tau)] = \sum_{k=0}^{K-1} \lambda_i \mu_i = \mathrm{Var}N(\tau).$$

44. *Location-Dependent Compound Poisson Process.* Let $Z(t) = \sum_{n=0}^{N(t)} Y_n$ be a location-dependent compound Poisson process, where $N = \sum_n \delta_{T_n}$ is a Poisson process on $\Re_+$ with intensity measure μ and Y_n are p-marks of T_n. Show that the process $Z(t)$ has independent increments (the increments will not be stationary in general), and

$$E[Z(t)] = \int_{(0,t]} \int_{\Re} yp(s, dy)\mu(ds).$$

Suppose the moment generating function $\phi_{s,t}(\alpha) \equiv \int_{\Re} e^{\alpha y} F_{s,t}(dy)$ exists, where

$$F_{s,t}(y) = \int_{(s,t]} p(u, (0, y])\mu(du)/\mu(s, t].$$

Show that, for $s < t$,

$$E[e^{\alpha[Z(t)-Z(s)]}] = e^{-\mu(s,t][1-\phi_{s,t}(\alpha)]}. \tag{16.1}$$

(This is the moment generating function of a compound Poisson distribution with rate $\mu(s,t]$ and distribution $F_{s,t}$.) Use the fact

$$E[e^{\alpha[Z(t)-Z(s)]}] = E[e^{\int_{\Re} yM((s,t]\times dy)}] = E[e^{-Mh_{s,t}}],$$

where $h_{s,t}(u,y) = -\alpha y \mathbf{1}(u \in (s,t])$ and $M = \sum_n \delta_{(T_n,Y_n)}$.

45. Suppose $Z_1(t),\dots,Z_m(t)$, are independent compound Poisson processes with respective rates $\lambda_1,\dots,\lambda_m$, and distributions $F_1,\dots,F_m$. Show that $Z(t) = \sum_{i=1}^m Z_i(t)$ is a compound Poisson process with rate $\lambda = \sum_{i=1}^n \lambda_i$ and distribution $F = \sum_{i=1} \frac{\lambda_i}{\lambda} F_i$.

46. *Partition of a Compound Poisson Process.* Suppose $Z(t) = \sum_{n=0}^{N(t)} Y_n$, for $t \geq 0$, is a location-dependent compound Poisson process with intensity measure μ and distribution $p(t,\cdot)$. Suppose the quantity Y_n at time T_n is partitioned into m pieces $\mathbf{Y}'_n \equiv (Y'_{n1},\dots,Y'_{nm})$ so that $Y_n = \sum_{i=1}^m Y'_{ni}$. These pieces are assigned to m processes defined by $Z_i(t) \equiv \sum_{n=0}^{N(t)} Y'_{ni}$. They form a partition of $Z(t)$ in that $Z(t) = \sum_{i=1}^m Z_i(t)$. Assume the $\mathbf{Y}'_n$ are p'-marks of (T_n,Y_n), where $p'((t,y),B_1\times\cdots\times B_m)$ is the conditional distribution of a typical vector $\mathbf{Y}'_n$ given $(T_n,Y_n)=(t,y)$. Prove that $Z_i(t)$ is a compound Poisson process with intensity μ and distribution $p'(t,\cdot)$, and specify $p'(t,\cdot)$. Use the idea that $\mathbf{Y}'_n$ are a second marking of the Poisson process N as discussed in Example 13.5, resulting in $M' = \sum_n \delta_{(T_n,Y_n,\mathbf{Y}'_n)}$.

47. *Origin-Dependent Cluster Processes.* The cluster process in Example 13.9 has the form

$$M(A\times B) = \sum_n N'_n(B)\delta_{X_n}(A),$$

where N'_n are point processes on a space S' generated by the points X_n in S. Instead of assuming the N'_n are independent of N, consider the more general setting in which the N'_n are p-marks of X_n. Let N'_x be a point process on S' such that $p(x,C) = P\{N'_x \in C\}$. Show (by conditioning on N) that the Laplace functional of M is

$$E[e^{-Mf}] = \exp[-\int_S (1-g(x))\mu(dx)],$$

where $g(x) \equiv E[e^{-\int_{S'} f(x,x')N'_x(dx')}]$.

48. The moments of a point process N on S are given by

$$\begin{aligned} &E[N(A_1)^{n_1}\cdots N(A_k)^{n_k}] \\ &\quad = (-1)^{n_1+\cdots+n_k}\frac{\partial^{n_1+\cdots+n_k}}{\partial t_1^{n_1}\cdots\partial t_k^{n_k}} E[e^{-Nf}]|_{t_1=\ldots=t_k=0}, \end{aligned}$$

where $f(x) = \sum_{i=1}^{k} t_i 1(x \in A_i)$. Prove this for $k = 1$ and $k = 2$. Use this fact to find expressions for the first two moments of the cluster process quantity $M(A \times B)$ in Exercise 47.

49. *Markov/Poisson Particle System.* Consider a particle system in a countable space S similar to the one in Example 9.8 with the following modifications. Each particle moves independently in continuous time according to a Markov jump process with stationary distribution p_i, $i \in S$. Let $P^t(i,j)$ denote the probability that a particle starting in state i is in state j at time t. Assume the system is empty at time 0 and that particles enter the system according to a space-time Poisson process M on $\Re_+ \times S$, where $M((0,t] \times B)$ is the number of arrivals in $(0,t]$ that enter $B \subset S$, and $E[M((0,t] \times \{i\})] = \lambda t p_i$. Let $Q_i(t)$ denote the quantity of particles in state i at time t. Show that

$$(Q_i(t) : i \in S) \xrightarrow{d} (Q_i : i \in S), \qquad \text{as } t \to \infty,$$

where Q_i are independent Poisson random variables with $E[Q_i] = \lambda p_i$.

50. *Continuation.* In the setting of the preceding exercise, suppose at time 0 the number of particles in the system is a point process with intensity μ that is independent of the space-time arrival process M of other particles and all the particles move independently as above. The quantity of particles in state i at time t is $X_i(t) \equiv Q_i^0(t) + Q_i(t)$, where $Q_i^0(t)$ denotes the quantity of particles in i at time t that were in the system at time 0. Show that $E[Q_i^0(t)] = \sum_{j \in S} P^t(j,i)\mu(i)$, and find $\alpha_i \equiv \lim_{t\to\infty} E[X_i(t)]$. Prove

$$Q_i^0(t) \xrightarrow{d} Q_i^0, \qquad \text{as } t \to \infty, \text{ for } i \in S$$

where Q_i^0 are independent Poisson random variables with $E[Q_i] = p_i$. Show that $\lim_{t\to\infty} P\{X_i(t) = n\} = e^{-\alpha_i}(\alpha_i)^n/n!$.

51. *Sums of Identically Distributed Renewal Processes.* Let $\tilde{N}_1, \tilde{N}_2 \dots$ denote independent renewal processes with inter-renewal distribution F. By the strong law of large numbers, we know that the sum $\sum_{i=1}^{n} \tilde{N}_i(t)$ converges to ∞ a.s. as $n \to \infty$. (The discussion prior to Example 15.5 addressed the opposite case where the sum tends to 0.) To normalize this sum (as in a central limit theorem) so that it converges to a non-degenerate limit, it is natural to rescale the time axis and consider the process $N_n(t) = \sum_{i=1}^{n} \tilde{N}_i(t/n)$. This is the sum with $1/n$ as the new unit of time. Assume the derivative $\lambda \equiv F'(0)$ exists and is positive. Show that $N_n \xrightarrow{d} N$, where N is a Poisson process with rate λ.

