Chapter 4

RENEWAL PROCESSES

4.1 Introduction

Recall that a renewal process is an arrival process in which the interarrival intervals are positive,[1] independent and identically distributed (IID) random variables (rv's). Renewal processes (since they are arrival processes) can be specified in three standard ways, first, by the joint distributions of the arrival epochs $S_1, S_2, \dots$, second, by the joint distributions of the interarrival times $X_1, X_2, \dots$, and third, by the joint distributions of the counting rv's, $N(t)$ for $t > 0$. Recall that $N(t)$ represents the number of arrivals to the system in the interval $(0, t]$.

The simplest characterization is through the interarrival times X_i, since they are IID. Each arrival epoch S_n is simply the sum $X_1 + X_2 + \cdots + X_n$ of n IID rv's. The characterization of greatest interest in this chapter is the renewal counting process, $\{N(t);\ t > 0\}$. Recall from (2.2) and (2.3) that the arrival epochs and the counting rv's are related in each of the following equivalent ways.

$$\{S_n \le t\} = \{N(t) \ge n\}; \qquad \{S_n > t\} = \{N(t) < n\}. \tag{4.1}$$

The reason for calling these processes *renewal processes* is that the process probabilistically starts over at each arrival epoch, S_n. That is, if the nth arrival occurs at $S_n = \tau$, then, counting from $S_n = \tau$, the j^{th} subsequent arrival epoch is at $S_{n+j} - S_n = X_{n+1} + \cdots + X_{n+j}$. Thus, given $S_n = \tau$, $\{N(\tau + t) - N(\tau); t \ge 0\}$ is a renewal counting process with IID interarrival intervals of the same distribution as the original renewal process. This interpretation of arrivals as renewals will be discussed in more detail later.

The major reason for studying renewal processes is that many complicated processes have randomly occurring instants at which the system returns to a state probabilistically equiva-

[1] Renewal processes are often defined in a slightly more general way, allowing the interarrival intervals X_i to include the possibility $1 > \Pr\{X_i = 0\} > 0$. All of the theorems in this chapter are valid under this more general assumption, as can be verified by complicating the proofs somewhat. Allowing $\Pr\{X_i = 0\} > 0$ allows multiple arrivals at the same instant, which makes it necessary to allow $N(0)$ to take on positive values, and appears to inhibit intuition about renewals. Exercise 4.3 shows how to view these more general renewal processes while using the definition here, thus showing that the added generality is not worth much.

lent to the starting state. These embedded renewal epochs allow us to separate the long term behavior of the process (which can be studied through renewal theory) from the behavior within each renewal period.

Example 4.1.1 (Visits to a given state for a Markov chain). Suppose a recurrent finite-state Markov chain with transition matrix $[P]$ starts in state i at time 0. Then on the first return to state i, say at time n, the Markov chain, from time n on, is a probabilistic replica of the chain starting at time 0. That is, the state at time 1 is j with probability P_{ij}, and, given a return to i at time n, the probability of state j at time $n+1$ is P_{ij}. In the same way, for any $m > 0$,

$$\Pr\{X_1 = j, \dots , X_m = k \mid X_0 = i\} = \Pr\{X_{n+1} = j, \dots , X_{n+m} = k \mid X_n = i\} . \qquad (4.2)$$

Each subsequent return to state i at a given time n starts a new probabilistic replica of the Markov chain starting in state i at time 0, . Thus the sequence of entry times to state i can be viewed as the arrival epochs of a renewal process.

This example is important, and will form the key to the analysis of Markov chains with a countably infinite set of states in Chapter 5. At the same time, (4.2) does not quite justify viewing successive returns to state i as a renewal process. The problem is that the time of the first entry to state i after time 0 is a random variable rather than a given time n. This will not be a major problem to sort out, but the resolution will be more insightful after developing some basic properties of renewal processes.

Example 4.1.2 (The G/G/m queue:). The customer arrivals to a G/G/m queue form a renewal counting process, $\{N(t); t > 0\}$. Each arriving customer waits in the queue until one of m identical servers is free to serve it. The service time required by each customer is a rv, IID over customers, and independent of arrival times and servers. The system is assumed to be empty for $t < 0$, and an arrival, viewed as customer number 0, is assumed at time 0. The subsequent interarrival intervals $X_1, X_2, \dots$, are IID. Note that $N(t)$ for each $t > 0$ is the number of arrivals in $(0, t]$, so arrival number 0 at $t = 0$ is not counted in $N(t)$.[2]

We define a new counting process, $\{N^r(t); t > 0\}$, for which the renewal epochs are those particular arrival epochs in the original process $\{N(t);\ t > 0\}$ at which an arriving customer sees an empty system (i.e., no customer in queue and none in service).[3] We will show in Section 4.5.3 that $\{N^r(t)\ t > 0\}$ is actually a renewal process, but give an intuitive explanation here. Note that customer 0 arrives at time 0 to an empty system, and given a first subsequent arrival to an empty system, at say epoch $S_1^r > 0$, the subsequent customer interarrival intervals are independent of the arrivals in $(0, S_1^r)$ and are identically distributed to those earlier arrivals. The service times after S_1^r are also IID from those earlier. Finally, the conditions that cause queueing starting from the arrival to an empty system at $t = S_1^r$ are the same as those starting from the arrival to an empty system at $t = 0$.

[2]There is always a certain amount of awkwardness in 'starting' a renewal process, and the assumption of an arrival at time 0 which is not counted in $N(t)$ seems strange, but simplifies the notation. The process is defined in terms of the IID inter-renewal intervals $X_1, X_2, \dots$. The first renewal epoch is at $S_1 = X_1$, and this is the point at which $N(t)$ changes from 0 to 1.

[3]Readers who accept without question that $\{N^r(t)\ t > 0\}$ is a renewal process should be proud of their probabilistic intuition, but should also question exactly how such a conclusion can be proven.

In most situations, we use the words *arrivals* and *renewals* interchangably, but for this type of example, the word *arrival* is used for the counting process $\{N(t);\ t > 0\}$ and the word *renewal* is used for $\{N^r(t);\ t > 0\}$. The reason for being interested in $\{N^r(t);\ t > 0\}$ is that it allows us to analyze very complicated queues such as this in two stages. First, $\{N(t);\ t > 0\}$ lets us analyze the distribution of the inter-renewal intervals X_n^r of $\{N^r(t);\ t > 0\}$. Second, the general renewal results developed in this chapter can be applied to the distribution on X_n^r to understand the overall behavior of the queueing system.

Throughout our study of renewal processes, we use $\overline{X}$ and $\mathsf{E}[X]$ interchangeably to denote the mean inter-renewal interval, and use σ_X^2 or simply σ^2 to denote the variance of the inter-renewal interval. We will usually assume that $\overline{X}$ is finite, but, except where explicitly stated, we need not assume that σ^2 is finite. This means, first, that σ^2 need not be calculated (which is often difficult if renewals are embedded into a more complex process), and second, since modeling errors on the far tails of the inter-renewal distribution typically affect σ^2 more than $\overline{X}$, the results are relatively robust to these kinds of modeling errors.

Much of this chapter will be devoted to understanding the behavior of $N(t)$ and $N(t)/t$ as t becomes large. As might appear to be intuitively obvious, and as is proven in Exercise 4.1, $N(t)$ is a rv (*i.e.*, not defective) for each $t > 0$. Also, as proven in Exercise 4.2, $\mathsf{E}[N(t)] < \infty$ for all $t > 0$. It is then also clear that $N(t)/t$, which is interpreted as the time-average renewal rate over (0,t], is also a rv with finite expectation.

One of the major results about renewal theory, which we establish shortly, concerns the behavior of the rv's $N(t)/t$ as $t \to \infty$. For each sample point $\omega \in \Omega$, $N(t,\omega)/t$ is a nonnegative number for each t and $\{N(t,\omega);\ t > 0\}$ is a sample path of the counting renewal process, taken from $(0,t]$ for each t. Thus $\lim_{t\to\infty} N(t,\omega)/t$, if it exists, is the time-average renewal rate over $(0,\infty)$ for the sample point ω.

The *strong law for renewal processes* states that this limiting time-average renewal rate exists for a set of ω that has probability 1, and that this limiting value is $1/\overline{X}$. We shall often refer to this result by the less precise statement that the time-average renewal rate is $1/\overline{X}$. This result is a direct consequence of the strong law of large numbers (SLLN) for IID rv's. In the next section, we first state and prove the SLLN for IID rv's and then establish the strong law for renewal processes.

Another important theoretical result in this chapter is the elementary renewal theorem, which states that $\mathsf{E}[N(t)/t]$ also approaches $1/\overline{X}$ as $t \to \infty$. Surprisingly, this is more than a trival consequence of the strong law for renewal processes, and we shall develop several widely useful results such as Wald's equality, in establishing this theorem.

The final major theoretical result of the chapter is Blackwell's theorem, which shows that, for appropriate values of δ, the expected number of renewals in an interval $(t, t+\delta]$ approaches $\delta/\overline{X}$ as $t \to \infty$. We shall thus interpret $1/\overline{X}$ as an ensemble-average renewal rate. This rate is the same as the above time-average renewal rate. We shall see the benefits of being able to work with both time-averages and ensemble-averages.

There are a wide range of other results, ranging from standard queueing results to results that are needed in all subsequent chapters.

4.2 The strong law of large numbers and convergence WP1

The concept of a sequence of rv's converging with probability 1 (WP1) was introduced briefly in Section 1.5.6. We discuss this type of convergence more fully here and establish some conditions under which it holds. Next the *strong law of large numbers* (SLLN) is stated for IID rv's (this is essentially the result that the partial sample averages of IID rv's converge to the mean WP1). A proof is given under the added condition that the rv's have a finite fourth moment. Finally, in the following section, we state the strong law for renewal processes and use the SLLN for IID rv's to prove it.

4.2.1 Convergence with probability 1 (WP1)

Recall that a sequence $\{Z_n;\, n \geq 1\}$ of rv's on a sample space Ω is defined to converge WP1 to a rv Z on Ω if

$$\Pr\Big\{\omega \in \Omega : \lim_{n\to\infty} Z_n(\omega) = Z(\omega)\Big\} = 1,$$

i.e., if the set of sample sequences $\{Z_n(\omega);\, n \geq 1\}$ that converge to $Z(\omega)$ has probability 1. This becomes slightly easier to understand if we define $Y_n = Z_n - Z$ for each n. The sequence $\{Y_n;\, n \geq 1\}$ then converges to 0 WP1 if and only if the sequence $\{Z_n;\, n \geq 1\}$ converges to Z WP1. Dealing only with convergence to 0 rather than to an arbitrary rv doesn't cut any steps from the following proofs, but it simplifies the notation and the concepts.

We start with a simple lemma that provides a useful condition under which convergence to 0 WP1 occurs. We shall see later how to use this lemma in an indirect way to prove the SLLN.

Lemma 4.2.1. *Let $\{Y_n; n \geq 1\}$ be a sequence of rv's, each with finite expectation. If $\sum_{n=1}^{\infty} \mathsf{E}[|Y_n|] < \infty$, then $\Pr\{\omega : \lim_{n\to\infty} Y_n(\omega) = 0\} = 1$.*

Proof: For any α, $0 < \alpha < \infty$ and any integer $m \geq 1$, the Markov inequality says that

$$\Pr\left\{\sum_{n=1}^{m} |Y_n| > \alpha\right\} \leq \frac{\mathsf{E}\left[\sum_{n=1}^{m} |Y_n|\right]}{\alpha} = \frac{\sum_{n=1}^{m} \mathsf{E}[|Y_n|]}{\alpha}. \tag{4.3}$$

Since $|Y_n|$ is non-negative, $\sum_{n=1}^{m} |Y_n| > \alpha$ implies that $\sum_{n=1}^{m+1} |Y_n| > \alpha$. Thus the left side of (4.3) is nondecreasing in m and we can go to the limit

$$\lim_{m\to\infty} \Pr\left\{\sum_{n=1}^{m} |Y_n| > \alpha\right\} \leq \frac{\sum_{n=1}^{\infty} \mathsf{E}[|Y_n|]}{\alpha}.$$

Now let $A_m = \{\omega : \sum_{n=1}^{m} |Y_n(\omega)| > \alpha\}$. As seen above, the sequence $\{A_m;\, m \geq 1\}$ is nested, $A_1 \subseteq A_2 \cdots$, so from property (1.9) of the axioms of probability,[4]

$$\lim_{m\to\infty} \Pr\left\{\sum_{n=1}^{m} |Y_n| > \alpha\right\} = \Pr\left\{\bigcup_{m=1}^{\infty} A_m\right\}$$
$$= \Pr\left\{\omega : \sum_{n=1}^{\infty} |Y_n(\omega)| > \alpha\right\}, \tag{4.4}$$

where we have used the fact that for any given ω, $\sum_{n=1}^{\infty} |Y_n(\omega)| > \alpha$ if and only if $\sum_{n=1}^{m} |Y_n(\omega)| > \alpha$ for some $m \geq 1$. Combining (4.3) with (4.4),

$$\Pr\left\{\omega : \sum_{n=1}^{\infty} |Y_n(\omega)| > \alpha\right\} \leq \frac{\sum_{n=1}^{\infty} \mathsf{E}[|Y_n|]}{\alpha}.$$

Looking at the complementary set and assuming $\alpha > \sum_{n=1}^{\infty} \mathsf{E}[|Y_n|]$,

$$\Pr\left\{\omega : \sum_{n=1}^{\infty} |Y_n(\omega)| \leq \alpha\right\} \geq 1 - \frac{\sum_{n=1}^{\infty} \mathsf{E}[|Y_n|]}{\alpha}. \tag{4.5}$$

For any ω such that $\sum_{n=1}^{\infty} |Y_n(\omega)| \leq \alpha$, we see that $\{|Y_n(\omega)|;\, n \geq 1\}$ is simply a sequence of non-negative numbers with a finite sum. Thus the individual numbers in that sequence must approach 0, *i.e.*, $\lim_{n\to\infty} |Y_n(\omega)| = 0$ for each such ω. It follows then that

$$\Pr\left\{\omega : \lim_{n\to\infty} |Y_n(\omega)| = 0\right\} \geq \Pr\left\{\omega : \sum_{n=1}^{\infty} |Y_n(\omega)| \leq \alpha\right\}.$$

Combining this with (4.5),

$$\Pr\left\{\omega : \lim_{n\to\infty} |Y_n(\omega)| = 0\right\} \geq 1 - \frac{\sum_{n=1}^{\infty} \mathsf{E}[|Y_n|]}{\alpha}.$$

This is true for all α, so $\Pr\{\omega : \lim_{n\to\infty} |Y_n| = 0\} = 1$, and thus $\Pr\{\omega : \lim_{n\to\infty} Y_n = 0\} = 1$. $\square$

It is instructive to recall Example 1.5.1, illustrated in Figure 4.1, where $\{Y_n;\, n \geq 1\}$ converges in probability but does not converge with probability one. Note that $\mathsf{E}[Y_n] = 1/(5^{j+1} - 1)$ for $n \in [5^j, 5^{j+1})$. Thus $\lim_{n\to\infty} \mathsf{E}[Y_n] = 0$, but $\sum_{n=1}^{\infty} \mathsf{E}[Y_n] = \infty$. Thus this sequence does not satisfy the conditions of the lemma. This helps explain how the conditions in the lemma exclude such sequences.

Before proceeding to the SLLN, we want to show that convergence WP1 implies convergence in probability. We give an incomplete argument here with precise versions both in Exercise 4.5 and Exercise 4.6. Exercise 4.6 has the added merit of expressing the set $\{\omega : \lim_n Y_n(\omega) = 0\}$ explicitly in terms of countable unions and intersections of simple events involving finite sets of the Y_n. This representation is valid whether or not the conditions of the lemma are satisfied and shows that this set is indeed an event.

[4]This proof probably appears to be somewhat nitpicking about limits. The reason for this is that the argument is quite abstract and it is difficult to develop the kind of intuition that ordinarily allows us to be somewhat more casual.

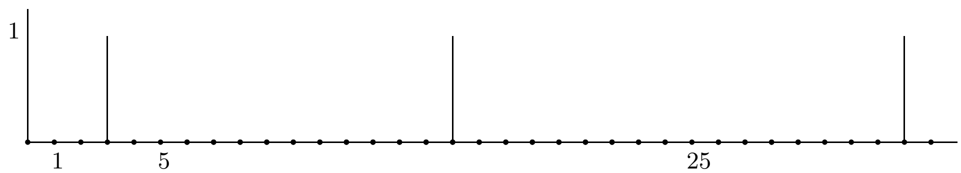


Figure 4.1: Illustration of a sample path of a sequence of rv's $\{Y_n; n \geq 0\}$ where, for each $j \geq 0$, $Y_n = 1$ for an equiprobable choice of $n \in [5^j, 5^{j+1})$ and $Y_n = 0$ otherwise.

Assume that $\{Y_n;\ n \geq 1\}$ is a sequence of rv's such that $\lim_{n\to\infty}(Y_n) = 0$ WP1. Then for any $\epsilon > 0$, each sample sequence $\{Y_n(\omega);\ n \geq 1\}$ that converges to 0 satisfies $|Y_n| \leq \epsilon$ for all sufficiently large n. This means (see Exercise 4.5) that $\lim_{n\to\infty} \Pr\{|Y_n| \leq \epsilon\} = 1$. Since this is true for all $\epsilon > 0$, $\{Y_n;\ n \geq 0\}$ converges in probability to 0.

4.2.2 Strong law of large numbers (SLLN)

We next develop the strong law of large numbers. We do not have the mathematical tools to prove the theorem in its full generality, but will give a fairly insightful proof under the additional assumption that the rv under discussion has a finite 4th moment. The theorem has a remarkably simple and elementary form, considering that it is certainly one of the most important theorems in probability theory. Most of the hard work in understanding the theorem comes from understanding what convergence WP1 means, and that has already been discussed. Given this understanding, the theorem is relatively easy to understand and surprisingly easy to prove (assuming a 4th moment).

Theorem 4.2.1 (Strong Law of Large Numbers (SLLN)). *For each integer $n \geq 1$, let $S_n = X_1 + \cdots + X_n$, where $X_1, X_2, \dots$ are IID rv's satisfying $\mathsf{E}[|X|] < \infty$. Then*

$$\Pr\left\{\omega : \lim_{n\to\infty} \frac{S_n(\omega)}{n} = \overline{X}\right\} = 1. \tag{4.6}$$

Proof (for the case where $\overline{X} = 0$ and $\mathsf{E}\left[X^4\right] < \infty$):

Assume that $\overline{X} = 0$ and $\mathsf{E}\left[X^4\right] < \infty$. Denote $\mathsf{E}\left[X^4\right]$ by γ. For any real number x, if $|x| \leq 1$, then $x^2 \leq 1$, and if $|x| > 1$, then $x^2 < x^4$. Thus $x^2 \leq 1 + x^4$ for all x. It follows $\sigma^2 = \mathsf{E}\left[X^2\right] \leq 1 + \mathsf{E}\left[X^4\right]$. Thus σ^2 is finite if $\mathsf{E}\left[X^4\right]$ is.

Now let $S_n = X_1 + \cdots + X_n$ where $X_1, \dots, X_n$ are IID with the distribution of X.

$$\begin{aligned}
\mathsf{E}\left[S_n^4\right] &= \mathsf{E}\left[(X_1 + \cdots + X_n)(X_1 + \cdots + X_n)(X_1 + \cdots + X_n)(X_1 + \cdots + X_n)\right] \\
&= \mathsf{E}\left[\left(\sum_{i=1}^n X_i\right)\left(\sum_{j=1}^n X_j\right)\left(\sum_{k=1}^n X_k\right)\left(\sum_{\ell=1}^n X_\ell\right)\right] \\
&= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{\ell=1}^n \mathsf{E}\left[X_i X_j X_k X_\ell\right],
\end{aligned}$$

where we have multiplied out the product of sums to get a sum of n^4 terms.

For each i, $1 \le i \le n$, there is a term in this sum with $i = j = k = \ell$. For each such term, $\mathsf{E}[X_i X_j X_k X_\ell] = \mathsf{E}\left[X^4\right] = \gamma$. There are n such terms (one for each choice of i, $1 \le i \le n$) and they collectively contribute $n\gamma$ to the sum $\mathsf{E}\left[S_n^4\right]$. Also, for each $i, k \neq i$, there is a term with $j = i$ and $\ell = k$. For each of these $n(n-1)$ terms, $\mathsf{E}[X_i X_i X_k X_k] = \sigma^4$. There are another $n(n-1)$ terms with $j \neq i$ and $k = i$, $\ell = j$. Each such term contributes σ^4 to the sum. Finally, for each $i \neq j$, there is a term with $\ell = i$ and $k = j$. Collectively all of these terms contribute $3n(n-1)\sigma^4$ to the sum. Each of the remaining terms is 0 since at least one of i, j, k, ℓ is different from all the others, Thus we have

$$\mathsf{E}\left[S_n^4\right] = n\gamma + 3n(n-1)\sigma^4.$$

Now consider the sequence of rv's $\{S_n^4/n^4;\ n \ge 1\}$.

$$\sum_{n=1}^{\infty} \mathsf{E}\left[\left|\frac{S_n^4}{n^4}\right|\right] \;=\; \sum_{n=1}^{\infty} \frac{n\gamma + 3n(n-1)\sigma^4}{n^4} \;<\; \infty,$$

where we have used the facts that the series $\sum_{n\ge1} 1/n^2$ and the series $\sum_{n\ge1} 1/n^3$ converge.

Using Lemma 4.2.1 applied to $\{S_n^4/n^4;\, n \ge 1\}$, we see that $\lim_{n\to\infty} S_n^4/n^4 = 0$ WP1. For each ω such that $\lim_{n\to\infty} S_n^4(\omega)/n^4 = 0$, the nonnegative fourth root of that sequence of nonnegative numbers also approaches 0. Thus $\lim_{n\to\infty} |S_n/n| = 0$ WP1. □

The above proof assumed that $\mathsf{E}[X] = 0$. It can be extended trivially to the case of an arbitrary finite $\overline{X}$ by replacing X in the proof with $X - \overline{X}$. A proof using the weaker condition that $\sigma_X^2 < \infty$ will be given in Section 7.9.1.

The technique that was used at the end of this proof provides a clue about why the concept of convergence WP1 is so powerful. The technique showed that if one sequence of rv's ($\{S_n^4/n^4;\, n \ge 1\}$) converges to 0 WP1, then another sequence ($|S_n/n|;\, n \ge 1\}$) also converges WP1. We will formalize and generalize this technique in Lemma 4.3.2 as a major step toward establishing the strong law for renewal processes.

4.3 Strong law for renewal processes

To get an intuitive idea why $N(t)/t$ should approach $1/\overline{X}$ for large t, consider Figure 4.2. For any given sample function of $\{N(t);\ t > 0\}$, note that, for any given t, $N(t)/t$ is the slope of a straight line from the origin to the point $(t, N(t))$. As t increases, this slope decreases in the interval between each adjacent pair of arrival epochs and then jumps up at the next arrival epoch. In order to express this as an equation, note that t lies between the $N(t)$th arrival (which occurs at $S_{N(t)}$) and the $(N(t)+1)$th arrival (which occurs at $S_{N(t)+1}$). Thus, for all sample points,

$$\frac{N(t)}{S_{N(t)}} \ge \frac{N(t)}{t} > \frac{N(t)}{S_{N(t)+1}}. \tag{4.7}$$

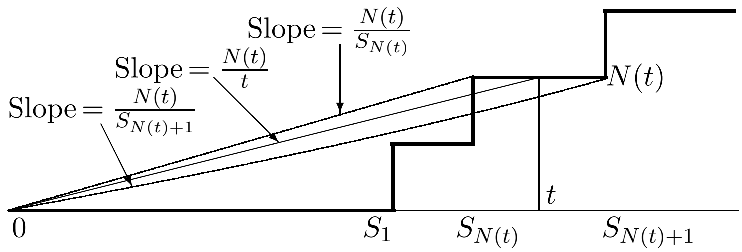


Figure 4.2: Comparison of a sample function of $N(t)/t$ with $\frac{N(t)}{S_{N(t)}}$ and $\frac{N(t)}{S_{N(t)+1}}$ for the same sample point. Note that for the given sample point, $N(t)$ is the number of arrivals up to and including t, and thus $S_{N(t)}$ is the epoch of the last arrival before or at time t. Similarly, $S_{N(t)+1}$ is the epoch of the first arrival strictly after time t.

We want to show intuitively why the slope $N(t)/t$ in the figure approaches $1/\overline{X}$ as $t \to \infty$. As t increases, we would guess that $N(t)$ increases without bound, *i.e.*, that for each arrival, another arrival occurs eventually. Assuming this, the left side of (4.7) increases with increasing t as $1/S_1, 2/S_2, \dots, n/S_n, \dots$, where $n = N(t)$. Since S_n/n converges to $\overline{X}$ WP1 from the strong law of large numbers, we might be brave enough or insightful enough to guess that n/S_n converges to $1/\overline{X}$.

We are now ready to state the strong law for renewal processes as a theorem. Before proving the theorem, we formulate the above two guesses as lemmas and prove their validity.

Theorem 4.3.1 (Strong Law for Renewal Processes). *For a renewal process with mean inter-renewal interval* $\overline{X} < \infty$, $\lim_{t\to\infty} N(t)/t = 1/\overline{X}$ *WP1.*

Lemma 4.3.1. *Let* $\{N(t);\, t > 0\}$ *be a renewal counting process with inter-renewal rv's* $\{X_n;\ n \geq 1\}$*. Then (whether or not* $\overline{X} < \infty$*),* $\lim_{t\to\infty} N(t) = \infty$ *WP1 and* $\lim_{t\to\infty} \mathsf{E}[N(t)] = \infty$*.*

Proof of Lemma 4.3.1: Note that for each sample point ω, $N(t, \omega)$ is a nondecreasing real-valued function of t and thus either has a finite limit or an infinite limit. Using (4.1), the probability that this limit is finite with value less than any given n is

$$\lim_{t\to\infty} \Pr\{N(t) < n\} \;=\; \lim_{t\to\infty} \Pr\{S_n > t\} \;=\; 1 - \lim_{t\to\infty} \Pr\{S_n \leq t\}\,.$$

Since the X_i are rv's, the sums S_n are also rv's (*i.e.*, nondefective) for each n (see Section 1.3.7), and thus $\lim_{t\to\infty} \Pr\{S_n \leq t\} = 1$ for each n. Thus $\lim_{t\to\infty} \Pr\{N(t) < n\} = 0$ for each n. This shows that the set of sample points ω for which $\lim_{t\to\infty} N(t(\omega)) < n$ has probability 0 for all n. Thus the set of sample points for which $\lim_{t\to\infty} N(t, \omega)$ is finite has probability 0 and $\lim_{t\to\infty} N(t) = \infty$ WP1.

Next, $\mathsf{E}[N(t)]$ is nondecreasing in t, and thus has either a finite or infinite limit as $t \to \infty$. For each n, $\Pr\{N(t) \geq n\} \geq 1/2$ for large enough t, and therefore $\mathsf{E}[N(t)] \geq n/2$ for such t. Thus $\mathsf{E}[N(t)]$ can have no finite limit as $t \to \infty$, and $\lim_{t\to\infty} \mathsf{E}[N(t)] = \infty$. □

The following lemma is quite a bit more general than the second guess above, but it will be useful elsewhere. This is the formalization of the technique used at the end of the proof of the SLLN.

Lemma 4.3.2. *Let* $\{Z_n;\, n \geq 1\}$ *be a sequence of rv's such that* $\lim_{n\to\infty} Z_n = \alpha$ *WP1. Let* f *be a real valued function of a real variable that is continuous at* α*. Then*

$$\lim_{n\to\infty} f(Z_n) = f(\alpha) \qquad \text{WP1.} \tag{4.8}$$

Proof of Lemma 4.3.2: First let $z_1, z_2, \dots ,$ be a sequence of real numbers such that $\lim_{n\to\infty} z_n = \alpha$. Continuity of f at α means that for every $\epsilon > 0$, there is a $\delta > 0$ such that $|f(z) - f(\alpha)| < \epsilon$ for all z such that $|z - \alpha| < \delta$. Also, since $\lim_{n\to\infty} z_n = \alpha$, we know that for every $\delta > 0$, there is an m such that $|z_n - \alpha| \leq \delta$ for all $n \geq m$. Putting these two statements together, we know that for every $\epsilon > 0$, there is an m such that $|f(z_n) - f(\alpha)| < \epsilon$ for all $n \geq m$. Thus $\lim_{n\to\infty} f(z_n) = f(\alpha)$.

If ω is any sample point such that $\lim_{n\to\infty} Z_n(\omega) = \alpha$, then $\lim_{n\to\infty} f(Z_n(\omega)) = f(\alpha)$. Since this set of sample points has probability 1, (4.8) follows. □

Proof of Theorem 4.3.1, Strong law for renewal processes: Since $\Pr\{X > 0\} = 1$ for a renewal process, we see that $\overline{X} > 0$. Choosing $f(x) = 1/x$, we see that $f(x)$ is continuous at $x = \overline{X}$. It follows from Lemma 4.3.2 that

$$\lim_{n\to\infty} \frac{n}{S_n} = \frac{1}{\overline{X}} \qquad \text{WP1.}$$

From Lemma 4.3.1, we know that $\lim_{t\to\infty} N(t) = \infty$ with probability 1, so, with probability 1, $N(t)$ increases through all the nonnegative integers as t increases from 0 to ∞. Thus

$$\lim_{t\to\infty} \frac{N(t)}{S_{N(t)}} = \lim_{n\to\infty} \frac{n}{S_n} = \frac{1}{\overline{X}} \qquad \text{WP1.}$$

Recall that $N(t)/t$ is sandwiched between $N(t)/S_{N(t)}$ and $N(t)/S_{N(t)+1}$, so we can complete the proof by showing that $\lim_{t\to\infty} N(t)/S_{N(t)+1} = 1/\overline{X}$. To show this,

$$\lim_{t\to\infty} \frac{N(t)}{S_{N(t)+1}} = \lim_{n\to\infty} \frac{n}{S_{n+1}} = \lim_{n\to\infty} \frac{n{+}1}{S_{n+1}}\,\frac{n}{n{+}1} = \frac{1}{\overline{X}} \qquad \text{WP1.}$$

□

We have gone through the proof of this theorem in great detail, since a number of the techniques are probably unfamiliar to many readers. If one reads the proof again, after becoming familiar with the details, the simplicity of the result will be quite striking. The theorem is also true if the mean inter-renewal interval is infinite; this can be seen by a truncation argument (see Exercise 4.8).

As explained in Section 4.2.1, Theorem 4.3.1 also implies the corresponding weak law of large numbers for $N(t)$, *i.e.*, for any $\epsilon > 0$, $\lim_{t\to\infty} \Pr\big\{|N(t)/t - 1/\overline{X}| \geq \epsilon\big\} = 0$). This weak law could also be derived from the weak law of large numbers for S_n (Theorem 1.5.3). We do not pursue that here, since the derivation is tedious and uninstructive. As we will see, it is the strong law that is most useful for renewal processes.

Figure 4.3 helps give some appreciation of what the strong law for $N(t)$ says and doesn't say. The strong law deals with time-averages, $\lim_{t\to\infty} N(t,\omega)/t$, for individual sample points ω;

these are indicated in the figure as horizontal averages, one for each ω. It is also of interest to look at time and ensemble-averages, $\mathsf{E}[N(t)/t]$, shown in the figure as vertical averages. Note that $N(t,\omega)/t$ is the time-average number of renewals from 0 to t, whereas $\mathsf{E}[N(t)/t]$ averages also over the ensemble. Finally, to focus on arrivals in the vicinity of a particular time t, it is of interest to look at the ensemble-average $\mathsf{E}[N(t+\delta) - N(t)]/\delta$.

Given the strong law for $N(t)$, one would hypothesize that $\mathsf{E}[N(t)/t]$ approaches $1/\overline{X}$ as $t \to \infty$. One might also hypothesize that $\lim_{t\to\infty} \mathsf{E}[N(t+\delta) - N(t)]/\delta = 1/\overline{X}$, subject to some minor restrictions on δ. These hypotheses are correct and are discussed in detail in what follows. This equality of time-averages and limiting ensemble-averages for renewal processes carries over to a large number of stochastic processes, and forms the basis of *ergodic theory*. These results are important for both theoretical and practical purposes. It is sometimes easy to find time averages (just like it was easy to find the time-average $N(t,\omega)/t$ from the strong law of large numbers), and it is sometimes easy to find limiting ensemble-averages. Being able to equate the two then allows us to alternate at will between time and ensemble-averages.

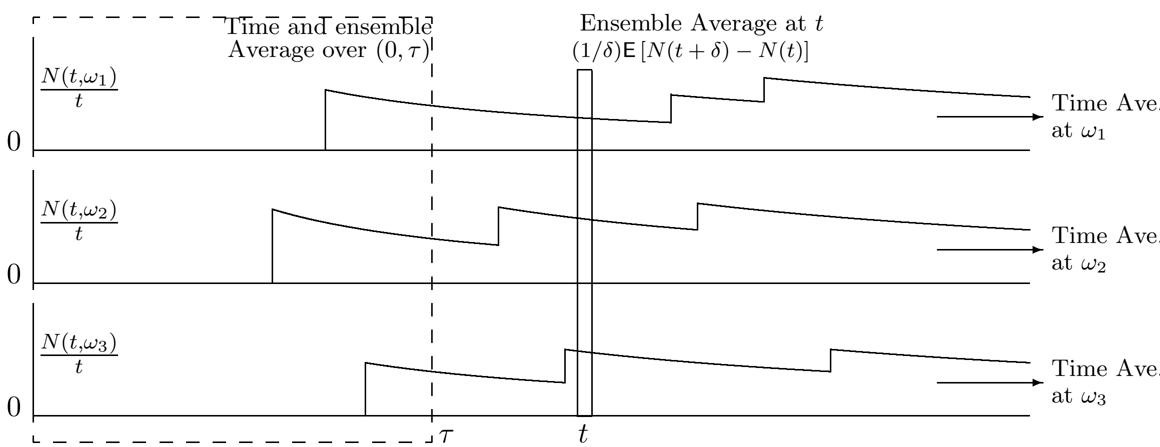


Figure 4.3: The time average at a sample point ω, the time and ensemble average from 0 to a given τ, and the ensemble-average in an interval $(t, t+\delta]$.

Note that in order to equate time-averages and limiting ensemble-averages, quite a few conditions are required. First, the time-average must exist in the limit $t \to \infty$ with probability one and also have a fixed value with probability one; second, the ensemble-average must approach a limit as $t \to \infty$; and third, the limits must be the same. The following example, for a stochastic process very different from a renewal process, shows that equality between time and ensemble averages is not always satisfied for arbitrary processes.

Example 4.3.1. Let $\{X_i; i \geq 1\}$ be a sequence of binary IID random variables, each taking the value 0 with probability 1/2 and 2 with probability 1/2. Let $\{M_n; n \geq 1\}$ be the product process in which $M_n = X_1X_2\cdots X_n$. Since $M_n = 2^n$ if X_1 to X_n each take the value 2 (an event of probability 2^{-n}) and $M_n = 0$ otherwise, we see that $\lim_{n\to\infty} M_n = 0$ with probability 1. Also $\mathsf{E}[M_n] = 1$ for all $n \geq 1$. Thus the time-average exists and equals 0 with probability 1 and the ensemble-average exists and equals 1 for all n, but the two are different. The problem is that as n increases, the atypical event in which $M_n = 2^n$ has a probability approaching 0, but still has a significant effect on the ensemble-average.

Further discussion of ensemble averages is postponed to Section 4.6. Before that, we briefly state and discuss the central limit theorem for counting renewal processes and then introduce the notion of rewards associated with renewal processes.

Theorem 4.3.2 (Central Limit Theorem (CLT) for $N(t)$). *Assume that the inter-renewal intervals for a renewal counting process $\{N(t);\, t > 0\}$ have finite standard deviation $\sigma > 0$. Then*

$$\lim_{t\to\infty} \Pr\left\{ \frac{N(t) - t/\overline{X}}{\sigma \overline{X}^{\,-3/2}\sqrt{t}} < \alpha \right\} \;=\; \Phi(\alpha). \tag{4.9}$$

where $\Phi(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)dx$.

This says that the distribution function of $N(t)$ tends to the Gaussian distribution with mean $t/\overline{X}$ and standard deviation $\sigma \overline{X}^{\,-3/2}\sqrt{t}$.

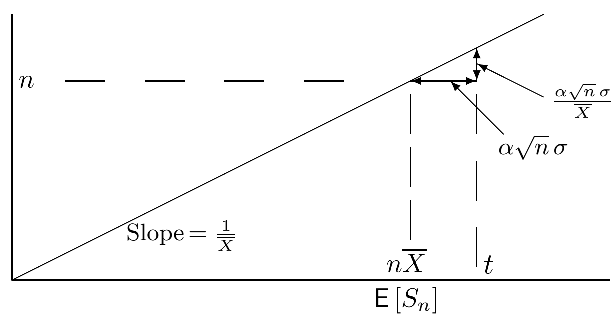


Figure 4.4: Illustration of the central limit theorem (CLT) for renewal processes. A given integer n is shown on the vertical axis, and the corresponding mean, $\mathsf{E}[S_n] = n\overline{X}$ is shown on the horizontal axis. The horizontal line with arrows at height n indicates α standard deviations from $\mathsf{E}[S_n]$, and the vertical line with arrows indicates the distance below $(t/\overline{X})$.

The theorem can be proved by applying Theorem 1.5.2 (the CLT for a sum of IID rv's) to S_n and then using the identity $\{S_n \le t\} = \{N(t) \ge n\}$. The general idea is illustrated in Figure 4.4, but the details are somewhat tedious, and can be found, for example, in [16]. We simply outline the argument here. For any real α, the CLT states that

$$\Pr\left\{S_n \le n\overline{X} + \alpha\sqrt{n}\sigma\right\} \approx \Phi(\alpha),$$

where $\Phi(\alpha) = \int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)\, dx$ and where the approximation becomes exact in the limit $n \to \infty$. Letting

$$t = n\overline{X} + \alpha\sqrt{n}\sigma,$$

and using $\{S_n \le t\} = \{N(t) \ge n\}$,

$$\Pr\{N(t) \ge n\} \approx \Phi(\alpha). \tag{4.10}$$

Since t is monotonic in n for fixed α, we can express n in terms of t, getting

$$n = \frac{t}{\overline{X}} - \frac{\alpha\sigma\sqrt{n}}{\overline{X}} \approx \frac{t}{\overline{X}} - \alpha\sigma t^{1/2}(\overline{X})^{-3/2}.$$

Substituting this into (4.10) establishes the theorem for $-\alpha$, which establishes the theorem since α is arbitrary. The omitted details involve handling the approximations carefully.

4.4 Renewal-reward processes; time-averages

There are many situations in which, along with a renewal counting process $\{N(t);\, t > 0\}$, there is another randomly varying function of time, called a *reward function* $\{R(t); t > 0\}$. $R(t)$ models a rate at which the process is accumulating a reward. We shall illustrate many examples of such processes and see that a "reward" could also be a cost or any randomly varying quantity of interest. The important restriction on these *reward functions* is that $R(t)$ at a given t depends only on the location of t within the inter-renewal interval containing t and perhaps other random variables local to that interval. Before defining this precisely, we start with several examples.

Example 4.4.1. (Time-average residual life) For a renewal counting process $\{N(t), t > 0\}$, let $Y(t)$ be the residual life at time t. The *residual life* is defined as the interval from t until the next renewal epoch, i.e., as $S_{N(t)+1} - t$. For example, if we arrive at a bus stop at time t and buses arrive according to a renewal process, $Y(t)$ is the time we have to wait for a bus to arrive (see Figure 4.5). We interpret $\{Y(t);\, t \geq 0\}$ as a reward function. The time-average of $Y(t)$, over the interval $(0, t]$, is given by[5] $(1/t)\int_0^t Y(\tau)d\tau$. We are interested in the limit of this average as $t \to \infty$ (assuming that it exists in some sense). Figure 4.5 illustrates a sample function of a renewal counting process $\{N(t);\, t > 0\}$ and shows the residual life $Y(t)$ for that sample function. Note that, for a given sample function $\{Y(t) = y(t)\}$, the integral $\int_0^t y(\tau)\, d\tau$ is simply a sum of isosceles right triangles, with part of a final triangle at the end. Thus it can be expressed as

$$\int_0^t y(\tau)d\tau = \frac{1}{2}\sum_{i=1}^{n(t)} x_i^2 + \int_{\tau=s_{n(t)}}^t y(\tau)d\tau,$$

where $\{x_i;\, 0 < i < \infty\}$ is the set of sample values for the inter-renewal intervals.

Since this relationship holds for every sample point, we see that the random variable $\int_0^t Y(\tau)d\tau$ can be expressed in terms of the inter-renewal random variables X_n as

$$\int_{\tau=0}^t Y(\tau)d\tau = \frac{1}{2}\sum_{n=1}^{N(t)} X_n^2 + \int_{\tau=S_{N(t)}}^t Y(\tau)d\tau.$$

[5] $\int_0^t Y(\tau)d\tau$ is a rv just like any other function of a set of rv's. It has a sample value for each sample function of $\{N(t);\, t > 0\}$, and its distribution function could be calculated in a straightforward but tedious way. For arbitrary stochastic processes, integration and differentiation can require great mathematical sophistication, but none of those subtleties occur here.

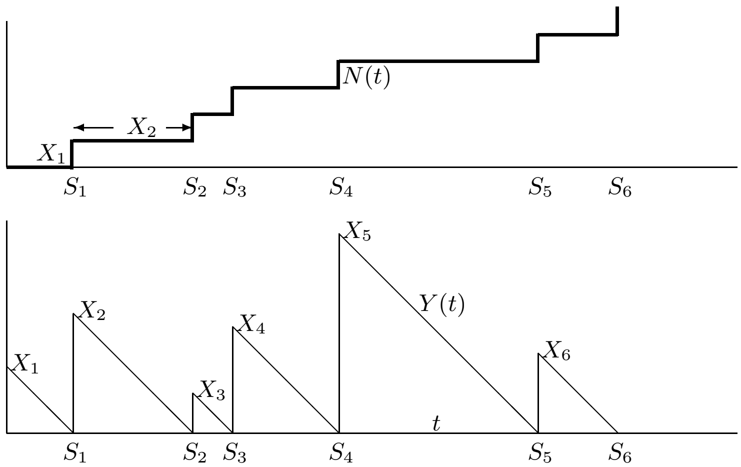


Figure 4.5: Residual life at time t. For any given sample function of the renewal process, the sample function of residual life decreases linearly with a slope of -1 from the beginning to the end of each inter-renewal interval.

Although the final term above can be easily evaluated for a given $S_{N(t)}(t)$, it is more convenient to use the following bound:

$$\frac{1}{2t}\sum_{n=1}^{N(t)} X_n^2 \leq \frac{1}{t}\int_{\tau=0}^{t} Y(\tau)d\tau \leq \frac{1}{2t}\sum_{n=1}^{N(t)+1} X_n^2. \tag{4.11}$$

The term on the left can now be evaluated in the limit $t \to \infty$ (for all sample functions except a set of probability zero) as follows:

$$\lim_{t\to\infty} \frac{\sum_{n=1}^{N(t)} X_n^2}{2t} = \lim_{t\to\infty} \frac{\sum_{n=1}^{N(t)} X_n^2}{N(t)} \; \frac{N(t)}{2t}. \tag{4.12}$$

Consider each term on the right side of (4.12) separately. For the first term, recall that $\lim_{t\to 0} N(t) = \infty$ with probability 1. Thus as $t \to \infty$, $\sum_{n=1}^{N(t)} X_n^2/N(t)$ goes through the same set of values as $\sum_{n=1}^{k} X_n^2/k$ as $k \to \infty$. Thus, using the SLLN,

$$\lim_{t\to\infty} \frac{\sum_{n=1}^{N(t)} X_n^2}{N(t)} = \lim_{k\to\infty} \frac{\sum_{n=1}^{k} X_n^2}{k} = \mathsf{E}\left[X^2\right] \qquad \text{WP1.}$$

The second term on the right side of (4.12) is simply $N(t)/2t$. By the strong law for renewal processes, $\lim_{t\to\infty} N(t)/2t = 1/(2\mathsf{E}[X])$ WP1. Thus both limits exist WP1 and

$$\lim_{t\to\infty} \frac{\sum_{n=1}^{N(t)} X_n^2}{2t} = \frac{\mathsf{E}\left[X^2\right]}{2\mathsf{E}[X]} \qquad \text{WP1.} \tag{4.13}$$

The right hand term of (4.11) is handled almost the same way:

$$\lim_{t\to\infty} \frac{\sum_{n=1}^{N(t)+1} X_n^2}{2t} = \lim_{t\to\infty} \frac{\sum_{n=1}^{N(t)+1} X_n^2}{N(t)+1} \; \frac{N(t)+1}{N(t)} \; \frac{N(t)}{2t} = \frac{\mathsf{E}\left[X^2\right]}{2\mathsf{E}[X]}. \tag{4.14}$$

Combining these two results, we see that, with probability 1, the time-average residual life is given by

$$\lim_{t\to\infty} \frac{\int_{\tau=0}^{t} Y(\tau)\, d\tau}{t} = \frac{\mathsf{E}\left[X^2\right]}{2\mathsf{E}\left[X\right]}. \tag{4.15}$$

Note that this time-average depends on the second moment of X; this is $\overline{X}^2 + \sigma^2 \geq \overline{X}^2$, so the time-average residual life is at least half the expected inter-renewal interval (which is not surprising). On the other hand, the second moment of X can be arbitrarily large (even infinite) for any given value of $\mathsf{E}[X]$, so that the time-average residual life can be arbitrarily large relative to $\mathsf{E}[X]$. This can be explained intuitively by observing that large inter-renewal intervals are weighted more heavily in this time-average than small inter-renewal intervals.

Example 4.4.2. As an example of the effect of improbable but large inter-renewal intervals, let X take on the value ϵ with probability $1-\epsilon$ and value $1/\epsilon$ with probability ϵ. Then, for small ϵ, $\mathsf{E}[X] \sim 1, \mathsf{E}\left[X^2\right] \sim 1/\epsilon$, and the time average residual life is approximately $1/(2\epsilon)$ (see Figure 4.6).

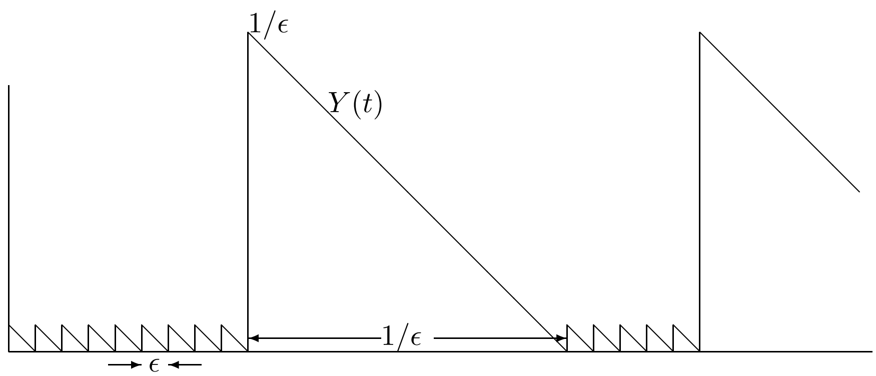


Figure 4.6: Average Residual life is dominated by large interarrival intervals. Each large interval has duration $1/\epsilon$, and the expected aggregate duration between successive large intervals is $1-\epsilon$

Example 4.4.3. (time-average Age) Let $Z(t)$ be the age of a renewal process at time t where *age* is defined as the interval from the most recent arrival before (or at) t until t, i.e., $Z(t) = t - S_{N(t)}$. By convention, if no arrivals have occurred by time t, we take the age to be t (*i.e.*, in this case, $N(t) = 0$ and we take S_0 to be 0).

As seen in Figure 4.16, the age process, for a given sample function of the renewal process, is almost the same as the residual life process—the isosceles right triangles are simply turned around. Thus the same analysis as before can be used to show that the time average of $Z(t)$ is the same as the time-average of the residual life,

$$\lim_{t\to\infty} \frac{\int_{\tau=0}^{t} Z(\tau)\, d\tau}{t} = \frac{\mathsf{E}\left[X^2\right]}{2\mathsf{E}\left[X\right]} \qquad \text{WP1.} \tag{4.16}$$

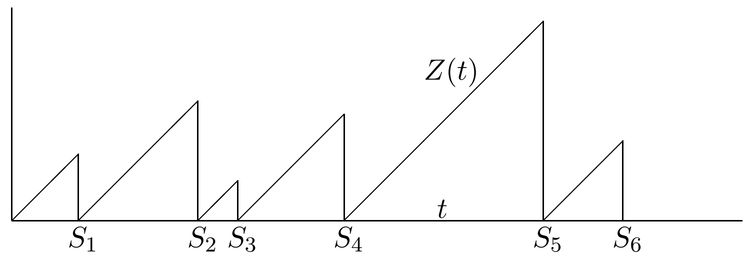


Figure 4.7: Age at time t: For any given sample function of the renewal process, the sample function of age increases linearly with a slope of 1 from the beginning to the end of each inter-renewal interval.

Example 4.4.4. (time-average Duration) Let $\widetilde{X}(t)$ be the duration of the inter-renewal interval containing time t, i.e., $\widetilde{X}(t) = X_{N(t)+1} = S_{N(t)+1} - S_{N(t)}$ (see Figure 4.8). It is clear that $\widetilde{X}(t) = Z(t) + Y(t)$, and thus the time-average of the duration is given by

$$\lim_{t\to\infty} \frac{\int_{\tau=0}^{t} \widetilde{X}(\tau)\, d\tau}{t} = \frac{\mathsf{E}\left[X^2\right]}{\mathsf{E}\left[X\right]} \qquad \text{WP1.} \tag{4.17}$$

Again, long intervals are heavily weighted in this average, so that the time-average duration is at least as large as the mean inter-renewal interval and often much larger.

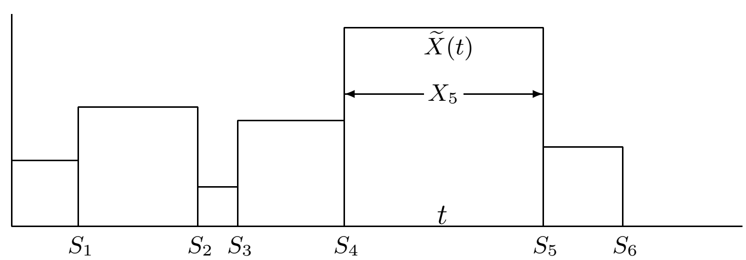


Figure 4.8: Duration $\widetilde{X}(t) = X_{N(t)}$ of the inter-renewal interval containing t.

4.4.1 General renewal-reward processes

In each of these examples, and in many other situations, we have a random function of time (i.e., $Y(t)$, $Z(t)$, or $\widetilde{X}(t)$) whose value at time t depends only on where t is in the current inter-renewal interval (i.e., on the age $Z(t)$ and the duration $\widetilde{X}(t)$ of the current inter-renewal interval). We now investigate the general class of reward functions for which the reward at time t depends at most on the age and the duration at t, *i.e.*, the reward $R(t)$ at time t is given explicitly as a function[6] $\mathcal{R}(Z(t), \widetilde{X}(t))$ of the age and duration at t.

[6]This means that $R(t)$ can be determined at any t from knowing $Z(t)$ and $X(t)$. It does not mean that $R(t)$ must vary as either of those quantities are changed. Thus, for example, $R(t)$ could depend on only one of the two or could even be a constant.

For the three examples above, the function $\mathcal{R}$ is trivial. That is, the residual life, $Y(t)$, is given by $\widetilde{X}(t) - Z(t)$ and the age and duration are given directly.

We now find the time-average value of $R(t)$, namely, $\lim_{t\to\infty} \frac{1}{t}\int_0^t R(\tau)\,d\tau$. As in examples 4.4.1 to 4.4.4 above, we first want to look at the accumulated reward over each inter-renewal period separately. Define R_n as the accumulated reward in the nth renewal interval,

$$\mathrm{R}_n = \int_{S_{n-1}}^{S_n} R(\tau)\,d(\tau) = \int_{S_{n-1}}^{S_n} \mathcal{R}[Z(\tau), \widetilde{X}(\tau)]\,d\tau. \tag{4.18}$$

For residual life (see Example 4.4.1), R_n is the area of the nth isosceles right triangle in Figure 4.5. In general, since $Z(\tau) = \tau - S_{n-1}$,

$$\mathrm{R}_n = \int_{S_{n-1}}^{S_n} \mathcal{R}(\tau - S_{n-1}, X_n)\,d\tau = \int_{z=0}^{X_n} \mathcal{R}(z, X_n)\,dz. \tag{4.19}$$

Note that R_n is a function only of X_n, where the form of the function is determined by $\mathcal{R}(Z, X)$. From this, it is clear that $\{\mathrm{R}_n; n \geq 1\}$ is essentially[7] a set of IID random variables. For residual life, $\mathcal{R}(z, X_n) = X_n - z$, so the integral in (4.19) is $X_n^2/2$, as calculated by inspection before. In general, from (4.19), the expected value of R_n is given by

$$\mathsf{E}[\mathrm{R}_n] = \int_{x=0}^{\infty}\int_{z=0}^{x} \mathcal{R}(z, x)\,dz\,d\mathsf{F}_X(x). \tag{4.20}$$

Breaking $\int_0^t R(\tau)\,d\tau$ into the reward over the successive renewal periods, we get

$$\begin{aligned}
\int_0^t R(\tau)\,d\tau &= \int_0^{S_1} R(\tau)\,d\tau + \int_{S_1}^{S_2} R(\tau)\,d\tau + \cdots + \int_{S_{N(t)-1}}^{S_{N(t)}} R(\tau)\,d\tau + \int_{S_{N(t)}}^{t} R(\tau)d\tau \\
&= \sum_{n=1}^{N(t)} \mathrm{R}_n + \int_{S_{N(t)}}^{t} R(\tau)\,d\tau. \qquad (4.21)
\end{aligned}$$

The following theorem now generalizes the results of Examples 4.4.1, 4.4.3, and 4.4.4 to general renewal-reward functions.

Theorem 4.4.1. *Let $\{R(t); t > 0\} \geq 0$ be a nonnegative renewal-reward function for a renewal process with expected inter-renewal time $\mathsf{E}[X] = \overline{X} < \infty$. If $\mathsf{E}[\mathrm{R}_n] < \infty$, then with probability 1*

$$\lim_{t\to\infty} \frac{1}{t}\int_{\tau=0}^{t} R(\tau)\,d\tau = \frac{\mathsf{E}[\mathrm{R}_n]}{\overline{X}}. \tag{4.22}$$

Proof: Using (4.21), the accumulated reward up to time t can be bounded between the accumulated reward up to the renewal before t and that to the next renewal after t,

$$\frac{\sum_{n=1}^{N(t)} \mathrm{R}_n}{t} \leq \frac{\int_{\tau=0}^{t} R(\tau)\,d\tau}{t} \leq \frac{\sum_{n=1}^{N(t)+1} \mathrm{R}_n}{t}. \tag{4.23}$$

[7]One can certainly define functions $\mathcal{R}(Z, X)$ for which the integral in (4.19) is infinite or undefined for some values of X_n, and thus R_n becomes a defective rv. It seems better to handle this type of situation when it arises rather than handling it in general.