52. *Poisson Limit of Thinned Processes.* Let N_n be a sequence of point processes on S. Suppose N_n is subject to a $p_n(x)$ thinning: A point of N_n at x is retained with probability $p_n(x)$ and is deleted with probability $1-p_n(x)$. Let N_n' denote the resulting thinned process on S. Assume the thinning is

uniformly null in that

$$\lim_{n\to\infty} \sup_x p_n(x) = 0, \quad B \in \hat{\mathcal{S}}.$$

Show that $N_n' \xrightarrow{d} N'$, a Poisson process on S with intensity measure μ, if

$$\int_B p_n(x) N_n(dx) \xrightarrow{d} \mu(B), \qquad B \in \hat{S}_\mu, \text{ as } n \to \infty.$$

CHAPTER 7

Appendix

This appendix covers background material from probability theory and real analysis. Included are a review of elementary notation and concepts or probability as well as theorems from measure theory, which are major tools of applied probability. More details can be found in the following textbooks:
Probability Theory — Billingsley 1968, Breiman 1992, Chung 1974, Durrett 2005, Feller 1972, Grimmett and Stirzaker 2001, Kallenberg 2004, Shiryaev 1995.
Real Analysis — Ash and Doléans-Dade 2000, Bauer 1972, Hewitt and Stromberg 1965.

1. Probability Spaces and Random Variables

The underlying frame of reference for random variables or a stochastic process is a probability space. A *probability space* is a triple $(\Omega, \mathcal{F}, P)$, where Ω is a set of *outcomes*, $\mathcal{F}$ is a family of subsets of Ω called *events*, and P is a *probability measure* defined on these events. The family $\mathcal{F}$ is a *σ-field* (or σ-algebra): If $A \in \mathcal{F}$, then so is its complement A^c, and if a sequence A_n is in $\mathcal{F}$, then so is its union $\cup_n A_n$. The probability measure P satisfies the properties of being a *measure*: It is a non-negative function on $\mathcal{F}$ such that, for any finite or countably infinite collection of disjoint sets A_n in $\mathcal{F}$,

$$P(\cup_n A_n) = \sum_n P(A_n).$$

Furthermore, P satisfies $P(\Omega) = 1$.

Under this definition, $P(A) \leq 1$, $P(A^c) = 1 - P(A)$, where $A^c = \Omega \backslash A$ (the complement of A), and

$$\begin{aligned} P(A) &\leq P(B), \qquad A \subset B, \\ P(A_n) &\to P(A), \qquad \text{if } A_n \uparrow A \text{ or } A_n \downarrow A. \end{aligned}$$

The definition of a random variable involves the notion of a measurable function. Suppose $(S, \mathcal{S})$ and $(S', \mathcal{S}')$ are measurable spaces (sets with associated σ-fields). A function $f : S \to S'$ is *measurable* if

$$f^{-1}(A) \equiv \{x \in S : f(x) \in A\} \in \mathcal{E}, \qquad A \in \mathcal{E}'.$$

That is, the set of all x's that f maps into A is in $\mathcal{E}$. Typically, S will be the outcome space Ω, the real line $\Re$, the d-dimensional Euclidean space $\Re^d$, or a metric space. We adopt the standard convention that the σ-field $\mathcal{S}$ for

S is its *Borel* σ-field — the smallest σ-field containing all open sets in S (or the σ-field consisting of countable unions of open sets and all complements of these). We sometimes write $\mathcal{B}$ and $\mathcal{B}_+$ for the Borel σ-fields of $\Re$ and $\Re_+$. A useful property is that if $f : S \to S'$ and $g : S' \to S''$ are measurable, then the *composition function* $g \circ f(x) \equiv g(f(x))$ is measurable.

A *random variable* X on a probability space $(\Omega, \mathcal{F}, P)$ is a measurable mapping from Ω to $\Re$. The measurability of X ensures that $\mathcal{F}$ contains all sets of the form

$$\{X \in B\} \equiv \{\omega \in \Omega : X(\omega) \in B\}, \qquad B \text{ a Borel set in } \Re. \tag{1.1}$$

These are the types of events for which P is defined. One usually constructs (or assumes) the σ-field $\mathcal{F}$ is large enough such that the random variables of interest are measurable. For instance, if X, Y and Z are of interest, one can let $\mathcal{F} = \sigma(X, Y, Z)$, the "smallest σ-field" containing all sets of the form (1.1) for X, Y and Z, so that they are measurable. A statement about events or random variables is said to hold *almost surely* (a.s.) if the statement holds with probability one (some say the statement is true almost everywhere (a.e.) on Ω with respect to P). For instance $X + Y \le Z$ a.s. Also, we sometimes omit a.s. from elementary statements like $X = Y$ and $X \le Y$ that hold a.s.

All of the probability information of X in "isolation" (not associated with other random quantities on the probability space) is contained in its *distribution function*

$$F(x) = P\{X \le x\}, \qquad x \in \Re.$$

Here $P\{X \le x\} = P(\{\omega : X(\omega) \le x\})$. A distribution function has at most a finite or countable number of discontinuities (which may be a dense set); this is a well-known property of any increasing function. We sometimes write the distribution as $F_X(x)$.

The random variable X is *discrete* if the range of X is a countable set S in $\Re$. In this case, the *probability function of* X is $P\{X = x\}$, $x \in S$; and

$$P\{X \in A\} = \sum_{x \in A} P\{X = x\}, \qquad A \subset S.$$

The random variable X is *continuous* if there is a (measurable) *density function* $f : \Re \to \Re_+$ such that $\int_{-\infty}^{\infty} f(x)\,dx = 1$ and $P\{X \in A\} = \int_A f(x)\,dx$, $A \subset \Re$. Then the distribution of X is $F(x) = \int_{-\infty}^{x} f(y)\,dy$, and so $f(x) = F'(x)$, the derivative of F. Standard distribution functions are in the next section.

2. Table of Distributions

The following are tables of some standard distributions and their means, variances, and moment generating functions.

Discrete Random Variables

Random Variable	$P\{X = x\}$	EX	Var$X$	$E[e^{sX}]$
Binomial $n \geq 1,\ p \in (0,1)$	$\binom{n}{x} p^x (1-p)^{n-x}$ $x = 0, 1, \ldots, n$	np	$np(1-p)$	$(pe^s + (1-p))^n$
Poisson $\lambda > 0$	$e^{-\lambda}\lambda^x/n!$ $x = 0, 1, \ldots$	λ	λ	$e^{-\lambda(e^s - 1)}$
Geometric $p \in (0,1)$	$p(1-p)^{x-1}$ $x = 1, 2, \ldots,$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^s}{1-(1-p)e^s}$
Negative Binomial $r \geq 1,\ p \in (0,1)$	$\binom{x-1}{r-1} p^r (1-p)^{x-r}$ $x = r, r+1, \ldots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$(\frac{pe^s}{1-(1-p)e^s})^r$

Continuous Random Variables

Random Variable	Density $f(x)$	EX	Var$X$	$E[e^{sX}]$
Uniform on $[a, b]$	$\frac{1}{b-a}, \quad x \in [a, b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bs} - e^{as}}{s(b-a)}$
Exponential $\lambda > 0$	$\lambda e^{-\lambda x}, \quad x \geq 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - s}$
Gamma $n \geq 1,\ \lambda > 0$	$\frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!}, \quad x \geq 0$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$	$(\frac{\lambda}{\lambda - s})^n$
Normal $\mu \in \Re,\ \sigma > 0$	$\frac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}, \quad x \in \Re$	μ	σ^2	$e^{\mu s + \sigma^2 s^2/2}$

3. Random Elements and Stochastic Processes

A unified way of discussing random vectors, stochastic processes and other random quantities is in terms of random elements. Suppose one is interested in a random element that takes values in a space S with a σ-field $\mathcal{S}$. A *random element in* S, defined on a probability space $(\Omega, \mathcal{F}, P)$, is a measurable mapping X from Ω to S.