The left hand side of (4.23) can now be broken into

$$\frac{\sum_{n=1}^{N(t)} \mathrm{R}_n}{t} = \frac{\sum_{n=1}^{N(t)} \mathrm{R}_n}{N(t)} \, \frac{N(t)}{t}. \tag{4.24}$$

Each R_n is a given function of X_n, so the R_n are IID. As $t \to \infty$, $N(t) \to \infty$, and, thus, as we have seen before, the strong law of large numbers can be used on the first term on the right side of (4.24), getting $\mathsf{E}[\mathrm{R}_n]$ with probability 1. Also the second term approaches $1/\overline{X}$ by the strong law for renewal processes. Since $0 < \overline{X} < \infty$ and $\mathsf{E}[\mathrm{R}_n]$ is finite, the product of the two terms approaches the limit $\mathsf{E}[\mathrm{R}_n]/\overline{X}$. The right-hand inequality of (4.23) is handled in almost the same way,

$$\frac{\sum_{n=1}^{N(t)+1} \mathrm{R}_n}{t} = \frac{\sum_{n=1}^{N(t)+1} \mathrm{R}_n}{N(t)+1} \, \frac{N(t)+1}{N(t)} \, \frac{N(t)}{t}. \tag{4.25}$$

It is seen that the terms on the right side of (4.25) approach limits as before and thus the term on the left approaches $\mathsf{E}[\mathrm{R}_n]/\overline{X}$ with probability 1. Since the upper and lower bound in (4.23) approach the same limit, $(1/t)\int_0^t R(\tau)\,d\tau$ approaches the same limit and the theorem is proved. □

The restriction to nonnegative renewal-reward functions in Theorem 4.4.1 is slightly artificial. The same result holds for non-positive reward functions simply by changing the directions of the inequalities in (4.23). Assuming that $\mathsf{E}[\mathrm{R}_n]$ exists (i.e., that both its positive and negative parts are finite), the same result applies in general by splitting an arbitrary reward function into a positive and negative part. This gives us the corollary:

Corollary 4.4.1. *Let $\{R(t); t > 0\}$ be a renewal-reward function for a renewal process with expected inter-renewal time $\mathsf{E}[X] = \overline{X} < \infty$. If $\mathsf{E}[\mathrm{R}_n]$ exists, then with probability 1*

$$\lim_{t\to\infty} \frac{1}{t} \int_{\tau=0}^{t} R(\tau)\,d\tau = \frac{\mathsf{E}[\mathrm{R}_n]}{\overline{X}}. \tag{4.26}$$

Example 4.4.5. (Distribution of Residual Life) Example 4.4.1 treated the time-average value of the residual life $Y(t)$. Suppose, however, that we would like to find the time-average distribution function of $Y(t)$, i.e., the fraction of time that $Y(t) \leq y$ as a function of y. The approach, which applies to a wide variety of applications, is to use an indicator function (for a given value of y) as a reward function. That is, define $R(t)$ to have the value 1 for all t such that $Y(t) \leq y$ and to have the value 0 otherwise. Figure 4.9 illustrates this function for a given sample path. Expressing this reward function in terms of $Z(t)$ and $\widetilde{X}(t)$, we have

$$R(t) = \mathcal{R}(Z(t), \widetilde{X}(t)) = \begin{cases} 1\,; & \widetilde{X}(t) - Z(t) \leq y \\ 0\,; & \text{otherwise} \end{cases}.$$

Note that if an inter-renewal interval is smaller than y (such as the third interval in Figure 4.9), then $R(t)$ has the value one over the entire interval, whereas if the interval is greater

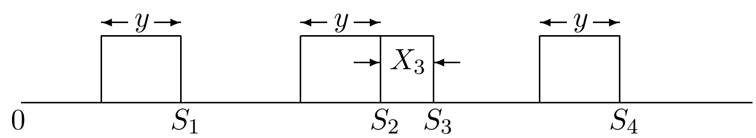


Figure 4.9: Reward function to find the time-average fraction of time that $\{Y(t) \le y\}$. For the sample function in the figure, $X_1 > y$, $X_2 > y$, and $X_4 > y$, but $X_3 < y$

than y, then $R(t)$ has the value one only over the final y units of the interval. Thus $\mathrm{R}_n = \min[y, X_n]$. Note that the random variable $\min[y, X_n]$ is equal to X_n for $X_n \le y$, and thus has the same distribution function as X_n in the range 0 to y. Figure 4.10 illustrates this in terms of the complementary distribution function. From the figure, we see that

$$\mathsf{E}[\mathrm{R}_n] = \mathsf{E}[\min(X, y)] = \int_{x=0}^{\infty} \Pr\{\min(X, y) > x\}\, dx = \int_{x=0}^{y} \Pr\{X > x\}\, dx. \qquad (4.27)$$

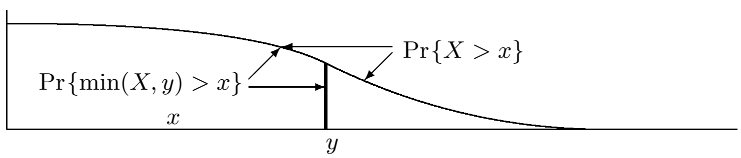


Figure 4.10: R_n for distribution of residual life.

Let $\mathsf{F}_Y(y) = \lim_{t\to\infty}(1/t)\int_0^t R(\tau)\, d\tau$ denote the time-average fraction of time that the residual life is less than or equal to y. From Theorem 4.4.1 and Eq.(4.27), we then have

$$\mathsf{F}_Y(y) = \frac{\mathsf{E}[\mathrm{R}_n]}{\overline{X}} = \frac{1}{\overline{X}} \int_{x=0}^{y} \Pr\{X > x\}\, dx \qquad \text{WP1.} \qquad (4.28)$$

As a check, note that this integral is increasing in y and approaches 1 as $y \to \infty$. Note also that the expected value of Y, calculated from (4.28), is given by $\mathsf{E}\left[X^2\right]/2\overline{X}$, in agreement with (4.15).

The same argument can be applied to the time-average distribution of age (see Exercise 4.12). The time-average fraction of time, $\mathsf{F}_Z(z)$, that the age is at most z is given by

$$\mathsf{F}_Z(z) = \frac{1}{\overline{X}} \int_{x=0}^{z} \Pr\{X > x\}\, dx \qquad \text{WP1.} \qquad (4.29)$$

In the development so far, the reward function $R(t)$ has been a function solely of the age and duration intervals, and the aggregate reward over the nth inter-renewal interval is a function only of X_n. In more general situations, where the renewal process is embedded in some more complex process, it is often desirable to define $R(t)$ to depend on other aspects of the process as well. The important thing here is for $\{\mathrm{R}_n;\ n \ge 1\}$ to be an IID sequence. How to achieve this, and how it is related to queueing systems, is described in Section 4.5.3. Theorem 4.4.1 clearly remains valid if $\{\mathrm{R}_n;\ n \ge 1\}$ is IID. This more general type of renewal-reward function will be required and further discussed in Sections 4.5.3 to Rss7 where we discuss Little's theorem and the M/G/1 expected queueing delay, both of which use this more general structure.

Limiting time-averages are sometimes visualized by the following type of experiment. For some given large time t, let T be a uniformly distributed random variable over $(0, t]$; T is independent of the renewal-reward process under consideration. Then $(1/t)\int_0^t R(\tau)\,d\tau$ is the expected value (over T) of $R(T)$ for a given sample path of $\{R(\tau);\ \tau{>}0\}$. Theorem 4.4.1 states that in the limit $t \to \infty$, all sample paths (except a set of probability 0) yield the same expected value over T. This approach of viewing a time-average as a random choice of time is referred to as *random incidence*. Random incidence is awkward mathematically, since the random variable T changes with the overall time t and has no reasonable limit. It also blurs the distinction between time and ensemble-averages, so it will not be used in what follows.

4.5 Random stopping trials

Visualize performing an experiment repeatedly, observing independent successive sample outputs of a given random variable (i.e., observing a sample outcome of $X_1, X_2, \dots$ where the X_i are IID). The experiment is stopped when enough data has been accumulated for the purposes at hand.

This type of situation occurs frequently in applications. For example, we might be required to make a choice from several hypotheses, and might repeat an experiment until the hypotheses are sufficiently discriminated. If the number of trials is allowed to depend on the outcome, the mean number of trials required to achieve a given error probability is typically a small fraction of the number of trials required when the number is chosen in advance. Another example occurs in tree searches where a path is explored until further extensions of the path appear to be unprofitable.

The first careful study of experimental situations where the number of trials depends on the data was made by the statistician Abraham Wald and led to the field of sequential analysis (see [21]). We study these situations now since one of the major results, Wald's equality, will be useful in studying $\mathsf{E}[N(t)]$ in the next section. Stopping trials are frequently useful in the study of random processes, and in particular will be used in Section 4.7 for the analysis of queues, and again in Chapter 7 as central topics in the study of Random walks and martingales.

An important part of experiments that stop after a random number of trials is the rule for stopping. Such a rule must specify, for each sample path, the trial at which the experiment stops, *i.e.*, the final trial after which no more trials are performed. Thus the rule for stopping should specify a positive, integer valued, random variable J, called the *stopping time*, or *stopping trial*, mapping sample paths to this final trial at which the experiment stops.

We view the sample space as including the set of sample value sequences for the never-ending sequence of random variables $X_1, X_2, \dots$. That is, even if the experiment is stopped at the end of the second trial, we still visualize the 3rd, 4th, ... random variables as having sample values as part of the sample function. In other words, we visualize that the experiment continues forever, but that the observer stops watching at the end of the stopping point. From the standpoint of applications, the experiment might or might not continue after the observer stops watching. From a mathematical standpoint, however, it is far preferable to view the experiment as continuing. This avoids confusion and ambiguity about the meaning of IID rv's when the very existence of later variables depends on earlier sample values.

The intuitive notion of stopping a sequential experiment should involve stopping based on the data (*i.e.*, the sample values) gathered up to and including the stopping point. For example, if $X_1, X_2, \dots$, represent the succesive changes in our fortune when gambling, we might want to stop when our cumulative gain exceeds some fixed value. The stopping trial n then depends on the sample values of $X_1, X_2, \dots, X_n$. At the same time, we want to exclude from stopping trials those rules that allow the experimenter to peek at subsequent values before making the decision to stop or not.[8] This leads to the following definition.

Definition 4.5.1. *A stopping trial (or stopping time[9]) J for a sequence of rv's $X_1, X_2, \dots,$ is a positive integer-valued rv such that for each $n \geq 1$, the indicator rv $\mathbb{I}_{\{J=n\}}$ is a function of $\{X_1, X_2, \dots, X_n\}$.*

The last clause of the definition means that any given sample value $x_1, \dots, x_n$ for $X_1, \dots, X_n$ uniquely determines whether the corresponding sample value of J is n or not. Note that since the stopping trial J is defined to be a positive integer-valued rv, the events $\{J = n\}$ and $\{J = m\}$ for $m < n$ are disjoint events, so stopping at trial m makes it impossible to also stop at n for a given sample path. Also the union of the events $\{J = n\}$ over $n \geq 1$ has probability 1. Aside from this final restriction, the definition does not depend on the probability measure and depends solely on the set of events $\{J = n\}$ for each n. In many situations, it is useful to relax the definition further to allow J to be a possibly-defective rv. In this case the question of whether stopping occurs with probability 1 can be postponed until after specifying the disjoint events $\{J = n\}$ over $n \geq 1$.

Example 4.5.1. Consider a Bernoulli process $\{X_n; n \geq 1\}$. A very simple stopping trial for this process is to stop at the first occurrence of the string $(1, 0)$. Figure 4.11 illustrates this stopping trial by viewing it as a truncation of the tree of possible binary sequences.

The event $\{J = 2\}$, *i.e.*, the event that stopping occurs at trial 2, is the event $\{X_1{=}1,\ X_2{=}0\}$. Similarly, the event $\{J = 3\}$ is $\{X_1{=}1,\ X_2{=}1,\ X_3{=}0\} \bigcup \{X_1{=}0,\ X_2{=}1,\ X_3{=}0\}$. The disjointness of $\{J = n\}$ and $\{J = m\}$ for $n \neq m$ is represented in the figure by terminating the tree at each stopping node. It can be seen that the tree never dies out completely, and in fact, for each trial n, the number of stopping nodes is $n - 1$. However, the probability that stopping has not occurred by trial n goes to zero exponentially with n, which ensures that J is a random variable.

[8]For example, poker players do not take kindly to a player who attempts to withdraw his bet when someone else wins the hand. Similarly, a statistician gathering data on product failures should not respond to a failure by then recording an earlier trial as a stopping time, thus not recording the failure.

[9]Stopping trials are more often called stopping times or optional stopping times in the literature. In our first major application of a stopping trial, however, the stopping trial is the first trial n at which a renewal epoch S_n exceeds a given time t . Viewing this *trial* as a *time* generates considerable confusion.

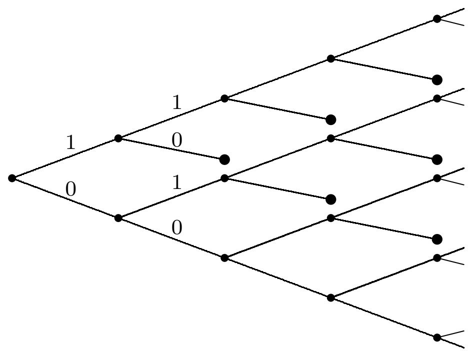


Figure 4.11: A tree representing the set of binary sequences, with a stopping rule viewed as a pruning of the tree. The particular stopping rule here is to stop on the first occurrence of the string $(1, 0)$. The leaves of the tree (*i.e.*, the nodes at which stopping occurs) are marked with large dots and the intermediate nodes (the other nodes) with small dots. Note that each leaf in the tree has a one-to-one correspondence with an initial segment of the tree, so the stopping nodes can be unambiguously viewed either as leaves of the tree or initial segments of the sample sequences.

Representing a stopping rule by a pruned tree can be used for any discrete random sequence, although the tree becomes quite unwieldy in all but trivial cases. Visualizing a stopping rule in terms of a pruned tree is useful conceptually, but stopping rules are usually stated in other terms. For example, we shortly consider a stopping trial for the interarrival intervals of a renewal process as the first n for which the arrival epoch S_n satisfies $S_n > t$ for some given $t > 0$.

4.5.1 Wald's equality

An important question that arises with stopping trials is to evaluate the sum S_J of the random variables up to the stopping trial, *i.e.*, $S_J = \sum_{n=1}^{J} X_n$. Many gambling strategies and investing strategies involve some sort of rule for when to stop, and it is important to understand the rv S_J (which can model the overall gain or loss up to that trial). Wald's equality is very useful in helping to find $\mathsf{E}[S_J]$.

Theorem 4.5.1 (Wald's equality). *Let $\{X_n; n \geq 1\}$ be a sequence of IID rv's, each of mean $\overline{X}$. If J is a stopping trial for $\{X_n; n \geq 1\}$ and if $\mathsf{E}[J] < \infty$, then the sum $S_J = X_1 + X_2 + \cdots + X_J$ at the stopping trial J satisfies*

$$\mathsf{E}[S_J] = \overline{X}\mathsf{E}[J]. \tag{4.30}$$

Proof: Note that X_n is included in $S_J = \sum_{n=1}^{J} X_n$ whenever $n \leq J$, *i.e.*, whenever the indicator function $\mathbb{I}_{\{J \geq n\}} = 1$. Thus

$$S_J = \sum_{n=1}^{\infty} X_n \mathbb{I}_{\{J \geq n\}}\ . \tag{4.31}$$

This includes X_n as part of the sum if stopping has not occurred before trial n. The event $\{J \geq n\}$ is the complement of $\{J < n\} = \{J = 1\} \bigcup \cdots \bigcup \{J = n-1\}$. All of these latter events are determined by $X_1, \dots, X_{n-1}$ and are thus independent of X_n. It follows that X_n and $\{J < n\}$ are independent and thus X_n and $\{J \geq n\}$ are also independent.[10] Thus

$$\mathsf{E}\left[X_n \mathbb{I}_{\{J \geq n\}}\right] = \overline{X}\mathsf{E}\left[\mathbb{I}_{\{J \geq n\}}\right].$$

We then have

$$\begin{aligned}
\mathsf{E}[S_J] &= \mathsf{E}\left[\sum\nolimits_{n=1}^{\infty} X_n \mathbb{I}_{\{J \geq n\}}\right] \\
&= \sum\nolimits_{n=1}^{\infty} \mathsf{E}\left[X_n \mathbb{I}_{\{J \geq n\}}\right] \qquad (4.32) \\
&= \sum\nolimits_{n=1}^{\infty} \overline{X}\mathsf{E}\left[\mathbb{I}_{\{J \geq n\}}\right] \\
&= \overline{X}\mathsf{E}[J]. \qquad (4.33)
\end{aligned}$$

The interchange of expectation and infinite sum in (4.32) is obviously valid for a finite sum, and is shown in Exercise 4.18 to be valid for an infinite sum if $\mathsf{E}[J] < \infty$. The example below shows that Wald's equality can be invalid when $\mathsf{E}[J] = \infty$. The final step above comes from the observation that $\mathsf{E}\left[\mathbb{I}_{\{J \geq n\}}\right] = \Pr\{J \geq n\}$. Since J is a positive integer rv, $\mathsf{E}[J] = \sum_{n=1}^{\infty} \Pr\{J \geq n\}$. One can also obtain the last step by using $J = \sum_{n=1}^{\infty} \mathbb{I}_{\{J \geq n\}}$ (see Exercise 4.13). □

What this result essentially says in terms of gambling is that strategies for when to stop betting are not really effective as far as the mean is concerned. This sometimes appears obvious and sometimes appears very surprising, depending on the application.

Example 4.5.2 (Stop when you're ahead in coin tossing). We can model a (biased) coin tossing game as a sequence of IID rv's $X_1, X_2, \dots$ where each X is 1 with probability p and -1 with probability $1-p$. Consider the possibly-defective stopping trial J where J is the first n for which $S_n = X_1 + \cdots + X_n = 1$, *i.e.*, the first trial at which the gambler is ahead.

We first want to see if J is a rv, *i.e.*, if the probability of eventual stopping, say $\theta = \Pr\{J < \infty\}$, is 1. We solve this by a frequently useful trick, but will use other more systematic approaches in Chapters 5 and 7 when we look at this same example as a birth-death Markov chain and then as a simple random walk. Note that $\Pr\{J = 1\} = p$, *i.e.*, $S_1 = 1$ with probability p and stopping occurs at trial 1. With probability $1-p$, $S_1 = -1$. Following $S_1 = -1$, the only way to become one ahead is to first return to $S_n = 0$ for some $n > 1$, and, after the first such return, go on to $S_m = 1$ at some later trial m. The probability of eventually going from -1 to 0 is the same as that of going from 0 to 1, *i.e.*, θ. Also, given a first return to 0 from -1, the probability of reaching 1 from 0 is θ. Thus,

$$\theta = p + (1-p)\theta^2.$$

[10] This can be quite confusing initially, since (as seen in the example of Figure 4.11) X_n is not necessarily independent of the event $\{J = n\}$, nor of $\{J = n+1\}$, etc. In other words, *given that* stopping has not occurred before trial n, then X_n can have a great deal to do with *the trial* at which stopping occurs. However, as shown above, X_n has nothing to do with *whether* $\{J < n\}$ or $\{J \geq n\}$.

This is a quadratic equation in θ with two solutions, $\theta = 1$ and $\theta = p/(1-p)$. For $p > 1/2$, the second solution is impossible since θ is a probability. Thus we conclude that J is a rv. For $p = 1/2$ (and this is the most interesting case), both solutions are the same, $\theta = 1$, and again J is a rv. For $p < 1/2$, the correct solution[11] is $\theta = p/(1-p)$. Thus $\theta < 1$ so J is a defective rv.

For the cases where $p \geq 1/2$, *i.e.*, where J is a rv, we can use the same trick to evaluate $\mathsf{E}[J]$,

$$\mathsf{E}[J] = p + (1-p)(1 + 2\mathsf{E}[J]).$$

The solution to this is

$$\mathsf{E}[J] \;=\; \frac{1}{2(1-p)} \;=\; \frac{1}{2p-1}.$$

We see that $\mathsf{E}[J]$ is finite for $p > 1/2$ and infinite for $p = 1/2$.

For $p > 1/2$, we can check that these results agree with Wald's equality. In particular, since S_J is 1 with probability 1, we also have $\mathsf{E}[S_J] = 1$. Since $\overline{X} = 2p-1$ and $\mathsf{E}[J] = 1/(2p-1)$, Wald's equality is satisfied (which of course it has to be).

For $p = 1/2$, we still have $S_J = 1$ with probability 1 and thus $\mathsf{E}[S_J] = 1$. However $\overline{X} = 0$ so $\overline{X}\mathsf{E}[J]$ has no meaning and Wald's equality breaks down. Thus we see that the restriction $\mathsf{E}[J] < \infty$ in Wald's equality is indeed needed. These results are tabulated below.

	$p > \frac{1}{2}$	$p = \frac{1}{2}$	$p < \frac{1}{2}$
$\Pr\{J < \infty\}$	$\frac{p}{1-p}$	1	1
$\mathsf{E}[J]$	∞	∞	$\frac{1}{2p-1}$

It is surprising that with $p = 1/2$, the gambler can eventually become one ahead with probability 1. This has little practical value, first because the required expected number of trials is infinite, and second (as will be seen later) because the gambler must risk a potentially infinite capital.

4.5.2 Applying Wald's equality to $m(t) = \mathsf{E}[N(t)]$

Let $\{S_n;\, n \geq 1\}$ be the arrival epochs and $\{X_n;\, n \geq 1\}$ the interarrival intervals for a renewal process. For any given $t > 0$, let J be the trial n for which S_n first exceeds t. Note that n is specified by the sample values of $\{X_1, \dots, X_n\}$ and thus J is a possibly-defective stopping trial for $\{X_n;\, n \geq 1\}$.

Since n is the first trial for which $S_n > t$, we see that $S_{n-1} \leq t$ and $S_n > t$. Thus $N(t)$ is $n-1$ and n is the sample value of $N(t)+1$. Since this is true for all sample sequences, $J = N(t) + 1$. Since $N(t)$ is a non-defective rv, J is also, so J is a stopping trial for $\{X_n;\, n \geq 1\}$.

[11] This will be shown when we view this example as a birth-death Markov chain in Chapter 5.

We can then employ Wald's equality to obtain

$$\mathsf{E}\left[S_{N(t)+1}\right] = \overline{X}\mathsf{E}\left[N(t)+1\right] = \overline{X}[m(t)+1]. \tag{4.34}$$

$$m(t) = \frac{\mathsf{E}\left[S_{N(t)+1}\right]}{\overline{X}} - 1. \tag{4.35}$$

As is often the case with Wald's equality, this provides a relationship between two quantities, $m(t)$ and $\mathsf{E}\left[S_{N(t)+1}\right]$, that are both unknown. This will be used in proving the elementary renewal theorem by upper and lower bounding $\mathsf{E}\left[S_{N(t)+1}\right]$. The lower bound is easy, since $\mathsf{E}\left[S_{N(t)+1}\right] \geq t$, and thus $m(t) \geq t/\overline{X} - 1$. It follows that

$$\frac{m(t)}{t} \geq \frac{1}{\overline{X}} - \frac{1}{t}. \tag{4.36}$$

We derive an upper bound on $\mathsf{E}\left[S_{N(t)+1}\right]$ in the next section. First, however, as a santity check, consider Figure 4.12 which illustrates (4.35) for the case where each X_n is a deterministic rv where $X_n = \overline{X}$ with probability 1.

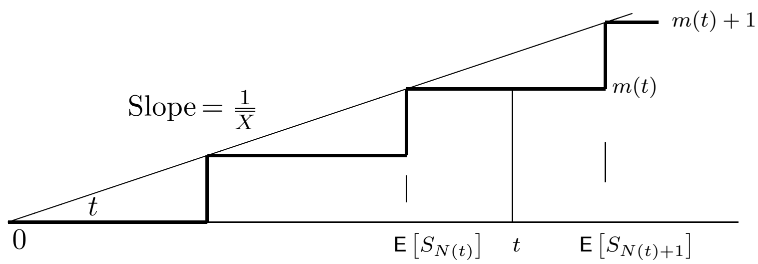


Figure 4.12: Illustration of (4.34) for the special case where X is deterministic. Note that $m(t)$, as a function of t, is then the illustrated staircase function. On each increment of t by $\overline{X}$, $m(t)$ increases by one. Then $m(t)+1$ and $\mathsf{E}\left[S_{N(t)+1}\right]$ are two sides of a right triangle of slope $1/\overline{X}$, yielding (4.34).

It might be puzzling why we used $N(t)+1$ rather than $N(t)$ as a stopping trial for the epochs $\{S_i;\, i \geq 1\}$ in this application of Wald's equality. To understand this, assume for example that $N(t) = n$. When an observer sees the sample values of $S_1, \ldots, S_n$, with $S_n < t$, the observer typically cannot tell (on the basis of $S_1, \ldots, S_n$ alone) whether any other arrivals will occur in the interval $(S_n, t]$. In other words, $N(t) = n$ implies that $S_n \leq t$, but $S_n < t$ does not imply that $N(t) = n$. On the other hand, still assuming $N(t) = n$, an observer seeing $S_1, \ldots, S_{n+1}$ knows that $N(t) = n$.

Any stopping trial (for an arbitrary sequence of rv's $\{S_n;\, n \geq 1\}$) can be viewed as an experiment where an observer views sample values $s_1, s_2, \ldots$, in turn until the stopping rule is satisfied. The stopping rule does not permit either looking ahead or going back to change an earlier decision. Thus the rule: stop at $N(t)+1$ (for the renewal epochs $\{S_n;\, n \geq 1\}$) means stop at the first sample value s_i that exceeds t. Stopping at the final sample value $s_n \leq t$ is not necessarily possible without looking ahead to the following sample value.

4.5.3 Stopping trials, embedded renewals, and G/G/1 queues

The above definition of a stopping trial is quite restrictive in that it refers only to a single sequence of rv's. In many queueing situations, for example, there is both a sequence of interarrival times $\{X_i;\, i \geq 1\}$ and a sequence of service times $\{V_i;\, i \geq 0\}$. Here X_i is the interarrival interval between customer $i-1$ and i, where an initial customer 0 is assumed to arrive at time 0, and X_1 is the arrival time for customer 1. The service time of customer 0 is then V_0 and each V_i, $i > 0$ is the service time of the corresponding ordinary customer. Customer number 0 is not ordinary in the sense that it arrives at the fixed time 0 and is not counted in the arrival counting process $\{N(t);\, t > 0\}$.

Example 4.5.3 (G/G/1 queues:). Consider a G/G/1 queue (the single server case of the G/G/m queue described in Example 4.1.2). We assume that the customers are served in First-Come-First-Served (FCFS) order.[12] Both the interarrival intervals $\{X_i;\, i \geq 1\}$ and the service times $\{V_i;\, i \geq 0\}$ are assumed to be IID and the service times are assumed to be independent of the interarrival intervals. Figure 4.13 illustrates a sample path for these arrivals and departures.

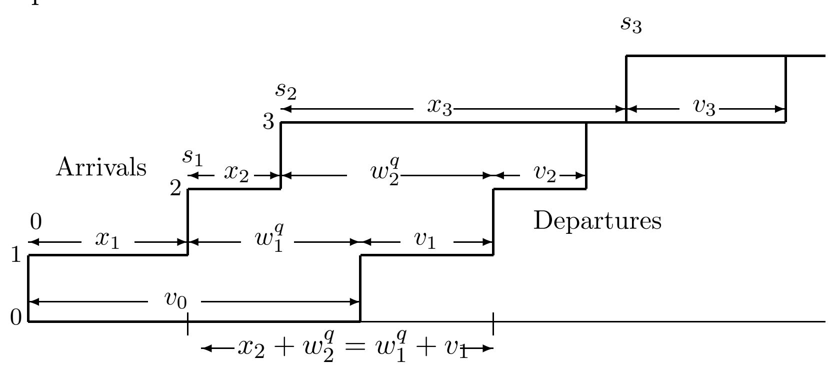


Figure 4.13: Sample path of arrivals and departures from a G/G/1 queue. Customer 0 arrives at time 0 and enters service immediately. Customer 1 arrives at time $s_1 = x_1$. For the case shown above, customer 0 has not yet departed, *i.e.*, $x_1 < v_0$, so customer 1 is queued for the interval $w_1^q = v_0 - x_1$ before entering service. As illustrated, customer 1's system time (queueing time plus service time) is $w_1 = w_1^q + v_1$. Note that the sample path of arrivals in the figure is one plus the sample path of the arrival counting process $\{N(t);\, t > 0\}$, since the counting process, by convention, does not count the initial arrival at time 0.

Customer 2 arrives at $s_2 = x_1 + x_2$. For the case shown above, this is before customer 1 departs at $v_0 + v_1$. Thus, customer 2's wait in queue is $w_2^q = v_0 + v_1 - x_1 - x_2$. As illustrated above, $x_2 + w_2^q$ is also equal to customer 1's system time, so $w_2^q = w_1^q + v_1 - x_2$. Customer 3 arrives when the system is empty, so it enters service immediately with no wait in queue, *i.e.*, $w_3^q = 0$.

The figure illustrates a sample path for which $X_1 < V_0$, so arrival number 1 waits in queue for $W_1^q = V_0 - X_1$. If $X_1 \geq V_0$, on the other hand, then customer one enters service immediately, *i.e.*, customer one 'sees an empty system.' In general, then $W_1^q = \max(V_0 - X_1,\, 0)$. In the same way, as illustrated in the figure, if $W_1^q > 0$, then customer 2

[12]For single server queues, this is sometimes referred to as First-In-First-Out (FIFO) service.

waits for $W_1^q + V_1 - X_2$ if positive and 0 otherwise. This same formula works if $W_1^q = 0$, so $W_2^q = \max(W_1^q + V_1 - X_2,\ 0)$. In general, it can be seen that

$$W_i^q = \max(W_{i-1}^q + V_{i-1} - X_i,\ 0). \tag{4.37}$$

This equation will be analyzed further in Section 7.2 where we are interested in queueing delay and system delay. Here our objectives are simpler, since we only want to show that the subsequence of customer arrivals i for which the event $\{W_i^q = 0\}$ is satisfied form the renewal epochs of a renewal process. To do this, first observe from (4.37) (using induction if desired) that W_i^q is a function of $(X_1, \dots, X_i)$ and $(V_0, \dots, V_{i-1})$. Thus, if we let J be the smallest $i > 0$ for which $W_i^q = 0$, then $\mathbb{I}_{J=i}$ is a function of $(X_1, \dots, X_i)$ and $(V_0, \dots, V_{i-1})$.

We now interrupt the discussion of G/G/1 queues with the following generalization of the definition of a stopping trial.

Definition 4.5.2 (Generalized stopping trials). *A generalized stopping trial J for a sequence of pairs of rv's $(X_1, V_1), (X_2, V_2)\dots,$ is a positive integer-valued rv such that, for each $n \geq 1$, the indicator rv $\mathbb{I}_{\{J=n\}}$ is a function of $X_1, V_1, X_2, V_2, \dots, X_n, V_n$.*

Wald's equality can be trivially generalized for these generalized stopping trials.

Theorem 4.5.2 (Generalized Wald's equality). *Let $\{(X_n, V_n);\ n \geq 1\}$ be a sequence of pairs of rv's, where each pair is independent and identically distributed (IID) to all other pairs. Assume that each X_i has finite mean $\overline{X}$. If J is a stopping trial for $\{(X_n, V_n);\ n \geq 1\}$ and if $\mathsf{E}[J] < \infty$, then the sum $S_J = X_1 + X_2 + \cdots + X_J$ satisfies*

$$\mathsf{E}[S_J] = \overline{X}\mathsf{E}[J]. \tag{4.38}$$

The proof of this will be omitted, since it is the same as the proof of Theorem 4.5.1. In fact, the definition of stopping trials could be further generalized by replacing the rv's V_i by vector rv's or by a random number of rv's , and Wald's equality would still hold.[13]

For the example of the G/G/1 queue, we take the sequence of pairs to be $\{(X_1, V_0), (X_2, V_1), \dots, \}$. Then $\{(X_n, V_{n-1};\ n \geq 1\}$ satisfies the conditions of Theorem 4.5.2 (assuming that $\mathsf{E}[X_i] < \infty$). Let J be the generalized stopping rule specifying the number of the first arrival to find an empty queue. Then the theorem relates $\mathsf{E}[S_J]$, the expected time $t > 0$ until the first arrival to see an empty queue, and $\mathsf{E}[J]$, the expected number of arrivals until seeing an empty queue.

It is important here, as in many applications, to avoid the confusion created by viewing J as a stopping *time*. We have seen that J is the *number* of the first customer to see an empty queue, and S_J is the *time* until that customer arrives.

There is a further possible timing confusion about whether a customer's service time is determined when the customer arrives or when it completes service. This makes no difference, since the ordered sequence of pairs is well-defined and satisfies the appropriate IID condition for using the Wald equality.

[13]In fact, J is sometimes defined to be a stopping rule if $\mathbb{I}_{\{J \geq n\}}$ is independent of $X_n, X_{n+1}, \dots$ for each n. This makes it easy to prove Wald's equality, but quite hard to see when the definition holds, especially since $\mathbb{I}_{\{J=n\}}$, for example, is typically dependent on X_n (see footnote 7).

As is often the case with Wald's equality, it is not obvious how to compute either quantity in (4.38), but it is nice to know that they are so simply related. It is also interesting to see that, although successive pairs (X_i, V_i) are assumed independent, it is not necessary for X_i and V_i to be independent. This lack of independence does not occur for the G/G/1 (or G/G/m) queue, but can be useful in situations such as packet networks where the interarrival time between two packets at a given node can depend on the service time (the length) of the first packet if both packets are coming from the same node.

Perhaps a more important aspect of viewing the first renewal for the G/G/1 queue as a stopping trial is the ability to show that successive renewals are in fact IID. Let $X_{2,1}, X_{2,2}, \dots$ be the interarrival times following J, the first arrival to see an empty queue. Conditioning on $J = j$, we have $X_{2,1}{=}X_{j+1},\ X_{2,2}{=}X_{j+2}, \dots,$. Thus $\{X_{2,k};\ k \geq 1\}$ is an IID sequence with the original interarrival distribution. Similarly $\{(X_{2,k}, V_{2,k});\ k \geq 1\}$ is a sequence of IID pairs with the original distribution. This is valid for all sample values j of the stopping trial J. Thus $\{(X_{2,k}, V_{2,k});\ k \geq 1\}$ is statistically independent of J and $(X_i, V_i);\ 1 \leq i \leq J$.

The argument above can be repeated for subsequent arrivals to an empty system, so we have shown that successive arrivals to an empty system actually form a renewal process.[14]

One can define many different stopping rules for queues, such as the first trial at which a given number of customers are in the queue. Wald's equality can be applied to any such stopping rule, but much more is required for the stopping trial to also form a renewal point. At the first time when n customers are in the system, the subsequent departure times depend partly on the old service times and partly on the new arrival and service times, so the required independence for a renewal point does not exist. Stopping rules are helpful in understanding embedded renewal points, but are by no means equivalent to embedded renewal points.

Finally, nothing in the argument above for the G/G/1 queue made any use of the FCFS service discipline. One can use any service discipline for which the choice of which customer to serve at a given time t is based solely on the arrival and service times of customers in the system by time t. In fact, if the server is never idle when customers are in the system, the renewal epochs will not depend on the service descipline. It is also possible to extend these arguments to the G/G/m queue, although the service discipline can affect the renewal points in this case.

4.5.4 Little's theorem

Little's theorem is an important queueing result stating that the expected number of customers in a queueing system is equal to the product of the arrival rate and the expected time each customer waits in the system. This result is true under very general conditions; we use the G/G/1 queue with FCFS service as a specific example, but the reason for the greater generality will be clear as we proceed. Note that the theorem does not tell us how

[14]Confession by author: For about 15 years, I mistakenly believed that it was obvious that arrivals to an empty system in a G/G/m queue form a renewal process. Thus I can not expect readers to be excited about the above proof. However, it is a nice example of how to use stopping times to see otherwise murky points clearly.

to find either the expected number or expected wait; it only says that if one can be found, the other can also be found.

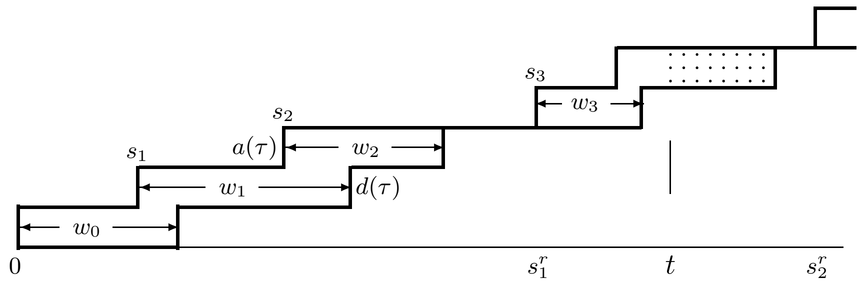


Figure 4.14: Sample path of arrivals, departures, and system waiting times for a G/G/1 queue with FCFS service. The upper step function is the number of customer arrivals, including the customer at time 0 and is denoted $a(\tau)$. Thus $a(\tau)$ is a sample path of $A(\tau) = N(\tau) + 1$, *i.e.*, the arrival counting process incremented by 1 for the initial arrival at $\tau = 0$. The lower step function, $d(\tau)$ is a sample path for $D(\tau)$, which is the number of departures (including customer 0) up to time τ. For each $i \geq 0$, w_i is the sample value of the system waiting time W_i for customer i. Note that $W_i = W_i^q + V_i$.

The figure also shows the sample values s_1^r and s_2^r of the first two arrivals that see an empty system (recall from Section 4.5.3 that the subsequence of arrivals that see an empty system forms a renewal process.)

Figure 4.14 illustrates a sample path for a G/G/1 queue with FCFS service. It illustrates a sample path $a(t)$ for the arrival process $A(t) = N(t) + 1$, *i.e.*, the number of customer arrivals in $[0, t]$, specifically including customer number 0 arriving at $t = 0$. Similarly, it illustrates the departure process $D(t)$, which is the number of departures up to time t, again including customer 0. The difference, $L(t) = A(t) - D(t)$, is then the number in the system at time t.

Recall from Section 4.5.3 that the subsequence of customer arrivals for $t > 0$ that see an empty system form a renewal process. Actually, we showed a little more than that. Not only are the inter-renewal intervals, $X_i^r = S_i^r - S_{i-1}^r$ IID, but the number of customer arrivals in each inter-renewal interval are IID, and the interarrival intervals and service times between inter-renewal intervals are IID. The sample values, s_1^r and s_2^r of the first two renewal epochs are shown in the figure.

The essence of Little's theorem can be seen by observing that $\int_0^{S_1^r} L(\tau)\, d\tau$ in the figure is the area between the upper and lower step functions, integrated out to the first time that the two step functions become equal (*i.e.*, the system becomes empty). For the sample value in the figure, this integral is equal to $w_0 + w_1 + w_2$. In terms of the rv's,

$$\int_0^{S_1^r} L(\tau)\, d\tau = \sum_{i=0}^{N(S_1^r)-1} W_i. \tag{4.39}$$

The same relationship exists in each inter-renewal interval, and in particular we can define

L_n for each $n \geq 1$ as

$$L_n = \int_{S^r_{n-1}}^{S^r_n} L(\tau)\, d\tau \;\; = \sum_{i=N(S^r_{n-1})}^{N(S^r_n)-1} W_i. \qquad (4.40)$$

The interpretation of this is far simpler than the notation. The arrival step function and the departure step function in Figure 4.14 are separated whenever there are customers in the system (the system is busy) and are equal whenever the system is empty. Renewals occur when the system goes from empty to busy, so the nth renewal is at the beginning of the nth busy period. Then L_n is the area of the region between the two step functions over the nth busy period. By simple geometry, this area is also the sum of the customer waiting times over that busy period. Finally, since the interarrival intervals and service times in each busy period are IID with respect to those in each other busy period, the sequence $L_1, L_2, \dots$, is a sequence of IID rv's.

The function $L(\tau)$ has the same behavior as a renewal reward function, but it is slightly more general, being a function of more than the age and duration of the renewal counting process $\{N^r(t); t > 0\}$ at $t = \tau$. However the fact that $\{L_n; n \geq 1\}$ is an IID sequence lets us use the same methodology to treat $L(\tau)$ as was used earlier to treat renewal-reward functions. We now state and prove Little's theorem. The proof is almost the same as that of Theorem 4.4.1, so we will not dwell on it.

Theorem 4.5.3 (Little). *For a FCFS G/G/1 queue in which the expected inter-renewal interval is finite, the limiting time-average number of customers in the system is equal, with probability* 1, *to a constant denoted as* $\overline{L}$. *The sample-path-average waiting time per customer is also equal, with probability 1, to a constant denoted as* $\overline{W}$. *Finally* $\overline{L} = \lambda\overline{W}$ *where* λ *is the customer arrival rate,* i.e., *the reciprocal of the expected interarrival time.*

Proof: Note that for any $t > 0$, $\int_0^t (L(\tau)\, d\tau$ can be expressed as the sum over the busy periods completed before t plus a residual term involving the busy period including t. The residual term can be upper bounded by the integral over that complete busy period. Using this with (4.40), we have

$$\sum_{n=1}^{N^r(t)} \mathrm{L}_n \leq \int_{\tau=0}^{t} L(\tau)\, d\tau \leq \sum_{i=0}^{N(t)} W_i \leq \sum_{n=1}^{N^r(t)+1} \mathrm{L}_n. \qquad (4.41)$$

Assuming that the expected inter-renewal interval, $\mathsf{E}[X^r]$, is finite, we can divide both sides of (4.41) by t and go to the limit $t \to \infty$. From the same argument as in Theorem 4.4.1,

$$\lim_{t\to\infty} \frac{\sum_{i=0}^{N(t)} W_i}{t} = \lim_{t\to\infty} \frac{\int_{\tau=0}^{t} L(\tau)\, d\tau}{t} = \frac{\mathsf{E}[\mathrm{L}_n]}{\mathsf{E}[X^r]} \qquad \text{with probability 1.} \qquad (4.42)$$

The equality on the right shows that the limiting time average of $L(\tau)$ exists with probability 1 and is equal to $\overline{L} = \mathsf{E}[L_n] / \mathsf{E}[X^r]$. The quantity on the left of (4.42) can now be broken up as waiting time per customer multiplied by number of customers per unit time, i.e.,

$$\lim_{t\to\infty} \frac{\sum_{i=0}^{N(t)} W_i}{t} = \lim_{t\to\infty} \frac{\sum_{i=0}^{N(t)} W_i}{N(t)} \lim_{t\to\infty} \frac{N(t)}{t}. \qquad (4.43)$$

From (4.42), the limit on the left side of (4.43) exists (and equals $\overline{L}$) with probability 1. The second limit on the right also exists with probability 1 by the strong law for renewal processes, applied to $\{N(t); t > 0\}$. This limit is called the *arrival rate* λ, and is equal to the reciprocal of the mean interarrival interval for $\{N(t)\}$. Since these two limits exist with probability 1, the first limit on the right, which is the sample-path-average waiting time per customer, denoted $\overline{W}$, also exists with probability 1. □

Reviewing this proof and the development of the G/G/1 queue before the theorem, we see that there was a simple idea, expressed by (4.39), combined with a lot of notational complexity due to the fact that we were dealing with both an arrival counting process $\{N(t);\ t > 0\}$ and an embedded renewal counting process $\{N^r(t);\ t > 0\}$. The difficult thing, mathematically, was showing that $\{N^r(t);\ t > 0\}$ is actually a renewal process and showing that the L_n are IID, and this was where we needed to understand stopping rules.

Recall that we assumed earlier that customers departed from the queue in the same order in which they arrived. From Figure 4.15, however, it is clear that FCFS order is not required for the argument. Thus the theorem generalizes to systems with multiple servers and arbitrary service disciplines in which customers do not follow FCFS order. In fact, all that the argument requires is that the system has renewals (which are IID by definition of a renewal) and that the inter-renewal interval is finite with probability 1.

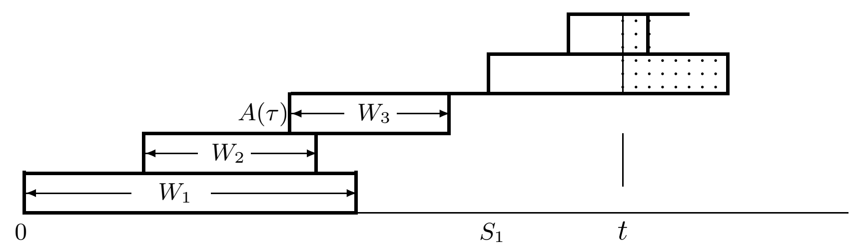


Figure 4.15: Arrivals and departures in a non-FCFS systems. The server, for example, could work simultaneously (at a reduced rate) on all customers in the system, and thus complete service for customers with small service needs before completing earlier arrivals with greater service needs. Note that the jagged right edge of the diagram does not represent number of deparatures, but this is not essential for the argument.

For example, if higher priority is given to customers with small service times, then it is not hard to see that the average number of customers in the system and the average waiting time per customer will be decreased. However, if the server is always busy when there is work to be done, it can be seen that the renewal times are unaffected. Service discipines will be discussed further in Section 5.6.

The same argument as in Little's theorem can be used to relate the average number of customers in a single server queue (not counting service) to the average wait in the queue (not counting service). Renewals still occur on arrivals to an empty system, and the integral of customers in queue over a busy period is still equal to the sum of the queue waiting times. Let $L^q(t)$ be the number in the queue at time t and let $\overline{L}^q = \lim_{t\to\infty}(1/t)\int_0^t L^q(\tau)d\tau$ be the time-average queue wait. Letting $\overline{W}^q$ be the sample-path-average waiting time in queue,

$$\overline{L}^q = \lambda \overline{W}^q. \tag{4.44}$$

The same argument can also be applied to the service facility of a single server queue. The time-average of the number of customers in the server is just the fraction of time that the server is busy. Denoting this fraction by ρ and the expected service time by $\overline{V}$, we get

$$\rho = \lambda \overline{V}. \tag{4.45}$$

4.5.5 Expected queueing time for an M/G/1 queue

For our last example of the use of renewal-reward processes, we consider the expected queueing time in an M/G/1 queue. We again assume that an arrival to an empty system occurs at time 0 and renewals occur on subsequent arrivals to an empty system. At any given time t, let $L^q(t)$ be the number of customers in the queue (not counting the customer in service, if any) and let $R(t)$ be the residual life of the customer in service. If no customer is in service, $R(t) = 0$, and otherwise $R(t)$ is the remaining time until the current service is completed. Let $U(t)$ be the waiting time in queue that would be experienced by a customer arriving at time t. This is often called the unfinished work in the queueing literature and represents the delay until all the customers currently in the system complete service. Thus the rv $U(t)$ is equal to $R(t)$, the residual life of the customer in service, plus the service times of each of the $L^q(t)$ customers currently waiting in the queue.

$$U(t) = \sum_{i=1}^{L^q(t)} V_{N(t)-i} + R(t), \tag{4.46}$$

where $N(t) - i$ is the customer number of the ith customer in the queue at time t. Since $L^q(t)$ is a function only of the interarrival times in $(0, t)$ and the service times of the customers that have already been served, we see that for each sample value $L^q(t) = \ell$, the rv's $\widetilde{V}_1, \dots, \widetilde{V}_\ell$ each have the service time distribution F_V. Thus, taking expected values,

$$\mathsf{E}[U(t)] = \mathsf{E}[L^q(t)]\,\mathsf{E}[V] + \mathsf{E}[R(t)]. \tag{4.47}$$

Figure 4.16 illustrates how to find the time-average of $R(t)$. Viewing $R(t)$ as a reward

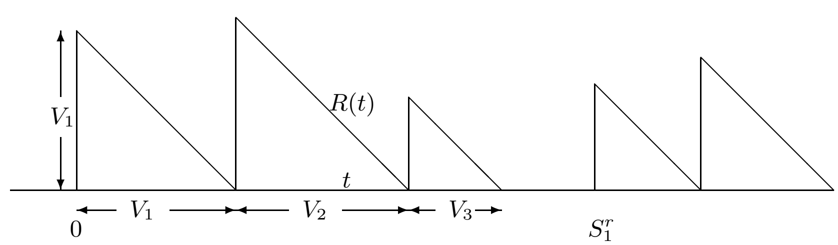


Figure 4.16: Sample value of the residual life function of customers in service.

function, we can find the accumulated reward up to time t as the sum of triangular areas.

First, consider $\int R(\tau)d\tau$ from 0 to $S^r_{N^r(t)}$, i.e., the accumulated reward up to the final renewal epoch in $[o,t]t$. Note that $S^r_{N^r(t)}$ is not only a renewal epoch for the renewal process, but also an arrival epoch for the arrival process; in particular, it is the $N(S^r_{N^r(t)})$th arrival epoch, and the $N(S^r_{N^r(t)}) - 1$ earlier arrivals are the customers that have received service up to time $S^r_{N^r(t)}$. Thus,

$$\int_0^{S^r_{N^r(t)}} R(\tau)\,d\tau \;=\; \sum_{i=0}^{N(S^r_{N^r(t)})-1} \frac{V_i^2}{2} \;\le\; \sum_{i=0}^{N(t)} \frac{V_i^2}{2}.$$

We can similarly upper bound the term on the right above by $\int_0^{S^r_{N^r(t)+1}} R(\tau)\,d\tau$. We also know (from going through virtually the same argument several times) that $(1/t)\int_{\tau=0}^{t} R(\tau)d\tau$ will approach a limit[15] with probability 1 as $t \to \infty$, and that the limit will be unchanged if t is replaced with $S^r_{N^r(t)}$ or $S^r_{N^r(t)+1}$. Thus, taking λ as the arrival rate,

$$\lim_{t\to\infty} \frac{\int_0^t R(\tau)\,d\tau}{t} \;=\; \lim_{t\to\infty} \frac{\sum_{i=1}^{A(t)} V_i^2}{2A(t)}\,\frac{A(t)}{t} \;=\; \frac{\lambda \mathsf{E}\left[V^2\right]}{2} \qquad \text{WP1}.$$

We will see in the next section that the time average above can be replaced with a limiting ensemble-average, so that

$$\lim_{t\to\infty} \mathsf{E}\left[R(t)\right] = \frac{\lambda \mathsf{E}\left[V^2\right]}{2}. \tag{4.48}$$

The next section also shows that there is a limiting ensemble-average form of (4.44), showing that $\lim_{t\to\infty} \mathsf{E}\left[L^q(t)\right] = \lambda \overline{W}^q$. Substituting this plus (4.48) into (4.47), we get

$$\lim_{t\to\infty} \mathsf{E}\left[U(t)\right] = \lambda \mathsf{E}\left[V\right]\, \overline{W}^q + \frac{\lambda \mathsf{E}\left[V^2\right]}{2}. \tag{4.49}$$

Thus $\mathsf{E}\left[U(t)\right]$ is asymptotically independent of t. It is now important to distinguish between $\mathsf{E}\left[U(t)\right]$ and $\overline{W}^q$. The first is the expected unfinished work at time t, which is the queueing delay that a customer would incur by arriving at t; the second is the sample-path-average expected queueing delay. For Poisson arrivals, the probability of an arrival in $(t, t+\delta]$ is independent of all earlier arrivals and service times, so it is independent of $U(t)$[16]. Thus, in the limit $t \to \infty$, each arrival faces an expected delay $\lim_{t\to\infty} \mathsf{E}\left[U(t)\right]$, so $\lim_{t\to\infty} \mathsf{E}\left[U(t)\right]$ must be equal to $\overline{W}^q$. Substituting this into (4.49), we obtain the celebrated *Pollaczek-Khinchin* formula,

$$\overline{W}^q = \frac{\lambda \mathsf{E}\left[V^2\right]}{2(1 - \lambda \mathsf{E}\left[V\right])}. \tag{4.50}$$

[15] In fact, one could simply take the limit without bringing in the renewal process, since it is clear by now that the renewal process justifies the limit with probability 1.

[16] This is often called the *PASTA* property, standing for Poisson arrivals see time-averages. This holds with great generality, requiring only that time-averages exist and that the parameters of interest at a given time t are independent of future arrivals. At the same time, this property is somewhat vague, so it should be used to help the intuition rather than to prove theorems.

This queueing delay has some of the peculiar features of residual life, and in particular, if $\mathsf{E}\left[V^2\right] = \infty$, the limiting expected queueing delay is infinite even when the expected service time is less than the expected interarrival interval.

In trying to visualize why the queueing delay is so large when $\mathsf{E}\left[V^2\right]$ is large, note that while a particularly long service is taking place, numerous arrivals are coming into the system, and all are being delayed by this single long service. In other words, the number of new customers held up by a long service is proportional to the length of the service, and the amount each of them are held up is also proportional to the length of the service. This visualization is rather crude, but does serve to explain the second moment of V in (4.50). This phenomenon is sometimes called the "slow truck effect" because of the pile up of cars behind a slow truck on a single lane road.

For a G/G/1 queue, (4.49) is still valid, but arrival times are no longer independent of $U(t)$, so that typically $\mathsf{E}[U(t)] \neq \overline{W}^q$. As an example, suppose that the service time is uniformly distributed between $1-\epsilon$ and $1+\epsilon$ and that the interarrival interval is uniformly distributed between $2-\epsilon$ and $2+\epsilon$. Assuming that $\epsilon < 1/2$, the system has no queueing and $\overline{W}^q = 0$. On the other hand, for small ϵ, $\lim_{t\to\infty} \mathsf{E}[U(t)] \sim 1/4$ (i.e., the server is busy half the time with unfinished work ranging from 0 to 1).

4.6 Expected number of renewals

The purpose of this section is to evaluate $\mathsf{E}[N(t)]$, denoted $m(t)$, as a function of $t > 0$ for arbitrary renewal processes. We first find an exact expression, in the form of an integral equation, for $m(t)$. This can be easily solved by Laplace transform methods in special cases. For the general case, however, $m(t)$ becomes increasingly messy for large t, so we then find the asymptotic behavior of $m(t)$. Since $N(t)/t$ approaches $1/\overline{X}$ with probability 1, we might expect $m(t)$ to grow with a derivative $m'(t)$ that asymptotically approaches $1/\overline{X}$. This is not true in general. Two somewhat weaker results, however, are true. The first, called the elementary renewal theorem (Theorem 4.6.1), states that $\lim_{t\to\infty} m(t)/t = 1/\overline{X}$. The second result, called Blackwell's theorem (Theorem 4.6.2), states that, subject to some limitations on $\delta > 0$, $\lim_{t\to\infty}[m(t+\delta) - m(t)] = \delta/\overline{X}$. This says essentially that the expected renewal rate approaches steady state as $t \to \infty$. We will find a number of applications of Blackwell's theorem throughout the remainder of the text.

The exact calculation of $m(t)$ makes use of the fact that the expectation of a nonnegative random variable is defined as the integral of its complementary distribution function,

$$m(t) = \mathsf{E}[N(t)] = \sum_{n=1}^{\infty} \Pr\{N(t) \geq n\}.$$

Since the event $\{N(t) \geq n\}$ is the same as $\{S_n \leq t\}$, $m(t)$ is expressed in terms of the distribution functions of S_n, $n \geq 1$, as follows.

$$m(t) = \sum_{n=1}^{\infty} \Pr\{S_n \leq t\}. \tag{4.51}$$

Although this expression looks fairly simple, it becomes increasingly complex with increasing t. As t increases, there is an increasing set of values of n for which $\Pr\{S_n \le t\}$ is significant, and $\Pr\{S_n \le t\}$ itself is not that easy to calculate if the interarrival distribution $\mathsf{F}_X(x)$ is complicated. The main utility of (4.51) comes from the fact that it leads to an integral equation for $m(t)$. Since $S_n = S_{n-1} + X_n$ for each $n \ge 1$ (interpreting S_0 as 0), and since X_n and S_{n-1} are independent, we can use the convolution equation (1.11) to get

$$\Pr\{S_n \le t\} = \int_{x=0}^{t} \Pr\{S_{n-1} \le t - x\}\, d\mathsf{F}_X(x) \quad \text{for } n \ge 2.$$

Substituting this in (4.51) for $n \ge 2$ and using the fact that $\Pr\{S_1 \le t\} = \mathsf{F}_X(t)$, we can interchange the order of integration and summation to get

$$\begin{aligned} m(t) &= \mathsf{F}_X(t) + \int_{x=0}^{t} \sum_{n=2}^{\infty} \Pr\{S_{n-1} \le t - x\}\, d\mathsf{F}_X(x) \\ &= \mathsf{F}_X(t) + \int_{x=0}^{t} \sum_{n=1}^{\infty} \Pr\{S_n \le t - x\}\, d\mathsf{F}_X(x) \\ &= \mathsf{F}_X(t) + \int_{x=0}^{t} m(t-x) d\mathsf{F}_X(x)\,; \qquad t \ge 0. \end{aligned} \tag{4.52}$$

An alternative derivation is given in Exercise 4.22. This integral equation is called the *renewal equation*. The following alternative form is achieved by integration by parts.[17]

$$m(t) = \mathsf{F}_X(t) + \int_{\tau=0}^{t} F_X(t-\tau)\, dm(\tau)\,; \qquad t \ge 0. \tag{4.53}$$

4.6.1 Laplace transform approach

If we assume that $X \ge 0$ has a density $\mathsf{f}_X(x)$, and that this density has a Laplace transform[18] $L_X(s) = \int_0^\infty \mathsf{f}_X(x) e^{-sx} dx$, then we can take the Laplace transform of both sides of (4.52). Note that the final term in (4.52) is the convolution of m with f_X, so that the Laplace transform of $m(t)$ satisfies

$$L_m(s) = \frac{L_X(s)}{s} + L_m(s) L_X(s).$$

Solving for $L_m(s)$,

$$L_m(s) = \frac{L_X(s)}{s[1 - L_X(s)]}. \tag{4.54}$$

[17] A mathematical subtlety with the Stieltjes integrals (4.52) and (4.53) will be discussed in Section 4.7.3.