For our purposes, the space S will be a countable set, a Euclidean space $\Re^d$, or a function space with a distance metric (for representing a stochastic process). To accommodate these and other spaces as well, we adopt the standard convention that $(S, \mathcal{S})$ is a *Polish space*. That is, S is a metric space that is complete (each Cauchy sequence is convergent) and separable (there is a countable dense set in S); and $\mathcal{E}$ is the Borel σ-field generated by the open sets. A *metric* on S is a map $d : S \times S \to \Re_+$ such that

$$d(x,y) = d(y,x), \qquad d(x,y) = 0 \text{ if and only if } x = y,$$
$$d(x,z) \leq d(x,y) + d(y,z), \qquad x, y, z \in S.$$

Our discussion of functions, integrals, convergence, etc. on S does not require a familiarity of Polish spaces, since these concepts are understandable by interpreting them as being on $\Re^d$. We use terminology involving Polish spaces and random elements in this appendix because it allows for a rigorous and unified presentation of background material, but this terminology is not used throughout the book.

The *probability distribution* of a random element X in S is the probability measure

$$F_X(B) \equiv P\{X \in B\} = P \circ X^{-1}(B), \qquad B \in \mathcal{E}.$$

If X and Y are random elements whose distributions are equal, we say that X is *equal in distribution* to Y and denote this by $X \stackrel{d}{=} Y$. The underlying probability spaces for X and Y need not be the same.

Loosely speaking, a stochastic process is a collection of random variables (or random elements) defined on a single probability space. Hereafter, we will simply use the term "random elements" (which includes random variables), and let $(S, \mathcal{S})$ denote the Polish space where they reside.

A *discrete-time stochastic process* (or random sequence) is a collection of random elements $X \equiv \{X_n : n \geq 0\}$ in S defined on a probability space $(\Omega, \mathcal{F}, P)$. The nonnegative integer n is a *time parameter* and S is the *state space* of the process. The value $X_n(\omega) \in S$ is the *state* of the process at time n associated with the outcome ω.

Note that X is also a random element in the infinite product space S^∞ with the product σ-field $\mathcal{S}^\infty$: the smallest σ-field generated by sets $B_1 \times \cdots \times B_n$, B_j's $\in \mathcal{S}$. Its distribution $P\{X \in B\}$, for $B \in \mathcal{S}^\infty$, is uniquely defined in terms of its *finite-dimensional distributions*

$$P\{X_1 \in B_1, \ldots, X_n \in B_n\}, \qquad B_j\text{'s} \in \mathcal{S},\ n \geq 1.$$

One consequence is that if Y is another random element in S^∞ whose finite-dimensional distributions are equal to those of X then $X \stackrel{d}{=} Y$. We sometimes refer to the process $\{X_n : n \geq 0\}$ simply by X_n.

Stochastic processes in continuous time are defined similarly to those in discrete time, but their evolutions over time are more subtle. A *continuous-time stochastic process* is a collection of random elements $\{X(t) : t \geq 0\}$ in S defined on a probability space, where $X(t, \omega)$ is the state at time t

associated with the outcome ω. The function $t \to X(t,\omega)$ from $\Re_+$ to S, for a fixed ω, is the *sample path* or *trajectory* associated with the outcome ω. Accordingly, $X(t)$ is a random function from $\Re_+$ to S. More precisely, the entire process $X \equiv \{X(t) : t \geq 0\}$ is a random element in a space of functions from $\Re_+$ to S. We sometimes refer to the process $\{X(t) : t \geq 0\}$ simply by $X(t)$.

A standard example is when the sample paths of X are in the set $D(\Re_+, S)$ of functions from $\Re_+$ to S that are right-continuous with left hand limits — often called *cadlag* functions (from the French *continu à droite, limites à gauche*). Then X is a random element in $D(\Re_+, S)$, with an appropriate metric depending on one's application (commonly a metric for the Skorohod topology [**5, 20**]), and $P\{X \in B\}$ is for a Borel set $B \subset D(\Re_+, S)$ of sample paths. The distribution of X is uniquely determined by its *finite-dimensional distributions*

$$P\{X(t_1) \in B_1, \ldots, X(t_n) \in B_n\}, \qquad t_1 < \cdots < t_n,\ B_j \in \mathcal{S},\ n \geq 1.$$

In summary, a stochastic process is a family of random variables or random elements defined on a probability space that contains all the probability information about the process. We will use the standard convention of suppressing the ω in random elements such as X_n or $X(t)$, and not displaying the underlying probability space $(\Omega, \mathcal{F}, P)$, unless it is essential for the exposition. Also, all the functions appearing in this book are measurable, and we will mention this property only when it is needed.

4. Expectation

The expectation (or expected value or mean) of a random variable is defined as follows. Recall the notation above for discrete and continuous random variables. We also say that an integral $\int_{-\infty}^{\infty} g(x)\,dx$ *exists* if it is absolutely convergent: $\int_{-\infty}^{\infty} |g(x)|\,dx < \infty$. Existence of an infinite sum is defined similarly.

Definition 15. Let X be a random variable, and denote its distribution function by $F(x) = P\{X \leq x\}$. The *expectation* of X is defined by

$$EX \equiv \int_{\Re} x dF(x), \tag{4.1}$$

(a Riemann-Stieltjes integral defined below) provided the integral exists. In particular,

$$EX \equiv \sum_{x \in S} x P\{X = x\}, \qquad \text{if } X \text{ is discrete,}$$

$$EX \equiv \int_{\Re} x f(x)\,dx, \qquad \text{if } X \text{ is continuous,}$$

provided the sum and integral exist.

The general formula (4.1) is needed for random variables that are not discrete or continuous. For instance if X has positive probabilities at points in a countable set S, and also has a density $f(x)$ elsewhere on $\Re$, then

$$EX = \sum_{x \in S} xP\{X = x\} + \int_{\Re} xf(x)\,dx.$$

Riemann-Stieltjes integrals are similar to Riemann integrals in calculus. A Riemann-Stieltjes integral of a function $g : [a, b] \to \Re$ with respect to F is constructed by limits of the upper and lower Darboux sums $\mathcal{D}^{\chi}$ and $\mathcal{D}_{\chi}$ defined on the set of points $\chi = \{a = x_0 < x_1 < \ldots < x_n = b\}$ by

$$\mathcal{D}^{\chi} \equiv \sum_{j=1}^{n} \sup\{g(x) : x_{j-1} \le x \le x_j\}[F(x_j) - F(x_{j-1})],$$

and $\mathcal{D}_{\chi}$ is defined similarly with sup replaced by inf. Specifically, the Riemann-Stieltjes integral of g exists if, for any $\epsilon > 0$, there is a set χ depending on g and ϵ such that $\mathcal{D}^{\chi} - \mathcal{D}_{\chi} < \epsilon$. When it exists, the integral has the form (e.g., see [**18**])

$$\int_{[a,b]} g(x)dF(x) = \inf_{\chi} \mathcal{D}^{\chi} = \sup_{\chi} \mathcal{D}_{\chi}.$$

This integral is a Riemann integral $\int_{[a,b]} g(x)f(x)dx$ when $dF(x) = f(x)dx$ and dx is the *Lebesgue* measure on $\Re$.

Riemann-Stieltjes integrals on infinite intervals are defined similar to Riemann integrals. For instance, for $g : \Re \to \Re$,

$$\int_{\Re} g(x)dF(x) = \lim_{a,b \to \infty} \int_{[-a,b]} g(x)dF(x),$$

provided the limit exists and is finite.

These Riemann-Stieltjes integrals on intervals in $\Re$ are equivalent to Lebesgue integrals [**18**], where F is a measure defined by $F((a, b]) = F(b) - F(a)$, $a < b$. For instance,

$$\int_{\Re} g(x)dF(x) = \int_{\Re} g(x)F(dx),$$

where the latter is a Lebesgue integral and $F(dx)$ denotes that F is to be interpreted as a measure. Riemann-Stieltjes integrals are defined only for measures on $\Re$, while Lebesgue integrals are defined for measures on general spaces. Another equivalent expression for the expectation of X in terms of the probability P is the Lebesgue integral

$$EX \equiv \int_{\Omega} X(\omega)P(d\omega).$$

This form is not used in the text.

A few properties of the expectation operator are $Ea = a$, for $a \in \Re$,

$$E[X + Y] = EX + EY, \qquad EX \le EY \quad \text{if } X \le Y,$$

$$E[\sum_{j=1}^{n} a_j X_j] = \sum_{j=1}^{n} a_j E X_j.$$

In addition to the mean EX, other summary measures of a random variable X are as follows. The *nth moment* of X is $E[X^n]$, and the *nth moment about its mean* $\mu = EX$ is $E[(X-\mu)^n]$. The *variance* of X is

$$\mathrm{Var}X \equiv E[(X-\mu))^2] = E[X^2] - \mu^2.$$

Whenever we refer to these moments, we assume they are finite.

5. Functions of Stochastic Processes

Many features of a stochastic process, or related quantities of interest, are expressed as functions of the process. This section contains several examples and a formula for evaluating expectations of real-valued functions of random elements and processes.

Suppose X is a random element in a Polish space S, such as a discrete- or continuous-time process $X \equiv \{X_n : n \geq 0\}$ or $X \equiv \{X(t) : t \geq 0\}$, and denote its distribution by $F_X(B) = P\{X \in B\}$, $B \in \mathcal{S}$. Consider a measurable function $g : S \to S'$ and define $Y = g(X)$. This Y is a random element in S' since it is a composition of X and g, which are measurable. When X is a continuous-time process, $g(x)$ is a function on the space of sample paths $x = \{x(t) : t \geq 0\}$. The distribution of Y is

$$P\{Y \in B\} = P\{g(X) \in B\} = P\{X \in g^{-1}(B)\},$$

where $g^{-1}(B) = \{x \in S : g(x) \in B\}$. Then the distribution of Y as a function of F_X is the probability measure

$$F_Y(B) = F_X \circ g^{-1}(B) \equiv F_X(g^{-1}(B)), \qquad B \in \mathcal{S}'.$$

In some cases, the function g is a standard measurable operation on real numbers such as addition, subtraction, maximum, etc. For instance, if $X \equiv \{X_n : n \geq 0\}$ is a family of random variables, then $Y = X_1 + \cdots + X_n$ is a random variable for fixed $n < \infty$, since the addition function $g(x) \equiv x_1 + \cdots + x_n$ from $\Re^\infty$ to $\Re$ is measurable. Other standard examples of random variables that are measurable functions of X include

$$\prod_{j=1}^{n} X_j, \qquad \max_{1 \leq j \leq n} X_j, \qquad \sum_{j=1}^{n} a_n (X_n - X_j).$$

Examples based on multiple compositions of functions are

$$\sup_{n \geq 0} X_n - \inf_{n \geq 0} X_n, \qquad \sup_{n \geq 0} \Big[e^{-aX_n} \sum_{j=1}^{n} (Y_j - Y_{j-1}) \Big],$$

provided they exist.

Examples of $g(X)$ for a continuous-time process X include analogues of those above as well as

$$\int_0^t X(t-s)\,ds, \qquad \int_{\Re_+} e^{-a(t)} \inf_{s\le t} X(s)\,dt.$$

In modelling a stochastic system, the state of the system, or a performance measure for it, are often of the form $Y(t) = g(t, X)$, where the process X represents the time-dependent system data, and the function $g(t, x)$ represents the dynamics of the system.

We will now describe a useful formula for the mean of a real-valued function $Y = g(X)$ of the random element X. The mean of the random variable Y in terms of the distribution F_X of X is

$$E[g(X)] = \int_S g(x)F_X(dx), \tag{5.1}$$

provided the (Lebesgue) integral exists. This follows since $F_Y = F_X \circ g^{-1}$ and, by the change-of-variable formula below, we have

$$EY = \int_\Re y F_X \circ g^{-1}(dy) = \int_S g(x)F_X(dx).$$

Change-of-variable Formula for Lebesgue integrals. Suppose F is a measure on S, and $g : S \to S'$ and $h : S' \to \Re$ are measurable. Then

$$\int_S h(g(x))F(dx) = \int_{S'} h(y)F \circ g^{-1}(dy), \tag{5.2}$$

provided both integrals exist (one exists if and only if the other one does).

Important functions of random variables are generating functions and transforms. They are tools for characterizing distributions and evaluating their moments. The *moment-generating function* of a random variable X is

$$m_X(s) \equiv E[e^{sX}] = \int_\Re e^{sx} dF_X(x),$$

provided the integral exists for s in some interval $[0, a]$, where $a > 0$. A major property is that a moment-generating function uniquely determines a distribution and vice versa ($m_X = m_Y$ if and only if $F_X = F_Y$). Also, the nth moment of X, when it exists, has the representation

$$E[X^n] = m_X^{(n)}(0),$$

which is the nth derivative of m_X at 0. Moment generating functions of some standard distributions are given in Section 2 below.

Two variations of moment generating functions for special types of random variables are as follows. For a nonnegative random variable X, its *Laplace transform* (or the Laplace-Stieltjes transform of F_X) is

$$E[e^{-sX}] = \int_{\Re_+} e^{-sx} dF_X(x), \qquad s \ge 0.$$

For a discrete random variable X whose range is contained in the nonnegative integers, its *generating function* is

$$E[s^X] = \sum_{n=0}^{\infty} s^n P\{X = n\}, \qquad -1 \le s \le 1.$$

Laplace transforms and generating functions play the same role as moment generating functions in that they uniquely determine distribution functions and yield moments of random variables. Laplace transforms are also defined for increasing functions F that need not be bounded, such as renewal functions, and they can also be extended to the complex plane. A similar statement applies to generating functions.

A generalization of a moment generating function is a characteristic function. The *characteristic function* of a random variable X (or the *Fourier-Stieltjes transform* of F_X) is defined by

$$E[e^{isX}] = \int_{\Re} e^{isx} dF_X(x) \qquad s \in \Re,$$

where $i = \sqrt{-1}$ and $e^{isx} = \cos sx + i \sin sx$. A characteristic function, which is complex-valued, exists for "any" random variable (or distribution function). In contrast, a moment generating function is real-valued, but it only exists for a random variable whose moments exist. There is a one-to-one correspondence between distribution functions and characteristic functions, and moments are expressible by derivatives of characteristic functions at 0. A characteristic function is typically used when the more elementary moment generating function, Laplace transform, or generating function are not applicable.