[18] Note that $L_X(s) = \mathsf{E}\left[e^{-sX}\right] = g_X(-s)$ where g is the MGF of X. Thus the argument here could be carried out using the MGF. We use the Laplace transform since the mechanics here are so familiar to most engineering students

Example 4.6.1. As a simple example of how this can be used to calculate $m(t)$, suppose $\mathsf{f}_X(x) = (1/2)e^{-x} + e^{-2x}$ for $x \geq 0$. The Laplace transform is given by

$$L_X(s) = \frac{1}{2(s+1)} + \frac{1}{s+2} = \frac{(3/2)s + 2}{(s+1)(s+2)}.$$

Substituting this into (4.54) yields

$$L_m(s) = \frac{(3/2)s + 2}{s^2(s + 3/2)} = \frac{4}{3s^2} + \frac{1}{9s} - \frac{1}{9(s + 3/2)}.$$

We can solve for $m(t), t \geq 0$, by taking the inverse Laplace transform,

$$m(t) = \frac{4t}{3} + \frac{1 - \exp[-(3/2)t]}{9}.$$

The procedure in this example can be used for any inter-renewal density $\mathsf{f}_X(x)$ for which the Laplace transform is a rational function, i.e., a ratio of polynomials. In such cases, $L_m(s)$ will also be a rational function. The Heaviside inversion formula (i.e., factoring the denominator and expressing $L_m(s)$ as a sum of individual poles as done above) can then be used to calculate $m(t)$. In the example above, there was a second order pole at $s = 0$ leading to the linear term $4t/3$ in $m(t)$, there was a first order pole at $s = 0$ leading to the constant $1/9$, and there was a pole at $s = -3/2$ leading to the exponentially decaying term.

We now show that a second order pole at $s = 0$ always occurs when $L_X(s)$ is a rational function. To see this, note that $L_X(0)$ is just the integral of $\mathsf{f}_X(x)$, which is 1; thus $1 - L_X(s)$ has a zero at $s = 0$ and $L_m(s)$ has a second order pole at $s = 0$. To evaluate the residue for this second order pole, we recall that the first and second derivatives of $L_X(s)$ at $s = 0$ are $-\mathsf{E}[X]$ and $\mathsf{E}[X^2]$ respectively. Expanding $L_X(s)$ in a power series around $s = 0$ then yields $L_X(s) = 1 - s\mathsf{E}[X] + (s^2/2)\mathsf{E}[X^2]$ plus terms of order s^3 or higher. This gives us

$$L_m(s) = \frac{1 - s\overline{X} + (s^2/2)\mathsf{E}[X^2] + \cdots}{s^2\left[\overline{X} - (s/2)\mathsf{E}[X^2] + \cdots\right]} = \frac{1}{s^2\overline{X}} + \frac{1}{s}\left(\frac{\mathsf{E}[X^2]}{2\overline{X}^2} - 1\right) + \cdots. \tag{4.55}$$

The remaining terms are the other poles of $L_m(s)$ with their residues. For values of s with $\Re(s) \geq 0$, we have $|L_X(s)| = |\int \mathsf{f}_X(x)e^{-sx}dx| \leq \int \mathsf{f}_X(x)|e^{-sx}|dx \leq \int \mathsf{f}_X(x)dx = 1$ with strict inequality except for $s = 0$. Thus $L_X(s)$ cannot have any poles on the imaginary axis or the right half plane, and $1 - L_X(s)$ cannot have any zeros there other than the one at $s = 0$. It follows that all the remaining poles of $L_m(s)$ are strictly in the left half plane. This means that the inverse transforms for all these remaining poles die out as $t \to \infty$. Thus the inverse Laplace transform of $L_m(s)$ is

$$\begin{aligned} m(t) &= \frac{t}{\overline{X}} + \frac{\mathsf{E}[X^2]}{2\overline{X}^2} - 1 + \epsilon(t) \\ &= \frac{t}{\overline{X}} + \frac{\sigma^2}{2\overline{X}^2} - \frac{1}{2} + \epsilon(t) \qquad \text{for } t \geq 0, \end{aligned} \tag{4.56}$$

where $\lim_{t\to\infty} \epsilon(t) = 0$.

We have derived (4.56) only for the special case in which $\mathsf{f}_X(x)$ has a rational Laplace transform. For this case, (4.56) implies both the elementary renewal theorem ($\lim_{t\to\infty} m(t)/t = 1/\overline{X}$) and also Blackwell's theorem ($\lim_{t\to\infty}[m(t+\delta) - m(t)] = \delta/\overline{X}$). We will interpret the meaning of the constant term $\sigma^2/(2\overline{X}^2) - 1/2$ in Section 4.8.

4.6.2 The elementary renewal theorem

Theorem 4.6.1 (The elementary renewal theorem). *Let $\{N(t);\, t > 0\}$ be a renewal counting process with mean inter-renewal interval $\overline{X}$. Then $\lim_{t\to\infty} \mathsf{E}[N(t)]/t = 1/\overline{X}$.*

Discussion: We have already seen that $m(t) = \mathsf{E}[N(t)]$ is finite for all $t > 0$ (see Exercise 4.2). The theorem is proven by establishing a lower and upper bound to $m(t)/t$ and showing that each approaches $1/\mathsf{E}[X]$ as $t \to \infty$. The key element for each bound is (4.35), repeated below, which comes from the Wald equality.

$$m(t) = \frac{\mathsf{E}\left[S_{N(t)+1}\right]}{\overline{X}} - 1. \tag{4.57}$$

Proof: The lower bound to $m(t)/t$ comes by recognizing that $S_{N(t)+1}$ is the epoch of the first arrival after t. Thus $\mathsf{E}\left[S_{N(t)+1}\right] > t$. Substituting this into (4.57),

$$\frac{m(t)}{t} > \frac{1}{\mathsf{E}[X]} - \frac{1}{t}.$$

Clearly this lower bound approaches $1/\mathsf{E}[X]$ as $t \to \infty$. The upper bound, which is more difficult[19] and might be omitted on a first reading, is established by first truncating $X(t)$ and then applying (4.57) to the truncated process.

For an arbitrary constant $b > 0$, let $\breve{X}_i = \min(b, X_i)$. Since these truncated random variables are IID, they form a related renewal counting process $\{\breve{N}(t); t > 0\}$ with $\breve{m}(t) = \mathsf{E}\left[\breve{N}(t)\right]$ and $\breve{S}_n = \breve{X}_1 + \cdots + \breve{X}_n$. Since $\breve{X}_i \leq X_i$ for all i, we see that $\breve{S}_n \leq S_n$ for all n. Since $\{S_n \leq t\} = \{N(t) \geq n\}$, it follows that $\breve{N}(t) \geq N(t)$ and thus $\breve{m}(t) \geq m(t)$. Finally, in the truncated process, $\breve{S}_{\breve{N}(t)+1} \leq t + b$ and thus $\mathsf{E}\left[\breve{S}_{\breve{N}(t)+1}\right] \leq t + b$. Thus, applying (4.57) to the truncated process,

$$\frac{m(t)}{t} \leq \frac{\breve{m}(t)}{t} = \frac{\mathsf{E}\left[S_{\breve{N}(t)+1}\right]}{t\mathsf{E}\left[\breve{X}\right]} - \frac{1}{t} \;\leq\; \frac{t+b}{t\mathsf{E}\left[\breve{X}\right]}.$$

Next, choose $b = \sqrt{t}$. Then

$$\frac{m(t)}{t} \leq \frac{1}{\mathsf{E}\left[\breve{X}\right]} + \frac{1}{\sqrt{t}\,\mathsf{E}\left[\breve{X}\right]}.$$

[19] The difficulty here, and the reason for using a truncation argument, comes from the fact that the residual life, $S_{N(t)+1} - t$ at t might be arbitrarily large. We saw in Section 4.4 that the time-average residual life is infinite if $\mathsf{E}\left[X^2\right]$ is infinite. Figure 4.6 also illustrates why residual life can be so large.

Note finally that $\mathsf{E}\left[\breve{X}\right] = \int_0^b [1-\mathsf{F}_X(x)]\,dx$. Since $b = \sqrt{t}$, we have $\lim_{t\to\infty} \mathsf{E}\left[\breve{X}\right] = \mathsf{E}[X]$, completing the proof. □

Note that this theorem (and its proof) have not assumed finite variance. It can also be seen that the theorem holds when $\mathsf{E}[X]$ is infinite, since $\lim_{t\to\infty} \mathsf{E}\left[\breve{X}\right] = \infty$ in this case.

Recall that $N[t, \omega]/t$ is the average number of renewals from 0 to t for a sample function ω, and $m(t)/t$ is the average of this over ω. Combining with Theorem 4.3.1, we see that the limiting time and ensemble-average equals the time-average renewal rate for each sample function except for a set of probability 0.

Another interesting question is to determine the expected renewal rate in the limit of large t without averaging from 0 to t. That is, are there some values of t at which renewals are more likely than others for large t? If the inter-renewal intervals have an integer distribution function (i.e., each inter-renewal interval must last for an integer number of time units), then each renewal epoch S_n must also be an integer. This means that $N(t)$ can increase only at integer times and the expected rate of renewals is zero at all non-integer times.

An obvious generalization of integer valued inter-renewal intervals is that of inter-renewals that occur only at integer multiples of some real number $d > 0$. Such a distribution is called an *arithmetic distribution*. The *span* of an arithmetic distribution is the largest number λ such that this property holds. Thus, for example if X takes on only the values 0, 2, and 6, its distribution is arithmetic with span $\lambda = 2$. Similarly, if X takes on only the values 1/3 and 1/5, then the span is $\lambda = 1/15$. The remarkable thing, for our purposes, is that any inter-renewal distribution that is not an arithmetic distribution leads to a uniform expected rate of renewals in the limit of large t. This result is contained in Blackwell's renewal theorem, which we state without proof.[20]. Recall, however, that for the special case of an inter-renewal density with a rational Laplace transform, Blackwell's renewal theorem is a simple consequence of (4.56).

Theorem 4.6.2 (Blackwell). *If a renewal process has an inter-renewal distribution that is non-arithmetic, then for each $\delta > 0$,*

$$\lim_{t\to\infty} [m(t+\delta) - m(t)] = \frac{\delta}{\mathsf{E}[X]}. \tag{4.58}$$

If the inter-renewal distribution is arithmetic with span λ, then

$$\lim_{t\to\infty} [m(t+\lambda) - m(t)] = \frac{\lambda}{\mathsf{E}[X]}. \tag{4.59}$$

Eq. (4.58) says that for non-arithmetic distributions, the expected number of arrivals in the interval $(t, t+\delta]$ is equal to $\delta/\mathsf{E}[X]$ in the limit $t \to \infty$. Since the theorem is true for arbitrarily small δ, the theorem almost seems to be saying that $m(t)$ has a derivative for large t, but this is not true. One can see the reason by looking at an example where X can take on only the values 1 and π. Then no matter how large t is, $N(t)$ can only increase at discrete points of time of the form $k + j\pi$ where k and j are nonnegative integers. Thus

[20] See Theorem 1 of Section 11.1, of [8]) for a proof

$dm(t)/dt$ is either 0 or ∞ for all t. As t gets larger, the jumps in $m(t)$ become both smaller in magnitude and more closely spaced from one to the next. Thus $\big[m(t+\delta) - m(t)\big]/\delta$ can approach $1/\mathsf{E}[X]$ as $t \to \infty$ for any fixed δ (as the theorem says), but as δ gets smaller, the convergence in t gets slower. For the above example (and for all discrete non-arithmetic distributions), $\big[m(t+\delta) - m(t)\big]/\delta$ does not approach[21] $1/\mathsf{E}[X]$ for any t as $\delta \to 0$.

For an arithmetic renewal process with span λ, the asymptotic behavior of $m(t)$ as $t \to \infty$ is much simpler. Renewals can only occur at multiples of λ, and since simultaneous renewals are not allowed, either 0 or 1 renewal occurs at each time $k\lambda$. Thus for any k. we have

$$\Pr\{\text{Renewal at } \lambda k\} = m(\lambda k) - m(\lambda(k-1)), \tag{4.60}$$

where, by convention, we take $m(0) = 0$. Thus (4.59) can be restated as

$$\lim_{k\to\infty} \Pr\{\text{Renewal at } k\lambda\} = \frac{\lambda}{\overline{X}}. \tag{4.61}$$

The limiting behavior of $m(t)$ is discussed further in the next section.

4.7 Renewal-reward processes; ensemble-averages

Theorem 4.4.1 showed that if a renewal-reward process has an expected inter-renewal interval $\overline{X}$ and an expected inter-renewal reward $\mathsf{E}[R_n]$, then the time-average reward is $\mathsf{E}[R_n]/\overline{X}$ with probability 1. In this section, we explore the ensemble average, $\mathsf{E}[R(t)]$, as a function of time t. It is easy to see that $\mathsf{E}[R(t)]$ typically changes with t, especially for small t, but a question of major interest here is whether $\mathsf{E}[R(t)]$ approaches a constant as $t \to \infty$.

In more concrete terms, if the arrival times of busses at a bus station forms a renewal process, then the waiting time for the next bus, *i.e.*, the residual life, starting at time t, can be represented as a reward function $R(t)$. We would like to know if the expected waiting time depends critically on t, where t is the time since the renewal process started, *i.e.*, the time since a hypothetical bus number 0 arrived. If $\mathsf{E}[R(t)]$ varies significantly with t, even as $t \to \infty$, it means that the choice of $t = 0$ as the beginning of the initial interarrival interval never dies out as $t \to \infty$.

Blackwell's renewal theorem (and common sense) tell us that there is a large difference between arithmetic inter-renewal times and non-arithmetic inter-renewal times. For the arithmetic case, all renewals occur at multiples of the span λ. Thus, for example, the expected waiting time (*i.e.*, the expected residual life) decreases at rate 1 from each multiple of λ to the next, and it increases with a jump equal to the probability of an arrival at each multiple of λ. For this reason, we usually consider various reward functions only at multiples of λ. We would guess, then, that $\mathsf{E}[R(n\lambda)]$ approaches a constant as $n \to \infty$.

For the non-arithmetic case, on the other hand, the expected number of renewals in any small interval of length δ becomes independent of t as $t \to \infty$, so we might guess that

[21] This must seem like mathematical nitpicking to many readers. However, $m(t)$ is the expected number of renewals in $(0, t]$, and how $m(t)$ varies with t, is central to this chapter and keeps reappearing.

$\mathsf{E}[R(t)]$ approaches a limit as $t \to \infty$. We would also guess that these asymptotic ensemble averages are equal to the appropriate time averages from Section 4.4.

The bottom line for this section is that under very broad conditions, the above guesses are essentially correct. Thus the limit as $t \to \infty$ of a given ensemble-average reward can usually be computed simply by finding the time-average and vice-versa. Sometimes time-averages are simpler, and sometimes ensemble-averages are. The advantage of the ensemble-average approach is both the ability to find $\mathsf{E}[R(t)]$ for finite values of t and to understand the rate of convergence to the asymptotic result.

The following subsection is restricted to the arithmetic case. We will derive the joint distribution function of age and duration for any given time t, and then look at the limit as $t \to \infty$. This leads us to arbitrary reward functions (such as residual life) for the arithmetic case. We will not look specifically at generalized reward functions that depend on other processes, but this generalization is quite similar to that for time-averages.

The non-arithmetic case is analyzed in the remainder of the subsections of this section. The basic ideas are the same as the arithmetic case, but a number of subtle mathematical limiting issues arise. The reader is advised to understand the arithmetic case first, since the limiting issues in the non-arithmetic case can then be viewed within the intuitive context of the arithmetic case.

4.7.1 Age and duration for arithmetic processes

Let $\{N(t); t > 0\}$ be an arithmetic renewal counting process with inter-renewal intervals $X_1, X_2, \dots$ and arrival epochs $S_1, S_2, \dots$, where $S_n = X_1 + \cdots + X_n$. To keep the notation as uncluttered as possible, we take the span to be one and then scale to an arbitrary λ later. Thus each X_i is a positive integer-valued rv.

Recall that the age $Z(t)$ at any given $t > 0$ is $Z(t) = t - S_{N(t)}$ (where by convention $S_0 = 0$) and the duration $\widetilde{X}(t)$ is $\widetilde{X}(t) = S_{N(t)+1}(t) - S_{N(t)}$. Since arrivals occur only at integer times, we initially consider age and duration only at integer times also. If an arrival occurs at integer time t, then $S_{N(t)} = t$ and $Z(t) = 0$. Also, if $S_1 > t$, then $N(t) = 0$ and $Z(t) = t$ (*i.e.*, the age is taken to be t if no arrivals occur up to and including time t). Thus, for integer t, $Z(t)$ is an integer-valued rv taking values from $[0, t]$. Since $S_{N(t)+1} > t$, it follows that $\widetilde{X}(t)$ is an integer-valued rv satisfying $\widetilde{X}(t) > Z(t)$. Since both are integer valued, $\widetilde{X}(t)$ must exceed $Z(t)$ by at least 1 (or by λ in the more general case of span λ).

In order to satisfy $Z(t) = i$ and $\widetilde{X}(t) = k$ for given $i < t$, it is necessary and sufficient to have an arrival epoch at $t - i$ followed by an interarrival interval of length k, where $k \geq i + 1$. For $Z(t) = t$ and $\widetilde{X}(t) = k$, it is necessary and sufficient that $k > t$, *i.e.*, that the first inter-renewal epoch occurs at $k > t$.

Theorem 4.7.1. *Let $\{X_n; n \geq 1\}$ be the interarrival intervals of an arithmetic renewal process with unit span. Then the the joint PMF of the age and duration at integer time*

$t \geq 1$ is given by

$$\mathsf{p}_{Z(t),\widetilde{X}(t)}(i,k) = \begin{cases} \mathsf{p}_X(k) & \text{for } i=t,\ k>t \\ q_{t-i}\,\mathsf{p}_X(k) & \text{for } 0 \leq i < t,\ k>i \\ 0 & \text{otherwise} \end{cases} \cdot \tag{4.62}$$

where $q_j = \Pr\{$arrival at time $j\}$. The limit as $t \to \infty$ for any given $0 \leq i < k$ is given by

$$\lim_{\text{integer } t\to\infty} \mathsf{p}_{Z(t),\widetilde{X}(t)}(i,k) = \frac{\mathsf{p}_X(k)}{\overline{X}}. \tag{4.63}$$

Proof: The idea is quite simple. For the upper part of (4.62), note that the age is t if and only there are no arrivals in $(0,t]$, which corresponds to $X_1 = k$ for some $k > t$. For the middle part, the age is i for a given $i < t$ if and only if there is an arrival at $t - i$ and the next arrival epoch is after t, which means that the corresponding interarrival interval k exceeds i. The probability of an arrival at $t-i$, *i.e.*, q_{t-i}, depends only on the arrival epochs up to and including time $t - i$, which should be independent of the subsequent interarrival time, leading to the product in the middle term of (4.62). To be more precise about this independence, note that for $i < t$,

$$q_{t-i} \;=\; \Pr\{\text{arrival at } t-i\} = \sum_{n\geq 1} \mathsf{p}_{S_n}(t-i). \tag{4.64}$$

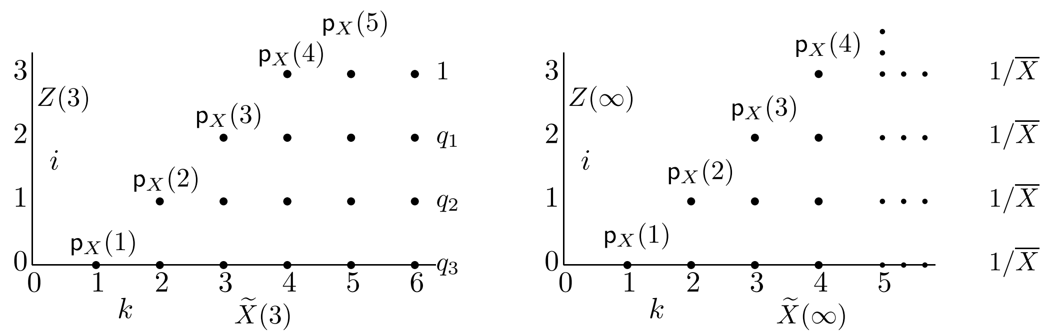
Given that $S_n = t - i$, the probability that $X_{n+1} = k$ is $\mathsf{p}_X(k)$. This is the same for all n, establishing (4.62).

For any fixed i, k with $i < k$, note that only the middle term in (4.62) is relevant as $t \to \infty$. Using Blackwell's theorem (4.61) to take the limit as $t \to \infty$, we get (4.63) □

The probabilities in the theorem, both for finite t and asymptotically as $t \to \infty$, are illustrated in Figure 4.17. The product form of the probabilities in (4.62) (as illustrated in the figure) might lead one to think that $Z(t)$ and $\widetilde{X}(t)$ are independent, but this is incorrect because of the constraint that $\widetilde{X}(t) > Z(t)$. It is curious that in the asymptotic case, (4.63) shows that, for a given duration $\widetilde{X}(t) = k$, the age is equally likely to have any integer value from 0 to $k-1$, *i.e.*, for a given duration, the interarrival interval containing t is uniformly distributed around t.

The marginal PMF for $Z(t)$ is calculated below using (4.62).

$$\mathsf{p}_{Z(t)}(i) = \begin{cases} \mathsf{F}^{\mathsf{c}}_X(i) & \text{for } i=t \\ q_{t-i}\,\mathsf{F}^{\mathsf{c}}_X(i) & \text{for } 0 \leq i < t \\ 0 & \text{otherwise} \end{cases} \cdot \tag{4.65}$$



a) Case where $t = 3$. b) Asymptotic case where $t \to \infty$

Figure 4.17: Joint PMF, $\mathsf{p}_{\widetilde{X}(t)Z(t)}(k, i)$ of $\widetilde{X}(t)$ and $Z(t)$ in an arithmetic renewal process with span 1. In part a), $t = 3$ and the PMF at each sample point is the product of two terms, $q_{t-i} = \Pr\{\text{Arrival at } t - i\}$ and $\mathsf{p}_X(k)$. Part b) is the asymptotic case where $t \to \infty$. Here the arrival probabilities become uniform.

where $\mathsf{F}^{\mathsf{c}}_X(i) = \mathsf{p}_X(i{+}1) + \mathsf{p}_X(i{+}2) + \cdots$. The marginal PMF for $\widetilde{X}(t)$ can be calculated directly from (4.62), but it is simplified somewhat by recognizing that

$$q_j \;=\; m(j) - m(j-1). \tag{4.66}$$

Substituting this into (4.62) and summing over age,

$$\mathsf{p}_{\widetilde{X}(t)}(k) = \begin{cases} \mathsf{p}_X(k)\big[m(t) - m(t{-}k)\big] & \text{for } k < t \\ \mathsf{p}_X(k)\, m(t) & \text{for } k = t \\ \mathsf{p}_X(k)\big[m(t) + 1\big] & \text{for } k > t \end{cases} \;. \tag{4.67}$$

The term +1 in the expression for $k > t$ corresponds to the uppermost point for the given k in Figure 4.17a. This accounts for the possibility of no arrivals up to time t. It is not immediately evident that $\sum_k \mathsf{p}_{\widetilde{X}(t)}(k) = 1$, but this can be verified from the renewal equation, (4.52).

Blackwell's theorem shows that the arrival probabilities tend to $1/\overline{X}$ as $t \to \infty$, so the limiting marginal probabilities for age and duration become

$$\lim_{\text{integer } t\to\infty} \mathsf{p}_{Z(t)}(i) \;=\; \frac{\mathsf{F}^{\mathsf{c}}_X(i)}{\overline{X}}. \tag{4.68}$$

$$\lim_{\text{integer } t\to\infty} \mathsf{p}_{\widetilde{X}(t)}(k) \;=\; \frac{k\,\mathsf{p}_X(k)}{\overline{X}}. \tag{4.69}$$

The expected value of $Z(t)$ and $\widetilde{X}(t)$ can also be found for all t from (4.62) and (4.63) respectively, but they don't have a particularly interesting form. The asymptotic values as $t \to \infty$ are more simple and interesting. The asymptotic expected value for age is derived below from (4.68).[22]

$$
\begin{aligned}
\lim_{\text{integer } t\to\infty} \mathsf{E}[Z(t)] &= \sum_i i \lim_{\text{integer } t\to\infty} \mathsf{p}_{Z(t)}(i) \\
&= \frac{1}{\overline{X}} \sum_{i=1}^{\infty} \sum_{j=i+1}^{\infty} i\, \mathsf{p}_X(j) = \frac{1}{\overline{X}} \sum_{j=2}^{\infty} \sum_{i=1}^{j-1} i\, \mathsf{p}_X(j) \\
&= \frac{1}{\overline{X}} \sum_{j=2}^{\infty} \frac{j(j-1)}{2}\, \mathsf{p}_X(j) \\
&= \frac{\mathsf{E}\left[X^2\right]}{2\overline{X}} - \frac{1}{2}. \qquad (4.70)
\end{aligned}
$$

This limiting ensemble average age has the same dependence on $\mathsf{E}\left[X^2\right]$ as the time average in (4.16), but, perhaps surprisingly, it is reduced from that amount by $1/2$. To understand this, note that we have only calculated the expected age at integer values of t. Since arrivals occur only at integer values, the age for each sample function must increase with unit slope as t is increased from one integer to the next. The expected age thus also increases, and then at the next integer value, it drops discontinuously due to the probability of an arrival at that next integer. Thus the limiting value of $\mathsf{E}[Z(t)]$ has a saw tooth shape and the value at each discontinuity is the lower side of that discontinuity. Averaging this asymptotic expected age over a unit of time, the average is $\mathsf{E}\left[X^2\right]/2\overline{X}$, in agreement with (4.16).

As with the time average, the limiting expected age is infinite if $\mathsf{E}\left[X^2\right] = \infty$. However, for each t, $Z(t) \le t$, so $\mathsf{E}[Z(t)] < \infty$ for all t, increasing without bound as $t \to \infty$

The asymptotic expected duration is derived in a similar way, starting from (4.69)

$$
\begin{aligned}
\lim_{\text{integer } t\to\infty} \mathsf{E}\left[\widetilde{X}(t)\right] &= \sum_k \lim_{\text{integer } t\to\infty} k \mathsf{p}_{\widetilde{X}(t)}(k) \\
&= \sum_k \frac{k^2 \mathsf{p}_X(k)}{\overline{X}} = \frac{\mathsf{E}\left[X^2\right]}{\overline{X}}. \qquad (4.71)
\end{aligned}
$$

This agrees with the time average in (4.17). The reduction by $1/2$ seen in (4.70) is not present here, since as t is increased in the interval $[t, t{+}1)$, $X(t)$ remains constant.

Since the asymptotic ensemble-average age differs from the time-average age in only a trivial way, and the asymptotic ensemble-average duration is the same as the time-average duration, it might appear that we have gained little by this exploration of ensemble averages. What we have gained, however, is a set of results that apply to all t. Thus they show how (in principle) these results converge as $t \to \infty$.

[22] Finding the limiting expectations from the limiting PMF's requires interchanging a limit with an expectation. This can be justified (in both (4.70) and (4.71)) by assuming that X has a finite second moment and noting that all the terms involved are positive, that $\Pr\{\text{arrival at } j\} \le 1$ for all j, and that $\mathsf{p}_X(k) \le 1$ for all k.

4.7.2 Joint age and duration: non-arithmetic case

Non-arithmetic renewal processes are mathematically more complicated than arithmetic renewal processes, but the concepts are the same. We start by looking at the joint probability of the age and duration, each over an incremental interval. (see Figure 4.18 (a)).

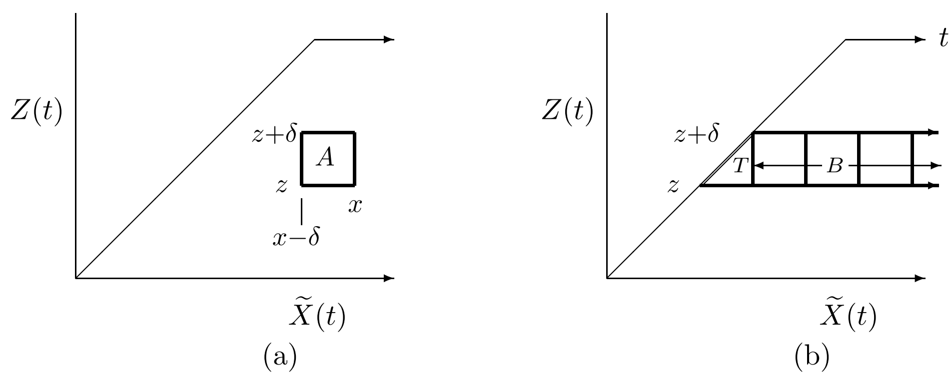


Figure 4.18: Part (a): The incremental square region A of sample values for $Z(t)$ and $\widetilde{X}(t)$ whose joint probability is specified in (4.73). The square is assumed to be inside the indicated semi-infinite trapezoidal region, *i.e.*, to satisfy $0 \leq z < z{+}\delta \leq t$ and $z{+}2\delta \leq x$.