The following are useful inequalities involving expectations of functions of random variables. For a random variable X and increasing $g : \Re \to \Re_+$,

$$P\{X \ge x\} \le E[g(X)]/g(x), \qquad \textit{Markov's Inequality.}$$

An example is

$$P\{|X - EX| \ge x\} \le \mathrm{Var}X/x^2, \qquad \textit{Chebyshev's Inequality.}$$

For random variables X, Y with finite second moments,

$$E|XY| \le \sqrt{E[X^2]E[Y^2]}, \qquad \textit{Cauchy-Buniakovsky-Schwarz.}$$

For random variables $X_1, \ldots, X_n$ and convex $g : \Re^n \to \Re$,

$$E[g(X_1, \ldots, X_n)] \ge g(EX_1, \ldots, EX_n), \qquad \textit{Jensen's Inequality.}$$

6. Independence

In this section, we define independent random variables and random elements, and describe several functions of them including summations.

Random variables $X_1, \ldots, X_n$ are *independent* if, for any Borel sets $B_1, \ldots, B_n$,

$$P\{X_1 \in B_1, \ldots, X_n \in B_n\} = \prod_{k=1}^{n} P\{X_k \in B_k\} = \prod_{k=1}^{n} F_{X_k}(B_k).$$

An infinite family of random variables are independent if any finite collection of the random variables are independent. The same definitions apply to independence of random elements such as stochastic processes.

There is the subtle issue of the existence of independent random elements. That is, given a sequence of distributions, does there exist a probability space and random elements on the space that are independent and have these distributions? The existence is justified by the following result.

THEOREM 6.1. (Existence of Independent Random Elements) *If $P_1, P_2, \ldots$ are probability measures on the respective Polish spaces $S_1, S_2, \ldots$, then there exist a probability space $(\Omega, \mathcal{F}, P)$ and independent random elements $X_1, X_2, \ldots$ defined on it such that $P\{X_n \in \cdot\} = P_n(\cdot)$ for each $n \geq 1$.*

Many properties of functions of random elements can be analyzed in terms of the separate distributions. Here is an important formula for expectations. Suppose X and Y are independent random elements in S and S', respectively, and $g : S \times S' \to \Re$ is measurable. Then by (5.1) and Fubini's theorem stated below, we have

$$\begin{aligned} E[g(X, Y)] &= \int_{S \times S'} g(x, y) F_X(dx) F_Y(dy) \\ &= \int_{S'} \Big[\int_S g(x, y) F_X(dx) \Big] F_Y(dy), \end{aligned} \tag{6.2}$$

provided the integral exists.

For the next result, we use the notion that a measure space $(S, \mathcal{S}, \mu)$ is *σ-finite* if there is a partition $B_1, B_2, \ldots$ of S such that $\mu(B_n) < \infty$, for each n. A Polish space has this property.

THEOREM 6.3. (Fubini) *Suppose μ and η are measures on the respective σ-finite spaces $(S, \mathcal{S})$ and $(S', \mathcal{S}')$, and $g : S \times S' \to \Re$ is measurable. If g is nonnegative or $\int_{S \times S'} |g(x, y)| \mu(dx) \eta(dy)$ is finite, then*

$$\begin{aligned} \int_{S \times S'} g(x, y) \mu(dx) \eta(dy) &= \int_{S'} \Big[\int_S g(x, y) \mu(dx) \Big] \eta(dy) \\ &= \int_S \Big[\int_{S'} g(x, y) \eta(dy) \Big] \mu(dx). \end{aligned}$$

This says that if the integral exists on the product space, then it equals the single-space integrals done separately (in either order).

A special case of (6.2) is

$$E[g(X)h(Y)] = [\int_S g(x) F_X(dx)][\int_{S'} h(y) F_Y(dy)] = E[g(X)]E[h(Y)].$$

This generalizes, for independent $X_1, \ldots, X_n$ in S and $g_j : S \to \Re$, to

$$E[\prod_{j=1}^{n} g_j(X_j)] = \prod_{j=1}^{n} E[g_j(X_j)].$$

provided the expectations exist.

EXAMPLE 6.4. Suppose $X_1, \ldots, X_n$ are nonnegative random variables. Then the moment generating function of $Y = \sum_{j=1}^{n} X_j$ is

$$E[e^{sY}] = \prod_{j=1}^{n} E[e^{sX_j}].$$

Now, assume each X_j has an exponential distribution with rate λ, and so $E[e^{sX_j}] = \lambda/(\lambda - s)$, $0 \le s < \lambda$. Consequently, $E[e^{sY}] = [\lambda/(\lambda - s)]^n$, which is the moment generating function of a gamma (or Erlang) distribution with parameters λ and n (see Section 2). Hence Y has this gamma distribution.

One can also determine distributions of sums of random variables by convolutions of their distributions. Specifically, if X and Y are independent random variables, then

$$P\{X + Y \le z\} = \int_{\Re} F_Y(z - x) dF_X(x). \tag{6.5}$$

That is, $F_{X+Y}(z) = F_X \star F_Y(z)$, where $\star$ is the convolution operator defined below. To prove (6.5), first note that by (5.1) (even for dependent X and Y), we have

$$P\{X + Y \le z\} = E[\mathbf{1}(X + Y \le z)] = \int_{x+y \le z} F_{X,Y}(dx \times dy).$$

Then applying $F_{X,Y}(dx \times dy) = dF_X(x) dF_Y(y)$ (from the independence) and Fubini's theorem to this double integral yields (6.5).

Properties of convolutions are as follows. The *convolution* of two distributions F and G is defined by

$$F \star G(z) = \int_{\Re} G(z - x) dF(x). \tag{6.6}$$

Note that $F \star G = G \star F$. Also, if $F(0-) = G(0-) = 0$, then

$$F \star G(z) = \int_0^z G(z - x) dF(x).$$

If F and G have respective densities f and g, then the derivative of (6.6) yields the formula

$$f \star g(z) = \int_{\Re} g(z - x) f(x) dx.$$

These formulas reduce to sums when F and G are discrete distributions.

Convolutions of several distributions are defined in the obvious way, for instance $F \star G \star H = F \star (G \star H)$, and if X, Y and Z are independent random variables, then $F_{X+Y+Z} = F_X \star F_Y \star F_Z$. The nth fold convolutions $F^{n\star}(x)$

of a distribution F, for $n \geq 0$, are defined recursively by $F^{0\star}(x) = \mathbf{1}(x \geq 0)$ and, for $n \geq 1$,

$$F^{n\star}(x) = F^{(n-1)\star} \star F(x) = F \star \cdots \star F \quad n \text{ convolutions.}$$

If $T_n = X_1 + \cdots + X_n$ where the X_j are independent with distribution F, then $F_{T_n} = F^{n\star}$.

The definition (6.6) of a convolution also extends to more general functions $\mu\star h$, where μ is a measure on $\Re$ and $h : \Re \to \Re$ is such that the integral exists. For example, renewal theory involves convolutions of the form

$$U \star h(t) = \int_{[0,t]} h(t-s)dU(s),$$

where $U(t) = \sum_{n=0}^{\infty} F^{n\star}(t)$, $F(0) = 0$ and $h(t)$ is bounded on compact sets and is 0 for $t < 0$.

7. Conditional Probabilities and Expectations

We will define conditional probabilities for random elements that apply to random variables as well. Suppose X and Y, defined on a common probability space $(\Omega, \mathcal{F}, P)$, are random elements in Polish spaces S and S', respectively. A *probability kernel* from S' to S is a function $p : S' \times \mathcal{S} \to [0, 1]$ such that $p(y, \cdot)$ is a probability measure on $\mathcal{S}$ for each $y \in S'$, and $p(\cdot, B)$ is a measurable function for each $B \in \mathcal{S}$. There exists a probability kernel $p(y, B)$ from S' to S such that

$$P\{X \in B\} = \int_{S'} p(y, B)F_Y(dy), \qquad B \in \mathcal{S}. \tag{7.1}$$

The kernel p is unique in the sense that if p' is another such kernel, then $p(Y, \cdot) = p'(Y, \cdot)$ a.s. (e.g., see [**20**]).

Definition 16. Using the preceding notation, the (random) probability measure

$$P\{X \in B|Y\} = p(Y, B), \qquad B \in \mathcal{S}$$

is the *conditional probability measure of* X *given* Y. When X is a random variable with a finite mean, the *conditional expectation* of X given Y is

$$E[X|Y] = \int_S xp(Y, dx). \tag{7.2}$$

Conditional probabilities and expectations — which are random quantities — are sometimes written as non-random quantities

$$P\{X \in B|Y = y\} = p(y, B), \qquad E[X|Y = y] = \int_S xp(y, dx) \qquad y \in S'.$$

An important property of conditional expectations, which follows from the definition, is

$$EX = E[E[X|Y]] = \int_{S'} E[X|Y = y]F_Y(dy).$$

Similarly, $P\{X \in B\} = E[P\{X \in B|Y\}]$. These formulas are useful tools for evaluating the mean or distribution of X in terms of the conditional means or distributions.

Another important property is that $E[X|Y]$ is a measurable function of Y, because $E[X|Y] = h(Y)$, where $h(y) = E[g(X,y)] = \int_S xp(y,dx)$ is measurable. Since Definition 16 is for random elements, Y may represent several random elements. For instance $E[X|Y,Z]$ is a measurable function of Y, Z and

$$E[X|Z] = E\Big[E[X|Y,Z]\Big|Z\Big].$$

Definition 16 is consistent with the definition used for elementary random variables. For instance, in case Y is a discrete random variable, the probability kernel that satisfies (7.1) is

$$p(y,B) = P\{X \in B|Y = y\} = P\{X \in B, Y = y\}/P\{Y = y\}.$$

Next suppose X and Y are continuous random variables such that

$$P\{X \in A, Y \in B\} = \int_{A\times B} f(x,y)\,dx\,dy,$$

where $f(x,y)$ is their joint density. Then the probability kernel satisfying (7.1) is

$$p(y,B) = \int_B f(x,y)/f(y)\,dx.$$

Probability texts usually define conditional probabilities $P\{X \in B|\mathcal{F}\}$ and expectations $E[X|\mathcal{F}]$ for conditioning on a σ-field $\mathcal{F}$ instead of a random element. Definition 16 includes these cases when $\mathcal{F} = \sigma(Y)$, the smallest σ-field generated by Y. Since we will only be dealing with conditioning on random elements and not abstract σ-fields, Definition 16 is adequate for our purposes.

Here are some more properties of conditional expectations (assuming they exist) for measurable $f : S' \to \Re$ and $g : S \times S' \to \Re$:

$$\begin{aligned} E[Xf(Y)|Y] &= f(Y)E[X|Y], \\ E[g(X,Y)|Y = y] &= E[g(X,y)|Y = y]. \end{aligned}$$

Furthermore, if X and Y are independent, then

$$E[g(X,Y)|Y = y] = E[g(X,y)],$$

or equivalently $E[g(X,Y)|Y] = E[h(Y)]$, where $h(y) = E[g(X,y)]$. Most of the standard properties of expectations extend to conditional expectations. For instance, $P\{X \in B|Y\} = E[\mathbf{1}(X \in B)|Y]$,

$$\begin{aligned} E[X|Y] &\le E[Z|Y] \qquad \text{if } X \le Z, \\ E[f(X) + g(Z)|Y] &= E[f(X)|Y] + E[g(Z)|Y]. \end{aligned}$$

8. Convergence Concepts

Many properties of stochastic processes are expressed in terms of convergence of random variables and elements. There are several types of convergence, but for our purposes we primarily use convergence with probability one and convergence in distribution, which we now describe.

We begin with a review of convergence of real numbers. A sequence of real numbers x_n *converges* to some $x \in \Re$, denoted by $\lim_{n\to\infty} x_n = x$, if, for any $\epsilon > 0$, there exists a number N such that

$$|x_n - x| < \epsilon, \qquad n \geq N.$$

We sometimes refer to this convergence as $x_n \to x$.

One often establishes convergence with the quantities

$$\liminf_{n\to\infty} x_n \equiv \lim_{n\to\infty} \inf_{k\geq n} x_k, \qquad \limsup_{n\to\infty} x_n \equiv \lim_{n\to\infty} \sup_{k\geq n} x_k. \tag{8.1}$$

This *limit inferior* and *limit superior* clearly satisfy

$$-\infty \leq \liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n \leq \infty.$$

If both of these quantities are equal to a finite x, then $x_n \to x$.

For insight into this result, let a and b denote the lim inf and lim sup in (8.1) and assume they are finite. By its definition, a is the lower "cluster value" of the x_n's in that x_n is in the interval $[a, a+\epsilon]$ infinitely often (i.o.), for fixed $\epsilon > 0$. Similarly, x_n is in the interval $[b-\epsilon, b]$ i.o. Now x_n does not converge to a limit if $a < b$ (since x_n is arbitrarily close to both a and b i.o.). On the other hand, $x_n \to x$ if and only if $a = b = x$.

The preceding properties of real numbers readily extend to random variables. Suppose $X, X_1, X_2, \ldots$ are random variables on a probability space. The sequence X_n *converges with probability one* to X if $\lim_{n\to\infty} X_n(\omega) = X(\omega)$, for $\omega \in \Omega' \subset \Omega$, where $P(\Omega') = 1$. We denote this convergence by

$$X_n \to X, \quad \text{a.s.} \quad \text{as } n \to \infty.$$

Now, the quantities

$$\liminf_{n\to\infty} X_n, \qquad \limsup_{n\to\infty} X_n$$

are random variables, since the functions $\liminf_{n\to\infty} x_n$ and $\limsup_{n\to\infty} x_n$ are measurable. If there is a random variable X such that

$$\liminf_{n\to\infty} X_n = \limsup_{n\to\infty} X_n = X \quad a.s.,$$

then $X_n \to X$, a.s. as $n \to \infty$. This follows by the analogous property for a sequence of real numbers.

Next, we consider the convergence in probability as well as convergence a.s. of random elements.

Definition 17. Let $X, X_1, X_2, \ldots$ be random elements in a metric space S, where $d(x, y)$ denotes the metric distance between x and y. The sequence

X_n *converges a.s. to* X in S, denoted by $X_n \to X$ a.s., if $d(X_n, X) \to 0$ a.s. The X_n *converges in probability* to X, denoted by $X_n \xrightarrow{P} X$, if

$$\lim_{n\to\infty} P\{d(X_n, X) > \epsilon\} = 0, \qquad \epsilon > 0.$$

Many applications involve establishing the convergence of a function $f(X_n)$, when X_n converges. A useful tool for this is the following.

PROPOSITION 8.2. (Continuous-mapping Property) *Suppose* $X, X_1, X_2, \ldots$ *are random elements in a metric space* S*, and* $f : S \to S'$ *where* S' *is a metric space. Assume* f *is continuous on* S*, or on a set* B *such that* $X \in B$ *a.s. If* $X_n \to X$ *a.s. in* S*, then* $f(X_n) \to f(X)$ *a.s. The same statement is true for convergence in probability.*

For instance, if $(X_n, Y_n) \to (X, Y)$ a.s. in $\Re^2$, then

$$X_n + Y_n \to X + Y, \qquad X_n Y_n \to XY \qquad \text{a.s. in } \Re\ .$$

This follows since the addition and multiplication functions from $\Re^2$ to $\Re$ are continuous. Similarly, $X_n/Y_n \to X/Y$ if $Y \neq 0$ a.s. These statements also hold for convergence in probability.