Part (b): Summing over discrete sample regions of $\widetilde{X}(t)$ to find the marginal probability $\Pr\{z \leq Z(t) < z{+}\delta\} = \Pr\{T\} + \Pr\{B\}$, where T is the triangular area and B the set of indicated squares.

Theorem 4.7.2. *Consider an arbitrary renewal process with age $Z(t)$ and duration $\widetilde{X}(t)$ at any given time $t > 0$. Let A be the event*

$$A = \{z \leq Z(t) < z{+}\delta\} \bigcap \{x{-}\delta < \widetilde{X}(t) \leq x\}, \tag{4.72}$$

where $0 \leq z < z{+}\delta \leq t$ and $z{+}2\delta \leq x$. Then

$$\Pr\{A\} = \big[m(t{-}z) - m(t{-}z{-}\delta)\big]\, \big[\mathsf{F}_X(x) - \mathsf{F}_X(x{-}\delta)\big]. \tag{4.73}$$

If in addition the renewal process is non-arithmetic,

$$\lim_{t\to\infty} \Pr\{A\} = \frac{\big[\mathsf{F}_X(x) - \mathsf{F}_X(x{-}\delta)\big]}{\overline{X}}. \tag{4.74}$$

Proof: Note that A is the box illustrated in Figure 4.18 (a) and that under the given conditions, $\widetilde{X}(t) > Z(t)$ for all sample points in A. Recall that $Z(t) = t - S_{N(t)}$ and $\widetilde{X}(t) = S_{N(t)+1} - S_{N(t)} = X_{N(t)+1}$, so A can also be expressed as

$$A = \{t{-}z{-}\delta < S_{N(t)} \leq t - z\} \bigcap \{x - \delta < X_{N(t)+1} \leq x\}. \tag{4.75}$$

We now argue that A can also be rewritten as

$$A = \bigcup_{n=1}^{\infty} \Big\{ \{t{-}z{-}\delta < S_n \leq t - z\} \bigcap \{x - \delta < X_{n+1} \leq x\} \Big\}. \tag{4.76}$$

To see this, first assume that the event in (4.75) occurs. Then $N(t)$ must have some positive sample value n, so (4.76) occurs. Next assume that the event in (4.76) occurs, which means that one of the events, say the nth, in the union occurs. Since $S_{n+1} = S_n + X_{n+1} > t$, we see that n is the sample value of $S_{N(t)}$ and (4.75) must occur.

To complete the proof, we must find $\Pr\{A\}$. First note that A is a union of disjoint events. That is, although more than one arrival epoch might occur in $(t{-}z{-}\delta,\ t{-}z]$, the following arrival epoch can exceed t for only one of them. Thus

$$\Pr\{A\} = \sum_{n=1}^{\infty} \Pr\Big\{ \{t{-}z{-}\delta < S_n \leq t - z\} \bigcap \{x - \delta < X_{n+1} \leq x\} \Big\}. \tag{4.77}$$

For each n, X_{n+1} is independent of S_n, so

$$\begin{aligned} &\Pr\Big\{ \{t{-}z{-}\delta < S_n \leq t - z\} \bigcap \{x - \delta < X_{n+1} \leq x\} \Big\} \\ &\qquad = \Pr\big\{t{-}z{-}\delta < S_n \leq t - z\big\} \big[\mathsf{F}^{\mathsf{c}}_X(x) - \mathsf{F}^{\mathsf{c}}_X(x{-}\delta\big]. \end{aligned}$$

Substituting this into (4.77) and using (4.51) to sum the series, we get

$$\Pr\{A\} = \big[m(t{-}z) - m(t{-}z{-}\delta)\big] \big[\mathsf{F}_X(x) - \mathsf{F}_X(x{-}\delta)\big].$$

This establishes (4.73). Blackwell's theorem then establishes (4.74). □

It is curious that $\Pr\{A\}$ has such a simple product expression, where one term depends only on the function $m(t)$ and the other only on the distribution function F_X. Although the theorem is most useful as $\delta \to 0$, the expression is exact for all δ such that the square region A satisfies the given constraints (*i.e.*, A lies in the indicated semi-infinite trapezoidal region).

4.7.3 Age $Z(t)$ for finite t: non-arithmetic case

In this section, we first use Theorem 4.7.2 to find bounds on the marginal incremental probability of $Z(t)$. We then find the distribution function, $\mathsf{F}_{Z(t)}(z)$, and the expected value, $\mathsf{E}[Z(t)]$ of $Z(t)$.

Corollary 4.7.1. *For $0 \leq z < z{+}\delta \leq t$, the following bounds hold on $\Pr\{z \leq Z(t) < z{+}\delta\}$.*

$$\Pr\{z \leq Z(t) < z{+}\delta\} \;\geq\; \big[m(t{-}z) - m(t{-}z{-}\delta)\big]\mathsf{F}^{\mathsf{c}}_X(z{+}\delta) \tag{4.78}$$

$$\Pr\{z \leq Z(t) < z{+}\delta\} \;\leq\; \big[m(t{-}z) - m(t{-}z{-}\delta)\big]\mathsf{F}^{\mathsf{c}}_X(z). \tag{4.79}$$

Proof*: As indicated in Figure 4.18 (b), $\Pr\{z \leq Z(t) < z{+}\delta\} = \Pr\{T\} + \Pr\{B\}$, where T is the triangular region,

$$T = \{z \leq Z(t) < z{+}\delta\} \bigcap \{Z(t) < \widetilde{X}(t) \leq z{+}\delta\},$$

and B is the rectangular region

$$B = \{z \le Z(t) < z+\delta\} \bigcap \{\widetilde{X}(t) > z+\delta\}.$$

It is easy to find $\Pr\{B\}$ by summing the probabilities in (4.73) for the squares indicated in Figure 4.18 (b). The result is

$$\Pr\{B\} = \big[m(t-z) - m(t-z-\delta)\big]\, \mathsf{F}^{\mathsf{c}}_X(z+\delta). \tag{4.80}$$

Since $\Pr\{T\} \ge 0$, this establishes the lower bound in (4.78). We next need an upper bound to $\Pr\{T\}$ and start by finding an event that includes T.

$$\begin{aligned} T &= \{z \le Z(t) < z+\delta\} \bigcap \{Z(t) < \widetilde{X}(t) \le z+\delta\} \\ &= \bigcup_{n\ge 1} \Big[\{t-z-\delta < S_n \le t-z\} \bigcap \{t - S_n < X_{n+1} \le z+\delta\}\Big] \\ &\subseteq \bigcup_{n\ge 1} \Big[\{t-z-\delta < S_n \le t-z\} \bigcap \{z < X_{n+1} \le z+\delta\}\Big], \end{aligned}$$

where the inclusion follows from the fact that $S_n \le t-z$ for the event on the left to occur, and this implies that $z \le t - S_n$.

Using the union bound, we then have

$$\begin{aligned} \Pr\{T\} &\le \Big[\sum_{n\ge 1} \Pr\{\{t-z-\delta < S_n \le t-z\}\Big] \big[\mathsf{F}^{\mathsf{c}}_X(z) - \mathsf{F}^{\mathsf{c}}_X(z+\delta)\big] \\ &= \big[m(t-z) - m(t-z-\delta)\big] \big[\mathsf{F}^{\mathsf{c}}_X(z) - \mathsf{F}^{\mathsf{c}}_X(z+\delta)\big]. \end{aligned}$$

Combining this with (4.80), we have (4.79). □

The following corollary uses Corollary 4.7.1 to determine the distribution function of $Z(t)$. The rather strange dependence on the existence of a Stieltjes integral will be explained after the proof.

Corollary 4.7.2. *If the Stieltjes integral* $\int_{t-z}^{t} \mathsf{F}^{\mathsf{c}}_X(t-\tau)\, dm(\tau)$ *exists for given* $t > 0$ *and* $0 < z < t$*, then*

$$\Pr\{Z(t) \le z\} = \int_{t-z}^{t} \mathsf{F}^{\mathsf{c}}_X(t-\tau)\, dm(\tau). \tag{4.81}$$

Proof: First, partition the interval $[0, z)$ into a given number ℓ of increments, each of size $\delta = z/\ell$. Then

$$\Pr\{Z(t) < z\} = \sum_{k=0}^{\ell-1} \Pr\{k\delta \le Z(t) < k\delta + \delta\}.$$

Applying the bounds in 4.78 and 4.79 to the terms in this sum,

$$\Pr\{Z(t) < z\} \ge \sum_{k=0}^{\ell-1} \Big[m\big(t-k\delta\big) - m\big(t-k\delta-\delta\big)\Big] \mathsf{F}^{\mathsf{c}}_X(k\delta+\delta) \tag{4.82}$$

$$\Pr\{Z(t) < z\} \le \sum_{k=0}^{\ell-1} \Big[m\big(t-k\delta\big) - m\big(t-k\delta-\delta\big)\Big] \mathsf{F}^{\mathsf{c}}_X\big(k\delta\big). \tag{4.83}$$

These are, respectively, lower and upper Riemann sums for the Stieltjes integral $\int_0^z \mathsf{F}^{\mathsf{c}}_X(t{-}\tau)\,dm(\tau)$. Thus, if this Stieltjes integral exists, then, letting $\delta = z/\ell \to 0$,

$$\Pr\{Z(t) < z\} = \int_{t-z}^{t} \mathsf{F}^{\mathsf{c}}_X(t-\tau)\,dm(\tau).$$

This is a convolution and thus the Stieltjes integral exists unless $m(\tau)$ and $\mathsf{F}^{\mathsf{c}}_X(t-\tau)$ both have a discontinuity at some $\tau \in [0, z]$ (see Exercise 1.12). If no such discontinuity exists, then $\Pr\{Z(t) < z\}$ cannot have a discontinuity at z. Thus, if the Stieltjes integral exists, $\Pr\{Z(t) < z\} = \Pr\{Z(t) \le z\}$, and, for $z < t$,

$$\mathsf{F}_{Z(t)}(z) = \int_{t-z}^{t} \mathsf{F}^{\mathsf{c}}_X(t-\tau)\,dm(\tau).$$

□

The above argument showed us that the values of z at which the Stieltjes integral in (4.81) fails to exist are those at which $\mathsf{F}_{Z(t)}(z)$ has a step discontinuity. At these values we know that $\mathsf{F}_{Z(t)}(z)$ (as a distribution function) should have the value at the top of the step (thus including the discrete probability that $\Pr\{Z(t) = z\}$). In other words, at any point z of discontinuity where the Stieltjes integral does not exist, $\mathsf{F}_{Z(t)}(z)$ is the limit[23] of $\mathsf{F}_{Z(t)}(z{+}\epsilon)$ as $\epsilon > 0$ approaches 0. Another way of expressing this is that for $0 \le z < t$, $\mathsf{F}_{Z(t)}(z)$ is the limit of the upper Riemann sum on the right side of (4.83).

The next corollary uses an almost identical argument to find $\mathsf{E}[Z(t)]$. As we will see, the Stieltjes integral fails to exist at those values of t at which there is a discrete positive probability of arrival. The expected value at these points is the lower Riemann sum for the Stieltjes integral.

Corollary 4.7.3. *If the Stieltjes integral $\int_0^t \mathsf{F}^{\mathsf{c}}_X(t-\tau)\,dm(\tau)$ exists for given $t > 0$, then*

$$\mathsf{E}[Z(t)] = \mathsf{F}^{\mathsf{c}}_X(t) + \int_0^t (t-\tau)\mathsf{F}^{\mathsf{c}}_X(t-\tau)\,dm(\tau). \tag{4.84}$$

Proof: Note that $Z(t) = t$ if and only if $X_1 > t$, which has probability $\mathsf{F}^{\mathsf{c}}_X(t)$. For the other possible values of $Z(t)$, we divide $[0,t)$ into ℓ equal intervals of length $\delta = t/\ell$ each. Then $\mathsf{E}[Z(t)]$ can be lower bounded by

$$\begin{aligned}
\mathsf{E}[Z(t)] \;&\ge\; \mathsf{F}^{\mathsf{c}}_X(t) + \sum_{k=0}^{\ell-1} k\delta \Pr\{k\delta \le Z(t) < k\delta+\delta\}\} \\
&\ge\; \mathsf{F}^{\mathsf{c}}_X(t) + \sum_{k=0}^{\ell-1} k\delta \left[m(t{-}k\delta) - m(t{-}k\delta{-}\delta)\right]\mathsf{F}^{\mathsf{c}}_X(k\delta+\delta).
\end{aligned}$$

[23]This seems to be rather abstract mathematics, but as engineers, we often evaluate functions with step discontinuities by ignoring the values at the discontinuities or evaluating these points by adhoc means.

where we used 4.78 for the second step. Similarly, $\mathsf{E}[Z(t)]$ can be upper bounded by

$$\begin{aligned}\mathsf{E}[Z(t)] &\leq \mathsf{F}^{\mathsf{c}}_X(t) + \sum_{k=0}^{\ell-1}(k\delta+\delta)\Pr\{k\delta \leq Z(t) < k\delta+\delta\}\} \\ &\leq \mathsf{F}^{\mathsf{c}}_X(t) + \sum_{k=0}^{\ell-1}(k\delta+\delta)\big[m(t-k\delta) - m(t-k\delta-\delta)\big]\mathsf{F}^{\mathsf{c}}_X(k\delta).\end{aligned}$$

where we used 4.79 for the second step. These provide lower and upper Riemann sums to the Stieltjes integral in (4.81), completing the proof in the same way as the previous corollary. □

4.7.4 Age $Z(t)$ as $t \to \infty$: non-arithmetic case

Next, for non-arithmetic renewal processes, we want to find the limiting values, as $t \to \infty$, for $\mathsf{F}_{Z(t)}(z)$ and $\mathsf{E}[Z(T)]$. Temporarily ignoring any subtleties about the limit, we first view $dm(t)$ as going to $\frac{dt}{\overline{X}}$ as $t \to \infty$. Thus from (4.81),

$$\lim_{t\to\infty} \Pr\{Z(t) \leq z\} = \frac{1}{\overline{X}} \int_0^z \mathsf{F}^{\mathsf{c}}_X(\tau)d\tau. \tag{4.85}$$

If X has a PDF, this simplifies further to

$$\lim_{t\to\infty} \mathsf{f}_{Z(t)}(z) = \frac{1}{\overline{X}}\mathsf{f}_X(z). \tag{4.86}$$

Note that this agrees with the time-average result in (4.29). Taking these limits carefully requires more mathematics than seems justified here, especially since the result uses Blackwell's theorem, which was not proven here. Thus we state (without proof) another theorem, equivalent to Blackwell's theorem, called the key renewal theorem, that simplifies taking this type of limit. Essentially Blackwell's theorem is easier to interpret, but the key renewal theorem is often easier to use.

Theorem 4.7.3 (Key renewal theorem). *Let $r(x) \geq 0$ be a directly Riemann integrable function, and let $m(t) = \mathsf{E}[N(t)]$ for a non-arithmetic renewal process. Then*

$$\lim_{t\to\infty} \int_{\tau=0}^{t} r(t-\tau)dm(\tau) = \frac{1}{\overline{X}} \int_0^\infty r(x)dx. \tag{4.87}$$

We first explain what directly Riemann integrable means. If $r(x)$ is nonzero only over finite limits, say $[0, b]$, then direct Riemann integration means the same thing as ordinary Riemann integration (as learned in elementary calculus). However, if $r(x)$ is nonzero over $[0, \infty)$, then the ordinary Riemann integral (if it exists) is the result of integrating from 0 to b and then taking the limit as $b \to \infty$. The direct Riemann integral (if it exists) is the result of taking a Riemann sum over the entire half line, $[0, \infty)$ and then taking the limit as the grid becomes finer. Exercise 4.25 gives an example of a simple but bizarre function that is Riemann integrable but not directly Riemann integrable. If $r(x) \geq 0$ can be upper bounded by a decreasing Riemann integrable function, however, then, as shown in Exercise 4.25,$r(x)$ must be directly Riemann integrable. The bottom line is that restricting $r(x)$ to be directly Riemann integrable is not a major restriction.

Next we interpret the theorem. If $m(t)$ has a derivative, then Blackwell's theorem would suggest that $dm(t)/dt \to (1/\overline{X})\,dt$, which leads to (4.87) (leaving out the mathematical details). On the other hand, if X is discrete but non-arithmetic, then $dm(t)/dt$ can be intuitively viewed as a sequence of impulses that become smaller and more closely spaced as $t \to \infty$. Then $r(t)$ acts like a smoothing filter which, as $t \to \infty$, smoothes these small impulses. The theorem says that the required smoothing occurs whenever $r(t)$ is directly Riemann integrable. The theorem does not assert that the Stieltjes integral exists for all t, but only that the limit exists. For most applications to discrete inter-renewal intervals, the Stieltjes integral does not exist everywhere. Using the key renewal theorem, we can finally determine the distribution function and expected value of $Z(t)$ as $t \to \infty$. These limiting ensemble averages are, of course, equal to the time averages found earlier.

Theorem 4.7.4. *For any non-arithmetic renewal process, the limiting distribution function and expected value of the age $Z(t)$ are given by*

$$\lim_{t\to\infty} \mathsf{F}_{Z(t)}(z) = \frac{1}{\overline{X}} \int_0^z \mathsf{F}^{\mathsf{c}}_X(x)\,dx. \tag{4.88}$$

Furthermore, if $\mathsf{E}\left[X^2\right] < \infty$, then

$$\lim_{t\to z} \mathsf{E}\left[Z(t)\right] = \frac{\mathsf{E}\left[X^2\right]}{2\overline{X}}. \tag{4.89}$$

Proof: For any given $z > 0$, let $r(x) = \mathsf{F}^{\mathsf{c}}_X(x)$ for $0 \le x \le z$ and $r(x) = 0$ elsewhere. Then (4.81) becomes

$$\Pr\{Z(t) \le z\} = \int_0^t r(t-\tau)\,dm(\tau).$$

Taking the limit as $t \to \infty$,

$$\begin{aligned} \lim_{t\to\infty} \Pr\{Z(t) \le z\} &= \lim_{t\to\infty} \int_0^t r(t-\tau)\,dm(\tau) \\ &= \frac{1}{\overline{X}} \int_0^\infty r(x)\,dx \;=\; \frac{1}{\overline{X}} \int_0^z \mathsf{F}^{\mathsf{c}}_X(x)\,dx, \end{aligned} \tag{4.90}$$

where in (4.90) we used the fact that $\mathsf{F}^{\mathsf{c}}_X(x)$ is decreasing to justify using (4.87). This establishes (4.88).

To establish (4.89), we again use the key renewal theorem, but here we let $r(x) = x\,\mathsf{F}^{\mathsf{c}}_X(x)$. Exercise 4.25 shows that $x\,\mathsf{F}^{\mathsf{c}}_X(x)$ is directly Riemann integrable if $\mathsf{E}\left[X^2\right] < \infty$. Then, taking the limit of (4.84 and then using (4.87), we have

$$\begin{aligned} \lim_{t\to\infty} \mathsf{E}\left[Z(t)\right] &= \lim_{t\to\infty} \mathsf{F}^{\mathsf{c}}_X(t) + \int_0^t r(t-\tau)\,dm(\tau) \\ &= \frac{1}{\overline{X}} \int_0^\infty r(x)\,dx \;=\; \frac{1}{\overline{X}} \int_0^\infty x\mathsf{F}^{\mathsf{c}}_X(x)\,dx. \end{aligned}$$

Integrating this by parts, we get (4.89). □

4.7.5 Arbitrary renewal-reward functions: non-arithmetic case

If we omit all the mathematical precision from the previous three subsections, we get a very simple picture. We started with (4.72), which gave the probability of an incremental square region A in the $(Z(t), \widetilde{X}(t))$ plane for given t. We then converted various sums over an increasingly fine grid of such regions into Stieltjes integrals. These integrals evaluated the distribution and expected value of age at arbitrary values of t. Finally, the key renewal theorem let us take the limit of these values as $t \to \infty$.

In this subsection, we will go through the same procedure for an arbitrary reward function, say $R(t) = \mathcal{R}(Z(t), \widetilde{X}(t))$, and show how to find $\mathsf{E}[R(T)]$. Note that $\Pr\{Z(t) \leq z\} = \mathsf{E}\left[\mathbb{I}_{Z(t)\leq z}\right]$ is a special case of $\mathsf{E}[R(T)]$ where $R(t)$ is chosen to be $\mathbb{I}_{Z(t)\leq z}$. Similarly, finding the distribution function at a given argument for any rv can be converted to the expectation of an indicator function. Thus, having a methodology for finding the expectation of an arbitrary reward function also covers distribution functions and many other quantities of interest.

We will leave out all the limiting arguments here about converting finite incremental sums of areas into integrals, since we have seen how to do that in treating $Z(t)$. In order to make this general case more transparent, we use the following shorthand for A when it is incrementally small:

$$\Pr\{A\} = m'(t-z) f_X(x)\, dx\, dz, \tag{4.91}$$

where, if the derivatives exist, $m'(\tau) = dm(\tau)/d\tau$ and $f_X(x) = d\mathsf{F}_X(x)/dx$. If the derivatives do not exist, we can view $m'(\tau)$ and $f_X(x)$ as generalized functions including impulses, or, more appropriately, view them simply as shorthand. After using the shorthand as a guide, we can put the results in the form of Stieltjes integrals and verify the mathematical details at whatever level seems appropriate.

We do this first for the example of the distribution function of duration, $\Pr\left\{\widetilde{X}(t) \leq x_0\right\}$, where we first assume that $x_0 \leq t$. As illustrated in Figure 4.19, the corresponding reward function $R(t)$ is 1 in the triangular region where $\widetilde{X}(t) \leq x_0$ and $Z(t) < \widetilde{X}(t)$. It is 0 elsewhere.

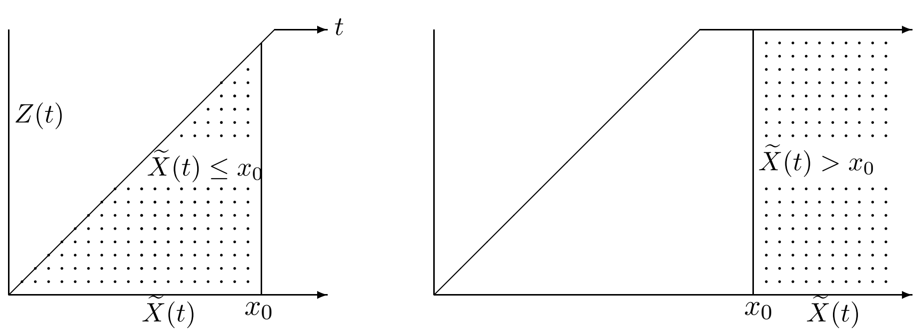


Figure 4.19: Finding $\mathsf{F}_{\widetilde{X}(t)}(x_0)$ for $x_0 \leq t$ and for $x_0 > t$.

$$
\begin{aligned}
\Pr\left\{\widetilde{X}(t) \le x_0\right\} &= \int_{z=0}^{x_0}\int_{x=z}^{x_0} m'(t-z)\mathsf{f}_X(x)\,dx\,dz \\
&= \int_{z=0}^{x_0} m'(t{-}z)\big[\mathsf{F}_X(x_0) - \mathsf{F}_X(z)\big]\,dz \\
&= \mathsf{F}_X(x_0)\big[m(t) - m(t{-}x_0)\big] - \int_{t-x_0}^{t} \mathsf{F}_X(t{-}\tau)\,dm(\tau). \qquad (4.92)
\end{aligned}
$$

For the opposite case, where $x_0 > t$, it is easier to find $\Pr\left\{\widetilde{X}(t) > x_0\right\}$. As shown in the figure, this is the region where $0 \le Z(t) \le t$ and $\widetilde{X}(t) > x_0$. There is a subtlety here in that the incremental areas we are using are only valid for $Z(t) < t$. If the age is equal to t, then no renewals have occured in $(0, t]$, so that $\Pr\left\{\widetilde{X}(t) > x_0; Z(t) = t\right\} = \mathsf{F}^{\mathsf{c}}_X(x_0)$. Thus

$$
\begin{aligned}
\Pr\left\{\widetilde{X}(t) > x_0\right\} &= \mathsf{F}^{\mathsf{c}}_X(x_0) + \int_{z=0}^{t^-}\int_{x=x_0}^{\infty} m'(t-z)\mathsf{f}_X(x)\,dx\,dz \\
&= \mathsf{F}^{\mathsf{c}}_X(x_0) + m(t)\mathsf{F}^{\mathsf{c}}_X(x_0). \qquad (4.93)
\end{aligned}
$$

As a sanity test, the renewal equation, (4.53), can be used to show that the sum of (4.92) and (4.93) at $x_0 = t$ is equal to 1 (as they must be if the equations are correct).

We can now take the limit, $\lim_{t\to\infty} \Pr\left\{\widetilde{X}(t) \le x_0\right\}$. For any given x_0, (4.92) holds for sufficiently large t, and the key renewal theorem can be used since the integral has a finite range. Thus,

$$
\begin{aligned}
\lim_{t\to\infty} \Pr\left\{\widetilde{X}(t) \le x_0\right\} &= \frac{1}{\overline{X}}\left[x_0\mathsf{F}_X(x_0) - \int_0^{x_0} \mathsf{F}_X(x)\,dx\right] \\
&= \frac{1}{\overline{X}}\int_0^{x_0}\Big[\mathsf{F}_X(x_0) - \mathsf{F}_X(x)\Big]\,dx \\
&= \frac{1}{\overline{X}}\int_0^{x_0}\Big[\mathsf{F}^{\mathsf{c}}_X(x) - dF^{\mathsf{c}}_X(x_0)\Big]\,dx. \qquad (4.94)
\end{aligned}
$$

It is easy to see that the right side of (4.94) is increasing from 0 to 1 with increasing x_0, *i.e.*, it is a distribution function.

After this example, it is now straightforward to write down the expected value of an arbitrary renewal-reward function $R(t)$ whose sample value at $Z(t) = z$ and $X(t) = x$ is denoted by $\mathcal{R}(z, x)$. We have

$$
\mathsf{E}\left[R(t)\right] = \int_{x=t}^{\infty} \mathcal{R}(t,x)\,d\mathsf{F}_X(x) + \int_{z=0}^{t}\int_{x=z}^{\infty} \mathcal{R}(z,x)\,d\mathsf{F}_X(x)\,dm(t-z). \qquad (4.95)
$$

The first term above arises from the subtlety discussed above for the case where $Z(t) = t$. The second term is simply the integral over the semi-infinite trapezoidal area in Figure 4.19.

The analysis up to this point applies to both the arithmetic and nonarithmetic cases, but we now must assume again that the renewal process is nonarithmetic. If the inner integral, *i.e.*, $\int_{x=z}^{\infty} \mathcal{R}(z,x)\, d\mathsf{F}_X(x)$, as a function of z, is directly Riemann integrable, then not only can the key renewal theorem be applied to this second term, but also the first term must approach 0 as $t \to \infty$. Thus the limit of (4.95) as $t \to \infty$ is

$$\lim_{t\to\infty} \mathsf{E}\left[R(t)\right] = \frac{1}{\overline{X}} \int_{z=0}^{\infty} \int_{x=z}^{\infty} \mathcal{R}(z,x)\, d\mathsf{F}_X(x)\, dz. \tag{4.96}$$

This is the same expression as found for the time-average renewal reward in Theorem 4.4.1. Thus, as indicated earlier, we can now equate any time-average result for the nonarithmetic case with the corresponding limiting ensemble average, and the same equations have been derived in both cases.

As a simple example of (4.96), let $\mathcal{R}(z,t) = x$. Then $\mathsf{E}\left[R(t)\right] = \mathsf{E}\left[\widetilde{X}(t)\right]$ and

$$\begin{aligned} \lim_{t\to\infty} \mathsf{E}\left[\widetilde{X}(t)\right] &= \frac{1}{\overline{X}} \int_{z=0}^{\infty} \int_{x=z}^{\infty} x\, d\mathsf{F}_X(x)\, dz \;=\; \int_{x=0}^{\infty} \int_{z=0}^{x} x\, dz\, d\mathsf{F}_X(x) \\ &= \frac{1}{\overline{X}} \int_{x=0}^{\infty} x^2\, d\mathsf{F}_X(x) \;=\; \frac{\mathsf{E}\left[X^2\right]}{\overline{X}}. \end{aligned} \tag{4.97}$$

After calculating the integral above by interchanging the order of integration, we can go back and assert that the key renewal theorem applies if $\mathsf{E}\left[X^2\right]$ is finite. If it is infinite, then it is not hard to see that $\lim_{t\to\infty} \mathsf{E}\left[\widetilde{X}(t)\right]$ is infinite also.