Another important mode of convergence of random variables and random elements is convergence in distribution or weak convergence. A sequence of distributions F_n on $\Re$ *converges weakly* to a distribution F, denoted by $F_n \xrightarrow{w} F$, if

$$\lim_{n\to\infty} F_n(x) = F(x),$$

for all continuity points x of F. A sequence of random variables X_n *converges in distribution* to a random variable X, denoted by $X_n \xrightarrow{d} X$, if $F_{X_n} \xrightarrow{w} F_X$, as $n \to \infty$. This notion extends to metric spaces as follows.

DEFINITION 18. A sequence of probability measures P_n on a metric space S *converge weakly* to a probability measure P, denoted by $P_n \xrightarrow{w} P$, if

$$\int_S f(x) P_n(dx) \to \int_S f(x) P(dx),$$

for any bounded continuous $f : S \to \Re$. A sequence of random elements X_n in S *converges in distribution* to a random element X in S, denoted by $X_n \xrightarrow{d} X$, if $F_{X_n} \xrightarrow{w} F_X$, or equivalently,

$$\lim_{n\to\infty} E[f(X_n)] = E[f(X)],$$

for any bounded continuous $f : S \to \Re$. Although the X_n take values in the same space S, their underlying probability spaces need not be the same, since the definition only depends on their distributions.

The preceding modes of convergence for random elements in a metric space S have the hierarchy

$$X_n \to X \text{ a.s.} \quad \Rightarrow \; X_n \xrightarrow{P} X \quad \Rightarrow \; X_n \xrightarrow{d} X.$$

Here are some equivalent ways of characterizing convergence in distribution (δB denotes the *boundary* of B).

THEOREM 8.3. (Portmanteau theorem) *For random elements $X, X_1, X_2, \ldots$ in a metric space S, the following statements are equivalent.*
(i) $X_n \xrightarrow{d} X$.
(ii) $\liminf_{n\to\infty} P\{X_n \in G\} \geq P\{X \in G\}$, *for any open set G.*
(iii) $\limsup_{n\to\infty} P\{X_n \in F\} \leq P\{X \in F\}$, *for any closed set F.*
(iv) $P\{X_n \in B\} \to P\{X \in B\}$, *for any Borel set B with $X \neq \delta B$ a.s.*

The convergence in distribution of random variables is equivalent to the convergence of their characteristic functions. Specifically, $X_n \xrightarrow{d} X$ in $\Re$ if and only if $E[e^{isX_n}] \to E[e^{isX}]$, $s \in \Re$. Similar statements hold for moment generating functions, Laplace transforms or generating functions.

The continuous-mapping property in Proposition 8.2 extends to convergence in distribution as follows.

THEOREM 8.4. (Continuous Mappings; Mann and Wald, Prohorov, Rubin) *Suppose $X_n \xrightarrow{d} X$ in a metric space S, and $B \in \mathcal{S}$ is such that $X \in B$ a.s. Let $f, f_1, f_2, \ldots$ be measurable functions from S to a metric space S'.*
(a) *If f is continuous on B, then $f(X_n) \xrightarrow{d} f(X)$.*
(b) *If $f_n(x_n) \to f(x)$, for any $x_n \to x \in B$, then $f_n(X_n) \xrightarrow{d} f(X)$.*

As an example, suppose $(X_n, Y_n) \xrightarrow{d} (X, Y)$ as $n \to \infty$. Then

$$X_n + Y_n \xrightarrow{d} X + Y, \qquad X_n Y_n \xrightarrow{d} XY.$$

This follows by the continuous mapping property for vectors since addition and multiplication are continuous functions from $\Re^d$ to $\Re$.

The next results address the following question for random variables X_n. If $X_n \to X$ a.s. (or in probability or distribution), what are the additional conditions under which $EX_n \to EX$ or $E|X_n - X| \to 0$?

THEOREM 8.5. (Fatou) *If X_n are nonnegative random variables (or are bounded from below), then* $\liminf_{n\to\infty} EX_n \geq E[\liminf_{n\to\infty} X_n]$.

THEOREM 8.6. (Monotone Convergence) *If X_n are nonnegative random variables (or are bounded from below) and $X_n \uparrow X$ a.s., then $EX_n \uparrow EX$ as $n \to \infty$, where $EX = \infty$ is possible.*

THEOREM 8.7. (Dominated Convergence) *If X_n are random variables such that $X_n \to X$ a.s., where $|X_n| \leq Y$ and $EY < \infty$, then $E|X|$ exists and $EX_n \to EX$ as $n \to \infty$.*

These results describing the convergence of $EX_n = \int_\Omega X_n(\omega) P(d\omega)$, are random-variable versions of basic theorems for Lebesgue integrals (and for summations as well). We will use the results a few times for real-valued functions f, f_n on a measurable space $(S, \mathcal{S})$ with a measure μ. For instance, the monotone convergence says that if $0 \le f_n \uparrow f$, then

$$\int_S f_n(x)\mu(dx) \to \int_S f(x)\mu(dx).$$

The next results address the convergence of EX_n when X_n converges in probability.

THEOREM 8.8. (Convergence in mean or in L^1) *Suppose $X_n \xrightarrow{P} X$ in $\Re$, and $E|X|$ and $E|X_n|$ are finite. Then the following statements are equivalent.*
(i) $E|X_n - X| \to 0$ $\quad$ *(X_n converges to X in L^1).*
(ii) $E|X_n| \to E|X|$.
(iii) *The X_n are* uniformly integrable, *meaning*

$$\sup_{n\ge 0} E[|X_n|\mathbf{1}(|X_n| \ge x)] \to 0, \ \ \text{as } x \to \infty.$$

Some convergence theorems such as the dominated convergence theorem also hold when $X_n \xrightarrow{d} X$ instead of $X_n \to X$ a.s. This is due to the following a.s. representation for convergence in distribution for random elements.

THEOREM 8.9. (Coupling; Skorohod, Dudley) *Suppose $X_n \xrightarrow{d} X$ in a Polish space S. Then there exist random elements Y_n and Y in S, defined on a single probability space, such that $Y_n \stackrel{d}{=} X_n$, $Y \stackrel{d}{=} X$, and $Y_n \to Y$ a.s.*

Loosely speaking, *coupling* is a method for comparing random elements on different probability spaces, usually to prove convergence theorems or stochastic ordering properties. For instance, suppose X_n is a random element in S_n defined on a probability space $(\Omega_n, \mathcal{F}_n, P_n)$, for $n \ge 0$. Random elements Y_n in S_n defined on a single probability space $(\Omega, \mathcal{F}, P)$ form a *coupling of* X_n if $Y_n \stackrel{d}{=} X_n$, $n \ge 0$.

Theorem 8.9 and the classical dominated convergence yield the following dominated convergence for convergence in distribution.

THEOREM 8.10. *Suppose $X_n \xrightarrow{d} X$ in $\Re$ and there exists a random variable Y with finite mean such that $X_n \stackrel{d}{\le} Y$, meaning*

$$P\{|X_n| > x\} \le P\{Y > x\}, \qquad x \ge 0.$$

Then $E|X|$ exists and $EX_n \to EX$ as $n \to \infty$.

Bibliography

[1] Ash, R. B., and C. A. Doléans-Dade (2000). *Probability and Measure Theory*, 2nd Ed. Academic Press, San Diego.

[2] Asmussen, S. (1987). *Applied Probability and Queues.* John Wiley & Sons, New York.

[3] Athreya, K, McDonald, D. and P. Ney (1972). Coupling and the renewal theorem. *Amer. Math. Monthly*, **85**, 809–814.

[4] Bauer, H. (1972). *Probability Theory and Elements of Measure Theory.* Holt, Rinehart & Winston, New York.

[5] Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley & Sons, New York.

[6] Brandt, A. and G. Last (1995). *Marked Point Processes on the Real Line: The Dynamic Approach.* Springer–Verlag, New York.

[7] Breiman, L. (1968). *Probability.* Addison-Wesley, Reading, Mass.

[8] Bremaud, P.ˇ(1999). *Markov Chains. Gibbs Fields, Monte Carlo Simulation, and Queues.* Springer, New York.

[9] Chao, X., Pinedo, M. and M. Miyazawa (1999). *Queueing Networks: Negative Customers, Signals and Product Form.* John Wiley & Sons, New York.

[10] Chung, K.L. (1974). *A Course in Probability Theory*, 2nd ed. Academic Press, New York.

[11] Çinlar, E. (1975). *Introduction to Stochastic Processes.* Prentic–Hall, Englewood Cliffs, New Jersey.

[12] Daley, D. J. and D. Vere-Jones (1988). *An Introduction to the Theory of Point Processes.* Springer–Verlag, New York.

[13] Deb, R.K. and R. Serfozo (1973). Optimal control of batch service queues. *Adv. Appl. Prob.*, **5**, 340–361.

[14] Durrett, R. (2005). *Probability: Theory and Examples*, 2nd ed. Duxbury Press, Belmont, CA.

[15] Feller, W. (1971). *An Introduction to Probability Theory and its Applications, Vol. II.* John Wiley & Sons, New York.

[16] Gross, D. and C.M. Harris (1985). *Fundamentals of Queueing Theory.* John Wiley & Sons, New York.

[17] Grimmett, G.R. and D.R. Stirzaker (2001). *Probability and Random Processes*, 3rd ed. Oxford University Press, Oxford.

[18] Hewitt, E. and K. Stromberg (1965). *Real and Abstract Analysis.* Springer-Verlag, New York.

[19] Kalashnikov, V. V. (1994). *Mathematical Methods in Queueing Theory.* Kluwer, Dordrecht.

[20] Kallenberg, O. (1983). *Random Measures.* 3rd ed. Akademie Verlag, Berlin; Academic Press, New York.

[21] Kallenberg, O. (2004). *Foundations of Modern Probability*, 2nd ed. Springer, New York.

[22] Karlin, S. and H.M. Taylor (1975). *A First Course in Stochastic Processes*, 2nd ed. Academic Press, New York.

[23] Kelly, F. P. (1979). *Reversibility and Stochastic Networks*. John Wiley & Sons, New York.

[24] Kingman, J.F.C. (1993). *Poisson Processes*. Oxford University Press, Oxford.

[25] Kulkarni, V. G. (1995). *Modeling and Analysis of Stochastic Systems*. Chapman and Hall, London.

[26] Lindvall, T. (1977). A probabatilistic proof of Blackwell's renewal theorem. *Ann. Probability*, **5**, 482–485.

[27] Lindvall, T. (1992). *Lectures on the Coupling Method*. John Wiley & Sons, New York.

[28] Resnick, S. I. (1992). *Adventures in Stochastic Processes*. Birkhäuser, Boston.

[29] Rolski, T., Schmidli, H., Schmidt, V. and J. Teugels. (1999). *Stochastic Processes for Insurance and Finance*. John Wiley & Sons, New York.

[30] Robert, P. (2003). *Stochastic Networks and Queues*. Springer-Verlag, Berlin.

[31] Ross, S. (1983). *Stochastic Processes*. John Wiley & Sons, New York.

[32] Serfozo, R. F. (1999). *Introduction to Stochastic Networks*. Springer-Verlag, New York.

[33] Shiryaev, A. N. (1995). *Probability*, 2nd ed. Spinger-Verlag, New York.

[34] Stroock, D. W. (2005). *An Introduction to Markov Processes*. Graduate Texts in Mathematics 230, Springer, Berlin.

[35] Whittle, P. (1986b). *Systems in Stochastic Equilibrium*. John Wiley & Sons, New York.

[36] Whitt, W. (2002). *Stochastic Process Limits*. Springer, New York.

[37] Wolff, R. W. (1989). *Stochastic Modelling and the Theory of Queues*. Prentice-Hall, New York.

Notation

$A(t) = t - T_{N(t)}$	Backward recurrence time
$B(t) = T_{N(t)+1} - t$	Forward recurrence time
$C_K^+(S)$	Set of continuous $f : S \to \Re_+$ with compact support
$\delta_x(A) = \mathbf{1}(x \in A)$	Dirac measure with unit mass at 1
DRI	Directly Riemann integrable
E	Expectation operator
$E[X\|Y],\ E[X\|A]$	Conditional expectations
$E[e^{sX}]$	Moment generating function of X
$\hat{F}(\alpha) = \int_{\Re_+} e^{-\alpha t} dF(t)$	Laplace transform of F
$(S, \mathcal{S})$	State space and its Borel subsets
$e_i = (0, \ldots, 0, 1, 0 \ldots, 0)$	ith unit vector
$f(t), g(t), h(t), H(t)$	Functions
$f(t) = f^+(t) - f^-(t)$	f equals its positive part minus its negative part
$F(t),\ G(t)$	Distribution functions
$F_e(x) = \frac{1}{\mu}\int_0^x [1 - F(s)]ds$	Equilibrium distribution of F
$H(t) = h(t) + F \star H(t)$	Renewal equation
$N(t)$	Counting process, or renewal process
$N(\mu, \sigma^2)$	Normal random variable with mean μ and variance σ^2
Random element X in S	X is measurable map from a probability space to S
$S,\ \mathcal{S}\ \hat{\mathcal{S}}$	Polish space and its Borel sets and bounded Borel sets
T_n	Time of nth event occurrence or nth renewal time
$U(t) = \sum_{n=0}^{\infty} F^{n\star}(t)$	Renewal function
a.s.	With probability one
$X(t),\ Y(t),\ Z(t)$	Continuous-time stochastic processes
$\xi_n = T_n - T_{n-1}$	nth inter-renewal time, or time between event occurrences

$x \vee y = \max\{x, y\}$	Maximum of x and y
$x \wedge y = \min\{x, y\}$	Minimum of x and y
$X(t) \xrightarrow{d} Y$	$X(t)$ converges in distribution to Y as $t \to \infty$
$X \stackrel{d}{=} Y$	The distributions of X and Y are equal
$x^+ = \max\{0, x\}$	Positive part of x
$x^- = -\min\{0, x\}$	Negative part of x and $x = x^+ - x^-$
$\sum_{n=1}^{N(t)}(\cdot) = 0$	When $N(t) = 0$
$\lfloor x \rfloor$	Largest integer $\leq x$, or the integer part of x
$\lceil x \rceil$	Smallest integer $\geq x$
$f(t) = o(g(t)) \quad$ as $t \to t_0$	$\lim_{t \to t_0} f(t)/g(t) = 0$
$f(t) = 0(g(t)) \quad$ as $t \to t_0$	$\limsup_{t \to t_0} \|f(t)\|/\|g(t)\| < \infty$
$a \Rightarrow b,\ a \Leftrightarrow a$	a implies b, and a is equivalent to b