It has been important, and theoretically reassuring, to be able to find ensemble-averages for nonarithmetic renewal-reward functions in the limit of large t and to show (not surprisingly) that they are the same as the time-average results. The ensemble-average results are quite tricky, though, and it is wise to check results achieved that way with the corresponding time-average results.

4.8 Delayed renewal processes

We have seen a certain awkwardness in our discussion of Little's theorem and the M/G/1 delay result because an arrival was assumed, but not counted, at time 0; this was necessary for the first inter-renewal interval to be statistically identical to the others. In this section, we correct that defect by allowing the epoch at which the first renewal occurs to be arbitrarily distributed. The resulting type of process is a generalization of the class of renewal processes known as *delayed renewal processes*. The word *delayed* does not necessarily imply that the first renewal epoch is in any sense larger than the other inter-renewal intervals. Rather, it means that the usual renewal process, with IID inter-renewal times, is delayed until after the epoch of the first renewal. What we shall discover is intuitively satisfying — both the time-average behavior and, in essence, the limiting ensemble behavior are not affected by the distribution of the first renewal epoch. It might be somewhat surprising, however, to find that this irrelevance of the distribution of the first renewal epoch holds even when the mean of the first renewal epoch is infinite.

To be more precise, we let $\{X_i;\, i{\geq}1\}$ be a set of independent nonnegative random variables. X_1 has a given distribution function $\mathsf{G}(x)$, whereas $\{X_i;\, i \geq 2\}$ are identically distributed with a given distribution function $\mathsf{F}(x)$. Thus a renewal process is a special case of a delayed renewal process for which $\mathsf{G}(x) = \mathsf{F}(x)$. Let $S_n = \sum_{i=1}^n X_i$ be the nth renewal epoch. We first show that the SLLN still holds despite the deviant behavior of X_1.

Lemma 4.8.1. *Let $\{X_i;\ i \geq 2$ be IID with a mean $\overline{X}$ satisfying $\mathsf{E}[|X|] < \infty$ and let X_1 be a rv, independent of $\{X_i;\ i \geq 2\}$. Let $S_n = \sum_{i=1}^n X_i$. Then $\lim S_n/n = \overline{X}$ WP1.*

Prrof: Note that

$$\frac{S_n}{n} = \frac{X_1}{n} + \frac{\sum_{i=2}^n X_i}{n}.$$

Since X_1 is finite WP1, the first term above goes to 0 WP1 as $n \to \infty$. The second term goes to $\overline{X}$, proving the lemma (which is thus a trivial variation of the SLLN). □

Now, for the given delayed renewal process, let $N(t)$ be the number of renewal epochs up to and including time t. This is still determined by the fact that $\{N(t) \geq n\}$ if and only if $\{S_n \leq t\}$). $\{N(t);\, t > 0\}$ is then called a delayed renewal counting process. The following simple lemma follows from lemma 4.3.1.

Lemma 4.8.2. *Let $\{N(t);\, t > 0\}$ be a delayed renewal counting process. Then $\lim_{t\to\infty} N(t) = \infty$ with probability 1 and $\lim_{t\to\infty} \mathsf{E}[N(t)] = \infty$.*

Proof: Conditioning on $X_1 = x$, we can write $N(t) = 1 + N'(t - x)$ where $N'\{t; t \geq 0\}$ is the ordinary renewal counting process with inter-renewal intervals $X_2, X_3, \ldots$. From Lemma 4.3.1, $\lim_{t\to\infty} N'(t - x) = \infty$ with probability 1, and $\lim_{t\to\infty} \mathsf{E}[N'(t - x)] = \infty$. Since this is true for every finite $x > 0$, and X_1 is finite with probability 1, the lemma is proven. □

Theorem 4.8.1 (Strong Law for Delayed Renewal Processes). *Let $N(t);\ t > 0$ be the renewal counting process for a delayed renewal process where the inter-renewal intervals $X_2, X_3, \ldots,$ have distribution function F and finite mean $\overline{X} = \int_{x=0}^{\infty}[1 - \mathsf{F}(x)]\, dx$. Then*

$$\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{\overline{X}} \qquad \textit{WP1.} \tag{4.98}$$

Proof: Using Lemma 4.8.1, the conditions for Theorem 4.3.2 are fulfilled, so the proof follows exactly as the proof of Theorem 4.3.1. □.

Next we look at the elementary renewal theorem and Blackwell's theorem for delayed renewal processes. To do this, we view a delayed renewal counting process $\{N(t);\, t > 0\}$ as an ordinary renewal counting process that starts at a random nonnegative epoch X_1 with some distribution function $\mathsf{G}(t)$. Define $N_o(t - X_1)$ as the number of renewals that occur in the interval $(X_1, t]$. Conditional on any given sample value x for X_1, $\{N_o(t-x);\, t-x > 0\}$ is an ordinary renewal counting process and thus, given $X_1 = x$, $\lim_{t\to\infty} \mathsf{E}[N_o(t - x)]/(t - x) = 1/\overline{X}$. Since $N(t) = 1 + N_o(t - X_1)$ for $t > X_1$, we see that, conditional on $X_1 = x$,

$$\lim_{t\to\infty} \frac{\mathsf{E}[N(t) \mid X_1{=}x]}{t} = \lim_{t\to\infty} \frac{\mathsf{E}[N_o(t - x)]}{t - x}\ \frac{t - x}{t} = \frac{1}{\overline{X}}. \tag{4.99}$$

Since this is true for every finite sample value x for X_1, we we have established the following theorem:

Theorem 4.8.2 (Elementary Delayed Renewal Theorem). *For a delayed renewal process with* $\mathsf{E}[X_i] = \overline{X}$ *for* $i \geq 2$*,*

$$\lim_{t\to\infty} \frac{\mathsf{E}[N(t)]}{t} = \frac{1}{\overline{X}}. \tag{4.100}$$

The same approach gives us Blackwell's theorem. Specifically, if $\{X_i;\, i{\geq}2\}$ is a sequence of IID non-arithmetic rv's, then, for any $\delta > 0$, Blackwell's theorem for ordinary renewal processes implies that

$$\lim_{t\to\infty} \frac{\mathsf{E}[N_o(t-x+\delta) - N_o(t-x)]}{\delta} = \frac{1}{\overline{X}}. \tag{4.101}$$

Thus, conditional on any sample value $X_1 = x$, $\lim_{t\to\infty} \mathsf{E}[N(t{+}\delta) - N(t) \,|\, X_1{=}x] = \delta/\overline{X}$. Taking the expected value over X_1 gives us $\lim_{t\to\infty} \mathsf{E}[N(t+\delta) - N(t)] = \delta/\overline{X}$. The case in which $\{X_i;\, i{\geq}2\}$ are arithmetic with span λ is somewhat more complicated. If X_1 is arithmetic with span λ (or a multiple of λ), then the first renewal epoch must be at some multiple of λ and $\lambda/\overline{X}$ gives the expected number of arrivals at time $i\lambda$ in the limit as $i \to \infty$. If X_1 is non-arithmetic or arithmetic with a span other than a multiple of λ, then the effect of the first renewal epoch never dies out, since all subsequent renewals occur at multiples of λ from this first epoch. We ignore this rather ugly case and state the following theorem for the nice situations.

Theorem 4.8.3 (Blackwell for Delayed Renewal). *If* $\{X_i;\, i{\geq}2\}$ *are non-arithmetic, then, for all* $\delta > 0$*,*

$$\lim_{t\to\infty} \frac{\mathsf{E}[N(t+\delta) - N(t)]}{\delta} = \frac{1}{\overline{X}}. \tag{4.102}$$

If $\{X_i; i \geq 2\}$ *are arithmetic with span* λ *and mean* $\overline{X}$ *and* X_1 *is arithmetic with span* $m\lambda$ *for some positive integer* m*, then*

$$\lim_{i\to\infty} \Pr\{\textit{renewal at } t = i\lambda\} = \frac{\lambda}{\overline{X}}. \tag{4.103}$$

4.8.1 Delayed renewal-reward processes

We have seen that the distribution of the first renewal epoch has no effect on the time or ensemble-average behavior of a renewal process (other than the ensemble dependence on time for an arithmetic process). This carries over to reward functions with almost no change. In particular, the generalized version of Theorem 4.4.1 is as follows:

Theorem 4.8.4. *Let* $\{N(t);\, t > 0\}$ *be a delayed renewal counting process where the inter-renewal intervals* $X_2, X_3, \ldots$ *have the distribution function* F*. Let* $Z(t) = t - S_{N(t)}$*, let* $\widetilde{X}(t) = S_{N(t)+1} - S_{N(t)}$*, and let* $R(t) = \mathcal{R}(Z(t), \widetilde{X}(t))$ *be a reward function. Assume that*

$$\mathsf{E}[\mathrm{R}_n] = \int_{x=0}^{\infty} \int_{z=0}^{x} \mathcal{R}(z,x)\, dz\, d\mathsf{F}(x) < \infty\,.$$

Then, with probability one,

$$\lim_{t\to\infty} \frac{1}{t} \int_{\tau=0}^{t} R(\tau)d\tau = \frac{\mathsf{E}[\mathrm{R}_n]}{\overline{X}_2} \ \textit{for}\ \ n \geq 2. \tag{4.104}$$

We omit the proof of this since it is a minor variation of that of theorem 4.4.1. Finally, the equivalence of time and limiting ensemble averages holds as before, yielding

$$\lim_{t\to\infty} \mathsf{E}[R(t)] = \frac{\mathsf{E}[\mathrm{R}_n]}{\overline{X}_2}. \tag{4.105}$$

4.8.2 Transient behavior of delayed renewal processes

Let $m(t) = \mathsf{E}[N(t)]$ for a delayed renewal process. As in (4.51), we have

$$m(t) = \sum_{n=1}^{\infty} \Pr\{N(t) \geq n\} = \sum_{n=1}^{\infty} \Pr\{S_n \leq t\}. \tag{4.106}$$

For $n \geq 2$, $S_n = S_{n-1} + X_n$ where X_n and S_{n-1} are independent. From the convolution equation (1.12),

$$\Pr\{S_n \leq t\} = \int_{x=0}^{t} \Pr\{S_{n-1} \leq t-x\}\ d\mathsf{F}(x) \quad \text{for}\, n \geq 2. \tag{4.107}$$

For $n = 1$, $\Pr\{S_n \leq t\} = \mathsf{G}(t)$. Substituting this in (4.106) and interchanging the order of integration and summation,

$$\begin{aligned} m(t) &= \mathsf{G}(t) + \int_{x=0}^{t} \sum_{n=2}^{\infty} \Pr\{S_{n-1} \leq t-x\}\, d\mathsf{F}(x) \\ &= \mathsf{G}(t) + \int_{x=0}^{t} \sum_{n=1}^{\infty} \Pr\{S_n \leq t-x\}\, d\mathsf{F}(x) \\ &= \mathsf{G}(t) + \int_{x=0}^{t} m(t-x)d\mathsf{F}(x)\,; \qquad t \geq 0. \end{aligned} \tag{4.108}$$

This is the *renewal equation* for delayed renewal processes and is a generalization of (4.52). It is shown to have a unique solution in [8], Section 11.1.

There is another useful integral equation very similar to (4.108) that arises from breaking up S_n as the sum of X_1 and $\widehat{S}_{n-1}$ where $\widehat{S}_{n-1} = X_2 + \cdots + X_n$. Letting $\widehat{m}(t)$ be the expected number of renewals in time t for an ordinary renewal process with interarrival distribution F, a similar argument to that above, starting with $\Pr\{S_n \leq t\} = \int_0^t \Pr\left\{\widehat{S}_{n-1} \leq t-x\right\}\ d\mathsf{G}(x)$ yields

$$m(t) = \mathsf{G}(t) + \int_{x=0}^{t} \widehat{m}(t-x)d\mathsf{G}(x). \tag{4.109}$$

This equation brings out the effect of the initial renewal interval clearly, and is useful in computation if one already knows $\widehat{m}(t)$.

Frequently, the most convenient way of dealing with $m(t)$ is through transforms. Following the same argument as that in (4.54), we get $L_m(r) = (1/r)L_{\mathsf{G}}(r) + L_m(r)L_{\mathsf{F}}(r)$. Solving, we get

$$L_m(r) = \frac{L_{\mathsf{G}}(r)}{r[1 - L_{\mathsf{F}}(r)]}. \tag{4.110}$$

We can find $m(t)$ from (4.110) by finding the inverse Laplace transform, using the same procedure as in Example 4.6.1. There is a second order pole at $r = 0$ again, and, evaluating the residue, it is $1/L'_{\mathsf{F}}(0) = 1/\overline{X}_2$, which is not surprising in terms of Blackwell's theorem. We can also expand numerator and denominator of (4.110) in a power series, as in (4.55). The inverse transform, corresponding to (4.56), is

$$m(t) = \frac{t}{\overline{X}} + \frac{\mathsf{E}\left[X_2^2\right]}{2\overline{X}} - \frac{\overline{X}_1}{\overline{X}} + \epsilon(t) \quad \text{for } t \to 0, \tag{4.111}$$

where $\lim_{t\to\infty} \epsilon(t) = 0$.

4.8.3 The equilibrium process

Consider an ordinary non-arithmetic renewal process with an inter-renewal interval X of distribution $\mathsf{F}(x)$. We have seen that the distribution of the interval from t to the next renewal approaches $\mathsf{F}_Y(y) = (1/\mathsf{E}[X]) \int_0^y [1 - \mathsf{F}(x)]dx$ as $t \to \infty$. This suggests that if we look at this renewal process starting at some very large t, we should see a delayed renewal process for which the distribution $\mathsf{G}(x)$ of the first renewal is equal to the residual life distribution $\mathsf{F}_Y(x)$ above and subsequent inter-renewal intervals should have the original distribution $\mathsf{F}(x)$ above. Thus it appears that such a delayed renewal process is the same as the original ordinary renewal process, except that it starts in "steady state." To verify this, we show that $m(t) = t/\overline{X}$ is a solution to (4.108) if $\mathsf{G}(t) = \mathsf{F}_Y(t)$. Substituting $(t - x)/\overline{X}$ for $m(t - x)$, the right hand side of (4.108) is

$$\frac{\int_0^t [1 - \mathsf{F}(x)]dx}{\overline{X}_2} + \frac{\int_0^t (t - x)d\mathsf{F}(x)}{\overline{X}} = \frac{\int_0^t [1 - \mathsf{F}(x)]dx}{\overline{X}} + \frac{\int_0^t \mathsf{F}(x)dx}{\overline{X}} = \frac{t}{\overline{X}},$$

where we have used integration by parts for the first equality. This particular delayed renewal process is called the *equilibrium process*, since it starts off in steady state, and thus has no transients.

4.9 Summary

Sections 4.1 to 4.7 developed the central results about renewal processes that frequently appear in subsequent chapters. The chapter starts with the strong law for renewal processes, showing that the time average rate of renewals, $N(t)/t$, approaches $1/\overline{X}$ with probability

1 as $t \to \infty$. This, combined with the strong law of large numbers, is the basis for most subsequent results about time-averages. Section 4.4 adds a reward function $R(t)$ to the underlying renewal process. These reward functions are defined to depend only on the inter-renewal interval containing t, and are used to study many surprising aspects of renewal processes such as residual life, age, and duration. For all sample paths of a renewal process (except a subset of probabiity 0), the time-average reward for a given $R(t)$ is a constant, and that constant is the expected aggregate reward over an inter-renewal interval divided by the expect length of an inter-renewal interval.

The next topic, in Section 4.5 is that of stopping trials. These have obvious applications to situations where an experiment or game is played until some desired (or undesired) outcome (based on the results up to and including the given trial) occurs. This is a basic and important topic in its right, but is also needed to understand both how the expected renewal rate $\mathsf{E}[N(t)]/t$ varies with time t and how renewal theory can be applied to queueing situations. Finally, we found that stopping rules were helpful in understanding G/G/1 queues, especially Little's theorem, and to derive an understand the Pollaczek-Khinchin expression for the expected delay in an M/G/1 queue.

This is followed, in Section 4.6, by an analysis of how $\mathsf{E}[N(t)]/t$ varies with t. This starts by using Laplace transforms to get a complete solution of the ensemble-average, $\mathsf{E}[N(t)]/t$, as a function of t, when the distribution of the inter-renewal interval has a rational Laplace transform. For the general case (where the Laplace transform is irrational or non-existant), the elementary renewal theorem shows that $\lim_{t\to\infty} \mathsf{E}[N(t)]/t = 1/\overline{X}$. The fact that the time-average (WP1) and the limiting ensemble-average are the same is not surprising, and the fact that the ensemble-average has a limit is not surprising. These results are so fundamental to other results in probability, however, that they deserve to be understood.

Another fundamental result in Section 4.6 is Blackwell's renewal theorem, showing that the distribution of renewal epochs reach a steady state as $t \to \infty$. The form of that steady state depends on whether the inter-renewal distribution is arithmetic (see (4.59)) or non-arithmetic (see (4.58)).

Section 4.7 ties together the results on rewards in 4.4 to those on ensemble averages in 4.6. Under some very minor restrictions imposed by the key renewal theorem, we found that, for non-arithmetic inter-renewal distributions, $\lim_{t\to\infty} \mathsf{E}[R(t)]$ is the same as the time-average value of reward.

Finally, all the results above were shown to apply to delayed renewal processes.

For further reading on renewal processes, see Feller,[8], Ross, [16], or Wolff, [22]. Feller still appears to be the best source for deep understanding of renewal processes, but Ross and Wolff are somewhat more accessible.

4.10 Exercises

Exercise 4.1. The purpose of this exercise is to show that for an arbitrary renewal process, $N(t)$, the number of renewals in $(0, t]$ is a (non-defective) random variable.

a) Let $X_1, X_2, \dots,$ be a sequence of IID inter-renewal rv's . Let $S_n = X_1 + \cdots + X_n$ be the corresponding renewal epochs for each $n \geq 1$. Assume that each X_i has a finite expectation $\overline{X} > 0$ and, for any given $t > 0$, use the weak law of large numbers to show that $\lim_{n\to\infty} \Pr\{S_n \leq t\} = 0$.

b) Use part a) to show that $\lim_{n\to\infty} \Pr\{N \geq n\} = 0$ and explain why this means that $N(t)$ is a rv, *i.e.*, is not defective.

c) Now suppose that the X_i do not have a finite mean. Consider truncating each X_i to $\breve{X}_i$, where for any given $b > 0$, $\breve{X}_i = \min(X_i, b)$. Let $\breve{N}(t)$ be the renewal counting process for the inter-renewal intervals $\breve{X}_i$. Show that $\breve{N}(t)$ is non-defective for each $t > 0$. Show that $N(t) \leq \breve{N}(t)$ and thus that $N(t)$ is non-defective. Note: Large inter-renewal intervals create small values of $N(t)$, and thus $\mathsf{E}[X] = \infty$ has nothing to do with potentially large values of $N(t)$, so the argument here was purely technical.

Exercise 4.2. The purpose of this exercise is to show that, for an arbitrary renewal process, $N(t)$, the number of renewals in $(0, t]$, has finite expectation.

a) Let the inter-renewal intervals have the distribution $\mathsf{F}_X(x)$, with, as usual, $\mathsf{F}_X(0) = 0$. Using whatever combination of mathematics and common sense is comfortable for you, show that numbers $\epsilon > 0$ and $\delta > 0$ must exist such that $\mathsf{F}_X(\delta) \leq 1 - \epsilon$. In other words, you are to show that a positive rv must take on some range of of values bounded away from zero with positive probability.

b) Show that $\Pr\{S_n \leq \delta\} \leq (1 - \epsilon)^n$.

c) Show that $\mathsf{E}[N(\delta)] \leq 1/\epsilon$.

d) Show that for every integer k, $\mathsf{E}[N(k\delta)] \leq k/\epsilon$ and thus that $\mathsf{E}[N(t)] \leq \frac{t+\delta}{\epsilon\delta}$ for any $t > 0$.

e) Use your result here to show that $N(t)$ is non-defective.

Exercise 4.3. Let $\{X_i; i \geq 1\}$ be the inter-renewal intervals of a renewal process generalized to allow for inter-renewal intervals of size 0 and let $\Pr\{X_i = 0\} = \alpha$, $0 < \alpha < 1$. Let $\{Y_i;\ i \geq 1\}$ be the sequence of non-zero interarrival intervals. For example, if $X_1 = x_1 > 0$, $X_2 = 0$, $X_3 = x_3 > 0, \dots,$ then $Y_1 = x_1$, $Y_2 = x_3$, $\dots$, .

a) Find the distribution function of each Y_i in terms of that of the X_i.

b) Find the PMF of the number of arrivals of the generalized renewal process at each epoch at which arrivals occur.

c) Explain how to view the generalized renewal process as an ordinary renewal process with inter-renewal intervals $\{Y_i;\ i \geq 1\}$ and bulk arrivals at each renewal epoch.

d) When a generalized renewal process is viewed as an ordinary renewal process with bulk arrivals, what is the distribution of the bulk arrivals? (The point of this part is to illustrate that bulk arrivals on an ordinary renewal process are considerably more general than generalized renewal processes.)

Exercise 4.4. Is it true for a renewal process that:

a) $N(t) < n$ if and only if $S_n > t$?

b) $N(t) \leq n$ if and only if $S_n \geq t$?

c) $N(t) > n$ if and only if $S_n < t$?

Exercise 4.5. (This shows that convergence WP1 implies convergence in probability.) Let $\{Y_n;\, n \geq 1\}$ be a sequence of rv's that converges to 0 WP1. For any positive integers m and k, let

$$A(m,k) = \{\omega : |Y_n(\omega)| \leq 1/k \quad \text{for all } n \geq m\}.$$

a) Show that if $\lim_{n\to\infty} Y_n(\omega) = 0$ for some given ω, then (for any given k) $\omega \in A(m,k)$ for some positive integer m.

b) Show that for all $k \geq 1$

$$\Pr\Big\{\bigcup_{m=1}^{\infty} A(m,k)\Big\} = 1.$$

c) Show that, for all $m \geq 1$, $A(m,k) \subseteq A(m{+}1,\, k)$. Use this (plus (1.9)) to show that

$$\lim_{m\to\infty} \Pr\{A(m,k)\} = 1.$$

d) Show that if $\omega \in A(m,k)$, then $|Y_m(\omega)| \leq 1/k$. Use this (plus part c) to show that

$$\lim_{m\to\infty} \Pr\{|Y_m| > 1/k\} = 0.$$

Since $k \geq 1$ is arbitrary, this shows that $\{Y_n;\, n \geq 1\}$ converges in probabiity.

Exercise 4.6. In this exercise, we find an explicit expression for $\{\omega : \lim_n Y_n = 0\}$. You may use whatever level of mathematical precision you feel comfortable with.

a) Let $\{Y_n;\, n \geq 1\}$ be a sequence of rv's. Using the definition of convergence for a sequence of numbers, justify the following set equivalences:

$$\begin{aligned}
\{\omega : \lim_n Y_n(\omega) = 0\} &= \bigcap_{k=1}^{\infty} \{\omega : \text{there exists an } m \text{ such that } |Y_n(\omega)| \leq 1/k \text{ for all } n \geq m\} \\
&= \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \{\omega : Y_n(\omega) \leq 1/k \text{ for all } n \geq m\} \\
&= \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \{\omega : Y_n(\omega) \leq 1/k\}
\end{aligned}$$

b) Explain how this shows that $\{\omega : \lim_n Y_n(\omega) = 0\}$ must be an event.

c) Use deMorgan's laws to show that the complement of the above equivalence is

$$\{\omega : \lim_n Y_n(\omega) = 0\}^{\mathsf{c}} = \bigcup_{k=1}^{\infty} \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{\omega : Y_n(\omega) > 1/k\}$$

d) Show that for $\{Y_n;\, n \geq 1\}$ to converge WP1, it is necessary and sufficient to satisfy

$$\Pr\Big\{\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{Y_n > 1/k\}\Big\} = 0 \qquad \text{for all } k \geq 1$$

e) Show that for $\{Y_n;\, n \geq 1\}$ to converge WP1, it is necessary and sufficient to satisfy

$$\lim_{m\to\infty} \Pr\Big\{\bigcup_{n=m}^{\infty}\{Y_n > 1/k\}\Big\} = 0 \qquad \text{for all } k \geq 1$$

Hint: Use part a) of Exercise 4.7. Note: Part e) provides an equivalent condition that is often useful in establishing convergence WP1. It also brings out quite clearly the difference between convergence WP1 and convergence in probability.

Exercise 4.7. Consider the event $\bigcap_{m\geq 1}\bigcup_{n\geq m} A_n$ where $A_1, A_2, \dots$, are arbitrary events.

a) Show that

$$\lim_{m\to\infty} \Pr\Big\{\bigcup_{n\geq m} A_n\Big\} = 0 \qquad \Longleftrightarrow \qquad \Pr\Big\{\bigcap_{m\geq 1}\bigcup_{n\geq m} A_n\Big\} = 0.$$

Hint: Apply the complement of (1.9).

b) Show that if $\sum_{m=1}^{\infty} \Pr\{A_m\} < \infty$, then $\mathsf{Pr}\{\bigcap_m \bigcup_{n\geq m} A_n\} = 0$. Hint: Recall that if $\sum_{m=1}^{\infty} \Pr\{A_m\} < \infty$, then $\lim_{m\to\infty} \mathsf{Pr}\{\bigcup_{n\geq m} A_n\} = 0$. Combine this with a). This well-known result is called the Borel-Cantelli lemma.

c) The set $\mathsf{Pr}\{\bigcap_m \bigcup_{n\geq m} A_n\}$ is often referred to as the set of ω that are contained in infinitely many of the A_n. Without trying to be precise about what this latter statement means, explain why it is a good way to think about $\mathsf{Pr}\{\bigcap_m \bigcup_{n\geq m} A_n\}$. Hint: Consider an ω that is contained in some finite number k of the sets A_n and argue that there must be an integer m such that $\omega \notin A_n$ for all $n > m$.

Exercise 4.8. Let $\{X_i;\, i{\geq}1\}$ be the inter-renewal intervals of a renewal process and assume that $\mathsf{E}[X_i] = \infty$. Let $b > 0$ be an arbitrary number and $\breve{X}_i$ be a truncated random variable defined by $\breve{X}_i = X_i$ if $X_i \leq b$ and $\breve{X}_i = b$ otherwise.

a) Show that for any constant $M > 0$, there is a b sufficiently large so that $\mathsf{E}\big[\breve{X}_i\big] \geq M$.

b) Let $\{\breve{N}(t);\, t{\geq}0\}$ be the renewal counting process with inter-renewal intervals $\{\breve{X}_i; i \geq 1\}$ and show that for all $t > 0$, $\breve{N}(t) \geq N(t)$.

c) Show that for all sample functions $N(t,\omega)$, except a set of probability 0, $N(t,\omega)/t < 2/M$ for all sufficiently large t. Note: Since M is arbitrary, this means that $\lim N(t)/t = 0$ with probability 1.

Exercise 4.9. Let $Y(t) = S_{N(t)+1} - t$ be the residual life at time t of a renewal process. First consider a renewal process in which the interarrival time has density $\mathsf{f}_X(x) = e^{-x};\ x \geq 0$, and next consider a renewal process with density

$$\mathsf{f}_X(x) = \frac{3}{(x+1)^4}; \quad x \geq 0.$$

For each of the above densities, use renewal-reward theory to find:

i) the time-average of $Y(t)$

ii) the second moment in time of $Y(t)$ (i.e., $\lim_{T\to\infty} \frac{1}{T}\int_0^T Y^2(t)dt$)

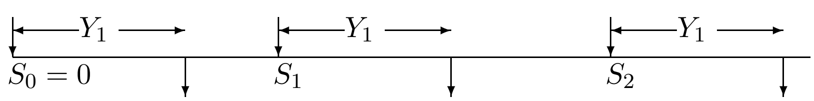
For the exponential density, verify your answers by finding $\mathsf{E}[Y(t)]$ and $\mathsf{E}\left[Y^2(t)\right]$ directly.

Exercise 4.10. Consider a variation of an M/G/1 queueing system in which there is no facility to save waiting customers. Assume customers arrive according to a Poisson process of rate λ. If the server is busy, the customer departs and is lost forever; if the server is not busy, the customer enters service with a service time distribution function denoted by $\mathsf{F}_Y(y)$.

Successive service times (for those customers that are served) are IID and independent of arrival times. Assume that customer number 0 arrives and enters service at time $t = 0$.

a) Show that the sequence of times $S_1, S_2, \dots$ at which successive customers enter service are the renewal times of a renewal process. Show that each inter-renewal interval $X_i = S_i - S_{i-1}$ (where $S_0 = 0$) is the sum of two independent random variables, $Y_i + U_i$ where Y_i is the ith service time; find the probability density of U_i.

b) Assume that a reward (actually a cost in this case) of one unit is incurred for each customer turned away. Sketch the expected reward function as a function of time for the sample function of inter-renewal intervals and service intervals shown below; the expectation is to be taken over those (unshown) arrivals of customers that must be turned away.



c) Let $\int_0^t R(\tau)d\tau$ denote the accumulated reward (i.e., cost) from 0 to t and find the limit as $t \to \infty$ of $(1/t)\int_0^t R(\tau)d\tau$. Explain (without any attempt to be rigorous or formal) why this limit exists with probability 1.

d) In the limit of large t, find the expected reward from time t until the next renewal. Hint: Sketch this expected reward as a function of t for a given sample of inter-renewal intervals and service intervals; then find the time-average.

e) Now assume that the arrivals are deterministic, with the first arrival at time 0 and the nth arrival at time $n-1$. Does the sequence of times $S_1, S_2, \dots$ at which subsequent customers start service still constitute the renewal times of a renewal process? Draw a sketch of arrivals, departures, and service time intervals. Again find $\lim_{t\to\infty}\left(\int_0^t R(\tau)\,d\tau\right)/t$.

Exercise 4.11. Let $Z(t) = t - S_{N(t)}$ be the age of a renewal process and $Y(t) = S_{N(t)+1} - t$ be the residual life. Let $\mathsf{F}_X(x)$ be the distribution function of the inter-renewal interval and find the following as a function of $\mathsf{F}_X(x)$:

a) $\Pr\{Y(t)>x \mid Z(t)=s\}$

b) $\Pr\{Y(t)>x \mid Z(t+x/2)=s\}$

c) $\Pr\{Y(t){>}x \mid Z(t{+}x){>}s\}$ for a Poisson process.

Exercise 4.12. Let $\mathsf{F}_Z(z)$ be the fraction of time (over the limiting interval $(0, \infty)$ that the age of a renewal process is at most z. Show that $\mathsf{F}_Z(z)$ satisfies

$$\mathsf{F}_Z(z) = \frac{1}{\overline{X}} \int_{x=0}^{y} \Pr\{X > x\}\, dx \qquad \text{WP1.}$$

Hint: Follow the argument in Example 4.4.5.

Exercise 4.13. **a)** Let J be a stopping rule and $\mathbb{I}_{\{J \geq n\}}$ be the indicator random variable of the event $\{J \geq n\}$. Show that $J = \sum_{n \geq 1} \mathbb{I}_{\{J \geq n}$.

b) Show that $\mathbb{I}_{J \geq 1} \geq \mathbb{I}_{J \geq 2} \geq \dots$, *i.e.*, show that for each $n > 1$, $\mathbb{I}_{J \geq n}(\omega) \geq \mathbb{I}_{J \geq n+1}(\omega)$ for each $\omega \in \Omega$ (except perhaps for a set of probability 0).

Exercise 4.14. **a)** Use Wald's equality to compute the expected number of trials of a Bernoulli process up to and including the kth success.

b) Use elementary means to find the expected number of trials up to and including the first success. Use this to find the expected number of trials up to and including the kth success. Compare with part a).

Exercise 4.15. A gambler with an initial finite capital of $d > 0$ dollars starts to play a dollar slot machine. At each play, either his dollar is lost or is returned with some additional number of dollars. Let X_i be his change of capital on the ith play. Assume that $\{X_i;\, i{=}1, 2, \dots\}$ is a set of IID random variables taking on integer values $\{-1, 0, 1, \dots\}$. Assume that $\mathsf{E}[X_i] < 0$. The gambler plays until losing all his money (i.e., the initial d dollars plus subsequent winnings).

a) Let J be the number of plays until the gambler loses all his money. Is the weak law of large numbers sufficient to argue that $\lim_{n \to \infty} \Pr\{J > n\} = 0$ (i.e., that J is a random variable) or is the strong law necessary?

b) Find $\mathsf{E}[J]$. Hint: The fact that there is only one possible negative outcome is important here.

Exercise 4.16. Let $\{X_i; i \geq 1\}$ be IID binary random variables with $P_X(0) = P_X(1) = 1/2$. Let J be a positive integer-valued random variable defined on the above sample space of binary sequences and let $S_J = \sum_{i=1}^{J} X_i$. Find the simplest example you can in which J is not a stopping trial for $\{X_i; i \geq 1\}$ and where $\mathsf{E}[X]\, \mathsf{E}[J] \neq \mathsf{E}[S_J]$. Hint: Try letting J take on only the values 1 and 2.

Exercise 4.17. Let $J = \min\{n \mid S_n{\leq}b \text{ or } S_n{\geq}a\}$, where a is a positive integer, b is a negative integer, and $S_n = X_1 + X_2 + \cdots + X_n$. Assume that $\{X_i;\, i{\geq}1\}$ is a set of zero mean IID rv's that can take on only the set of values $\{-1, 0, +1\}$, each with positive probability.

a) Is J a stopping rule? Why or why not? Hint: The more difficult part of this is to argue that J is a random variable (*i.e.*, non-defective); you do not need to construct a proof of this, but try to argue why it must be true.

b) What are the possible values of S_J?

c) Find an expression for $\mathsf{E}[S_J]$ in terms of p, a, and b, where $p = \Pr\{S_J \geq a\}$.

d) Find an expression for $\mathsf{E}[S_J]$ from Wald's equality. Use this to solve for p.

Exercise 4.18. Show that the interchange of expectation and sum in (4.32) is valid if $\mathsf{E}[J] < \infty$. Hint: First express the sum as $\sum_{n=1}^{k-1} X_n \mathbb{I}_{J \geq n} + \sum_{n=k}^{\infty} (X_n^+ + X_n^-) \mathbb{I}_{J \geq n}$ and then consider the limit as $k \to \infty$.

Exercise 4.19. Consider an amnesic miner trapped in a room that contains three doors. Door 1 leads him to freedom after two-day's travel; door 2 returns him to his room after four-day's travel; and door 3 returns him to his room after eight-day's travel. Suppose each door is equally likely to be chosen whenever he is in the room, and let T denote the time it takes the miner to become free.

a) Define a sequence of independent and identically distributed random variables $X_1, X_2, \ldots$ and a stopping rule J such that

$$T = \sum_{i=1}^{J} X_i.$$

b) Use Wald's equality to find $\mathsf{E}[T]$.

c) Compute $\mathsf{E}\left[\sum_{i=1}^{J} X_i \mid J{=}n\right]$ and show that it is not equal to $\mathsf{E}[\sum_{i=1}^{n} X_i]$.

d) Use part c) for a second derivation of $\mathsf{E}[T]$.

Exercise 4.20. a) Consider a renewal process for which the inter-renewal intervals have the PMF $p_X(1) = p_X(2) = 1/2$. Use elementary combinatorics to show that $m(1) = 1/2$, $m(2) = 5/4$, and $m(3) = 15/8$.

b) Use elementary means to show that $\mathsf{E}\left[S_{N(1)}\right] = 1/2$ and $\mathsf{E}\left[S_{N(1)+1}\right] = 9/4$. Verify (4.35) in this case (*i.e.*, for $t = 1$) and show that $N(1)$ is not a stopping trial. Note also that the expected duration, $\mathsf{E}\left[S_{N(1)+1} - S_{N(1)}\right]$ is not equal to $\overline{X}$.

c) Consider a more general form of part a) where $\Pr\{X = 1\} = 1 - p$ and $\Pr\{X = 2\} = p$. Let $\Pr\{W_n = 1\} = x_n$ and show that x_n satisfies the difference equation $x_n = 1 - p x_{n-1}$ for $n \geq 1$ where by convention $x_0 = 1$. Use this to show that

$$x_n = \frac{1 - (-p)^{n+1}}{1 + p}. \qquad (4.112)$$

From this, solve for $m(n)$ for $n \geq 1$.

Exercise 4.21. Let $\{N(t);\ t > 0\}$ be a renewal counting process generalized to allow for inter-renewal intervals $\{X_i\}$ of duration 0. Let each X_i have the PMF $\Pr\{X_i = 0\} = 1 - \epsilon$; $\Pr\{X_i = 1/\epsilon\} = \epsilon$.

a) Sketch a typical sample function of $\{N(t);\ t > 0\}$. Note that $N(0)$ can be non-zero (i.e., $N(0)$ is the number of zero interarrival times that occur before the first non-zero interarrival time).

b) Evaluate $\mathsf{E}[N(t)]$ as a function of t.

c) Sketch $\mathsf{E}[N(t)]/t$ as a function of t.

d) Evaluate $\mathsf{E}\left[S_{N(t)+1}\right]$ as a function of t (do this directly, and then use Wald's equality as a check on your work).

e) Sketch the lower bound $\mathsf{E}[N(t)]/t \geq 1/\mathsf{E}[X] - 1/t$ on the same graph with part c).

f) Sketch $\mathsf{E}\left[S_{N(t)+1} - t\right]$ as a function of t and find the time average of this quantity.

g) Evaluate $\mathsf{E}\left[S_{N(t)}\right]$ as a function of t; verify that $\mathsf{E}\left[S_{N(t)}\right] \neq \mathsf{E}[X]\,\mathsf{E}[N(t)]$.

Exercise 4.22. Let $\{N(t);\ t > 0\}$ be a renewal counting process and let $m(t) = \mathsf{E}[N(t)]$ be the expected number of arrivals up to and including time t. Let $\{X_i; i \geq 1\}$ be the inter-renewal times and assume that $\mathsf{F}_X(0) = 0$.

a) For all $x > 0$ and $t > x$ show that $\mathsf{E}[N(t)|X_1{=}x] = \mathsf{E}[N(t - x)] + 1$.

b) Use part a) to show that $m(t) = \mathsf{F}_X(t) + \int_0^t m(t - x)d\mathsf{F}_X(x)$ for $t > 0$. This equation is the renewal equation derived differently in (4.52).

c) Suppose that X is an exponential random variable of parameter λ. Evaluate $L_m(s)$ from (4.54); verify that the inverse Laplace transform is $\lambda t;\ t{\geq}0$.

Exercise 4.23. a) Let the inter-renewal interval of a renewal process have a second order Erlang density, $\mathsf{f}_X(x) = \lambda^2 x \exp(-\lambda x)$. Evaluate the Laplace transform of $m(t) = \mathsf{E}[N(t)]$.

b) Use this to evaluate $m(t)$ for $t \geq 0$. Verify that your answer agrees with (4.56).

c) Evaluate the slope of $m(t)$ at $t = 0$ and explain why that slope is not surprising.

d) View the renewals here as being the even numbered arrivals in a Poisson process of rate λ. Sketch $m(t)$ for the process here and show one half the expected number of arrivals for the Poisson process on the same sketch. Explain the difference between the two.

Exercise 4.24. a) Let $N(t)$ be the number of arrivals in the interval $(0, t]$ for a Poisson process of rate λ. Show that the probability that $N(t)$ is even is $[1 + \exp(-2\lambda t)]/2$. Hint: Look at the power series expansion of $\exp(-\lambda t)$ and that of $\exp(\lambda t)$, and look at the sum of the two. Compare this with $\sum_{n \text{ even}} \Pr\{N(t) = n\}$.

b) Let $\widetilde{N}(t)$ be the number of even numbered arrivals in $(0, t]$. Show that $\widetilde{N}(t) = N(t)/2 - \mathbb{I}_{\text{odd}}(t)/2$ where $\mathbb{I}_{\text{odd}}(t)$ is a random variable that is 1 if $N(t)$ is odd and 0 otherwise.

c) Use parts a) and b) to find $\mathsf{E}\left[\widetilde{N}(t)\right]$. Note that this is $m(t)$ for a renewal process with 2nd order Erlang inter-renewal intervals.

Exercise 4.25. **a)** Consider a function $r(z) \geq 0$ defined as follows for $0 \leq z < \infty$: For each integer $n \geq 1$ and each integer $k, 1 \leq k < n$, $r(z) = 1$ for $n+k/n \leq z \leq n+k/n+2^{-n}$. For all other z, $r(z) = 0$. Sketch this function and show that $r(z)$ is not directly Riemann integrable.

b) Evaluate the Riemann integral $\int_0^\infty r(z)dz$.

c) Suppose $r(z)$ is decreasing, *i.e.*, that $r(z) \geq r(y)$ for all $y > z > 0$. Show tht if $r(z)$ is Riemann integrable, it is also directly Riemann integrable.

d) Suppose $f(z) \geq 0$, defined for $z \geq 0$, is decreasing and Riemann integrable and that $f(z) \geq r(z)$ for all $z \geq 0$. Show that $r(z)$ is Directly Riemann integrable.

e) Let X be a non-negative rv for which $\mathsf{E}\left[X^2\right] < \infty$. Show that $x\,\mathsf{F}^{\mathsf{c}}_X(x)$ is directly Riemann integrable. Hint: Consider $y\mathsf{F}^{\mathsf{c}}_X(y) + \int_y^\infty \mathsf{F}_X(x)\,dx$ and use Figure 1.7 (or use integration by parts) to show that this expression is decreasing in y.

Exercise 4.26. Let $Z(t), Y(t), \widetilde{X}(t)$ denote the age, residual life, and duration at time t for a renewal counting process $\{N(t);\, t > 0\}$ in which the interarrival time has a density given by $f(x)$. Find the following probability densities; assume steady state.

a) $\mathsf{f}_{Y(t)}(y \,|\, Z(t{+}s/2){=}s)$ for given $s > 0$.

b) $\mathsf{f}_{Y(t),Z(t)}(y, z)$.

c) $\mathsf{f}_{Y(t)}(y \,|\, \widetilde{X}(t){=}x)$.

d) $\mathsf{f}_{Z(t)}(z \,|\, Y(t{-}s/2){=}s)$ for given $s > 0$.

e) $\mathsf{f}_{Y(t)}(y \,|\, Z(t{+}s/2){\geq}s)$ for given $s > 0$.

Exercise 4.27. **a)** Find $\lim_{t\to\infty}\{\mathsf{E}[N(t)] - t/\overline{X}\}$ for a renewal counting process $\{N(t);\, t > 0\}$ with inter-renewal times $\{X_i; i \geq 1\}$. Hint: use Wald's equation.

b) Evaluate your result for the case in which X is an exponential random variable (you already know what the result should be in this case).

c) Evaluate your result for a case in which $\mathsf{E}[X] < \infty$ and $\mathsf{E}\left[X^2\right] = \infty$. Explain (very briefly) why this does not contradict the elementary renewal theorem.

Exercise 4.28. Customers arrive at a bus stop according to a Poisson process of rate λ. Independently, buses arrive according to a renewal process with the inter-renewal interval distribution $\mathsf{F}_X(x)$. At the epoch of a bus arrival, all waiting passengers enter the bus and the bus leaves immediately. Let $R(t)$ be the number of customers waiting at time t.

a) Draw a sketch of a sample function of $R(t)$.

b) Given that the first bus arrives at time $X_1 = x$, find the expected number of customers picked up; then find $\mathsf{E}\left[\int_0^x R(t)dt\right]$, again given the first bus arrival at $X_1 = x$.

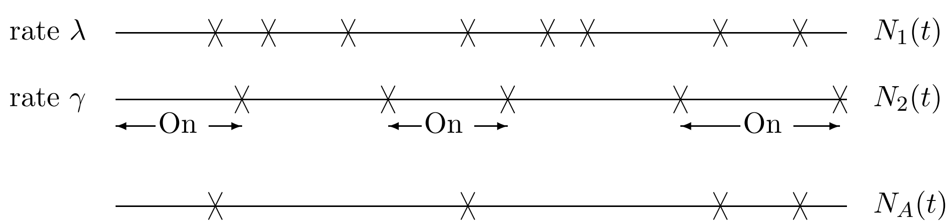
c) Find $\lim_{t\to\infty} \frac{1}{t}\int_0^t R(\tau)d\tau$ (with probability 1). Assuming that F_X is a non-arithmetic distribution, find $\lim_{t\to\infty}\mathsf{E}[R(t)]$. Interpret what these quantities mean.

d) Use part c) to find the time-average expected wait per customer.

e) Find the fraction of time that there are no customers at the bus stop. (Hint: this part is independent of a), b), and c); check your answer for $\mathsf{E}[X] \ll 1/\lambda$).

Exercise 4.29. Consider the same setup as in Exercise 4.28 except that now customers arrive according to a non-arithmetic renewal process independent of the bus arrival process. Let $1/\lambda$ be the expected inter-renewal interval for the customer renewal process. Assume that both renewal processes are in steady state (i.e., either we look only at $t \gg 0$, or we assume that they are equilibrium processes). Given that the nth customer arrives at time t, find the expected wait for customer n. Find the expected wait for customer n without conditioning on the arrival time.

Exercise 4.30. Let $\{N_1(t);\, t > 0\}$ be a Poisson counting process of rate λ. Assume that the arrivals from this process are switched on and off by arrivals from a non-arithmetic renewal counting process $\{N_2(t);\, t > 0\}$ (see figure below). The two processes are independent.



Let $\{N_A(t); t \geq 0\}$ be the switched process; that is $N_A(t)$ includes arrivals from $\{N_1(t);\, t > 0\}$ while $N_2(t)$ is even and excludes arrivals from $\{N_1(t);\, t > 0\}$ while $N_2(t)$ is odd.

a) Is $N_A(t)$ a renewal counting process? Explain your answer and if you are not sure, look at several examples for $N_2(t)$.

b) Find $\lim_{t\to\infty} \frac{1}{t} N_A(t)$ and explain why the limit exists with probability 1. Hint: Use symmetry—that is, look at $N_1(t) - N_A(t)$. To show why the limit exists, use the renewal-reward theorem. What is the appropriate renewal process to use here?

c) Now suppose that $\{N_1(t);\, t{>}0\}$ is a non-arithmetic renewal counting process but not a Poisson process and let the expected inter-renewal interval be $1/\lambda$. For any given δ, find $\lim_{t\to\infty} \mathsf{E}[N_A(t+\delta) - N_A(t)]$ and explain your reasoning. Why does your argument in (b) fail to demonstrate a time-average for this case?

Exercise 4.31. An M/G/1 queue has arrivals at rate λ and a service time distribution given by $\mathsf{F}_Y(y)$. Assume that $\lambda < 1/\mathsf{E}[Y]$. Epochs at which the system becomes empty define a renewal process. Let $\mathsf{F}_Z(z)$ be the distribution of the inter-renewal intervals and let $\mathsf{E}[Z]$ be the mean inter-renewal interval.

a) Find the fraction of time that the system is empty as a function of λ and $\mathsf{E}[Z]$. State carefully what you mean by such a fraction.

b) Apply Little's theorem, not to the system as a whole, but to the number of customers in the server (i.e., 0 or 1). Use this to find the fraction of time that the server is busy.

c) Combine your results in a) and b) to find $\mathsf{E}[Z]$ in terms of λ and $\mathsf{E}[Y]$; give the fraction of time that the system is idle in terms of λ and $\mathsf{E}[Y]$.

d) Find the expected duration of a busy period.

Exercise 4.32. Consider a sequence $X_1, X_2, \dots$ of IID binary random variables. Let p and $1-p$ denote $\Pr\{X_m = 1\}$ and $\Pr\{X_m = 0\}$ respectively. A *renewal* is said to occur at time m if $X_{m-1} = 0$ and $X_m = 1$.

a) Show that $\{N(m); m \geq 0\}$ is a renewal counting process where $N(m)$ is the number of renewals up to and including time m and $N(0)$ and $N(1)$ are taken to be 0.

b) What is the probability that a renewal occurs at time m, $m \geq 2$?

c) Find the expected inter-renewal interval; use Blackwell's theorem here.

d) Now change the definition of renewal; a renewal now occurs at time m if $X_{m-1} = 1$ and $X_m = 1$. Show that $\{N_m^*; m \geq 0\}$ is a delayed renewal counting process where N_m^* is the number of renewals up to and including m for this new definition of renewal ($N_0^* = N_1^* = 0$).

e) Find the expected inter-renewal interval for the renewals of part d).

f) Given that a renewal (according to the definition in (d)) occurs at time m, find the expected time until the next renewal, conditional, first, on $X_{m+1} = 1$ and, next, on $X_{m+1} = 0$. Hint: use the result in e) plus the result for $X_{m+1} = 1$ for the conditioning on $X_{m+1} = 0$.

g) Use your result in f) to find the expected interval from time 0 to the first renewal according to the renewal definition in d).

h) Which pattern requires a larger expected time to occur: 0011 or 0101

i) What is the expected time until the first occurrence of 0111111?

Exercise 4.33. A large system is controlled by n identical computers. Each computer independently alternates between an operational state and a repair state. The duration of the operational state, from completion of one repair until the next need for repair, is a random variable X with finite expected duration $\mathsf{E}[X]$. The time required to repair a computer is an exponentially distributed random variable with density $\lambda e^{-\lambda t}$. All operating durations and repair durations are independent. Assume that all computers are in the repair state at time 0.

a) For a single computer, say the ith, do the epochs at which the computer enters the repair state form a renewal process? If so, find the expected inter-renewal interval.

b) Do the epochs at which it enters the operational state form a renewal process?

c) Find the fraction of time over which the ith computer is operational and explain what you mean by fraction of time.

d) Let $Q_i(t)$ be the probability that the ith computer is operational at time t and find $\lim_{t\to\infty} Q_i(t)$.

e) The system is in failure mode at a given time if all computers are in the repair state at that time. Do the epochs at which system failure modes begin form a renewal process?

f) Let $\Pr\{t\}$ be the probability that the the system is in failure mode at time t. Find $\lim_{t\to\infty} \Pr\{t\}$. Hint: look at part d).

g) For δ small, find the probability that the system enters failure mode in the interval $(t, t+\delta]$ in the limit as $t \to \infty$.

h) Find the expected time between successive entries into failure mode.

i) Next assume that the repair time of each computer has an arbitrary density rather than exponential, but has a mean repair time of $1/\lambda$. Do the epochs at which system failure modes begin form a renewal process?

j) Repeat part f) for the assumption in (i).

Exercise 4.34. Let $\{N_1(t); t>0\}$ and $\{N_2(t); t>0\}$ be independent renewal counting processes. Assume that each has the same distribution function $\mathsf{F}(x)$ for interarrival intervals and assume that a density $f(x)$ exists for the interarrival intervals.

a) Is the counting process $\{N_1(t) + N_2(t);\ t > 0\}$ a renewal counting process? Explain.

b) Let $Y(t)$ be the interval from t until the first arrival (from either process) after t. Find an expression for the distribution function of $Y(t)$ in the limit $t \to \infty$ (you may assume that time averages and ensemble-averages are the same).

c) Assume that a reward R of rate 1 unit per second starts to be earned whenever an arrival from process 1 occurs and ceases to be earned whenever an arrival from process 2 occurs.

Assume that $\lim_{t\to\infty}(1/t)\int_0^t R(\tau)\,d\tau$ exists with probability 1 and find its numerical value.

d) Let $Z(t)$ be the interval from t until the first time after t that $R(t)$ (as in part c) changes value. Find an expression for $\mathsf{E}[Z(t)]$ in the limit $t \to \infty$. Hint: Make sure you understand why $Z(t)$ is not the same as $Y(t)$ in part b). You might find it easiest to first find the expectation of $Z(t)$ conditional on both the duration of the $\{N_1(t);\ t > 0\}$ interarrival interval containing t and the duration of the $\{N_2(t); t \geq 0\}$ interarrival interval containing t; draw pictures!

Exercise 4.35. This problem provides another way of treating ensemble-averages for renewal-reward problems. Assume for notational simplicity that X is a continuous random variable.

a) Show that $\Pr\{\text{one or more arrivals in } (\tau, \tau+\delta)\} = m(\tau+\delta) - m(\tau) - o(\delta)$ where $o(\delta) \geq 0$ and $\lim_{\delta\to 0} o(\delta)/\delta = 0$.

b) Show that $\Pr\Big\{Z(t) \in [z, z+\delta),\ \widetilde{X}(t) \in (x, x+\delta)\Big\}$ is equal to $[m(t-z) - m(t-z-\delta) - o(\delta)][\mathsf{F}_X(x+\delta) - \mathsf{F}_X(x)]$ for $x \geq z+\delta$.

c) Assuming that $m'(\tau) = dm(\tau)/d\tau$ exists for all τ, show that the joint density of $Z(t)$, $\widetilde{X}(t)$ is $\mathsf{f}_{Z(t),\widetilde{X}(t)}(z,x) = m'(t-z)\mathsf{f}_X(x)$ for $x > z$.

d) Show that $\mathsf{E}[R(t)] = \int_{z=0}^{t}\int_{x=z}^{\infty} \mathcal{R}(z,x)\mathsf{f}_X(x)dx\, m'(t-z)dz$

Exercise 4.36. This problem is designed to give you an alternate way of looking at ensemble-averages for renewal-reward problems. First we find an exact expression for $\Pr\{S_{N(t)} > s\}$. We find this for arbitrary s and t, $0 < s < t$.

a) By breaking the event $\{S_{N(t)} > s\}$into subevents $\{S_{N(t)} > s, N(t) = n\}$, explain each of the following steps:

$$\begin{aligned}
\Pr\{S_{N(t)} > s\} &= \sum_{n=1}^{\infty} \Pr\{t \geq S_n > s, S_{n+1} > t\} \\
&= \sum_{n=1}^{\infty} \int_{y=s}^{t} \Pr\{S_{n+1}{>}t \mid S_n{=}y\}\, d\mathsf{F}_{S_n}(y) \\
&= \int_{y=s}^{t} \mathsf{F}^{\mathsf{c}}_X(t{-}y)\, d\sum_{n=1}^{\infty} \mathsf{F}_{S_n}(y) \\
&= \int_{y=s}^{t} \mathsf{F}^{\mathsf{c}}_X(t{-}y)\, dm(y) \quad \text{where } m(y) = \mathsf{E}\left[N(y)\right].
\end{aligned}$$

b) Show that for $0 < s < t < u$,

$$\Pr\{S_{N(t)} > s, S_{N(t)+1} > u\} = \int_{y=s}^{t} \mathsf{F}^{\mathsf{c}}_X(u - y)\, dm(y).$$

c) Draw a two dimensional sketch, with age and duration as the axes, and show the region of (age, duration) values corresponding to the event $\{S_N(t) > s,\ S_{N(t)+1} > u\}$.

d) Assume that for large t, $dm(y)$ can be approximated (according to Blackwell) as $(1/\overline{X})dy$, where $\overline{X} = \mathsf{E}[X]$. Assuming that X also has a density, use the result in parts b) and c) to find the joint density of age and duration.

Exercise 4.37. In this problem, we show how to calculate the residual life distribution $Y(t)$ as a transient in t. Let $\mu(t) = dm(t)/dt$ where $m(t) = \mathsf{E}[N(t)]$, and let the interarrival distribution have the density $\mathsf{f}_X(x)$. Let $Y(t)$ have the density $\mathsf{f}_{Y(t)}(y)$.

a) Show that these densities are related by the integral equation

$$\mu(t + y) = \mathsf{f}_{Y(t)}(y) + \int_{u=0}^{y} \mu(t + u)\mathsf{f}_X(y - u)du.$$

b) Let $L_{\mu,t}(r) = \int_{y\geq 0} \mu(t + y)e^{-ry}dy$ and let $L_{Y(t)}(r)$ and $L_X(r)$ be the Laplace transforms of $\mathsf{f}_{Y(t)}(y)$ and $\mathsf{f}_X(x)$ respectively. Find $L_{Y(t)}(r)$ as a function of $L_{\mu,t}$ and L_X.

c) Consider the inter-renewal density $\mathsf{f}_X(x) = (1/2)e^{-x} + e^{-2x}$ for $x \geq 0$ (as in Example 4.6.1). Find $L_{\mu,t}(r)$ and $L_{Y(t)}(r)$ for this example.

d) Find $\mathsf{f}_{Y(t)}(y)$. Show that your answer reduces to that of (4.28) in the limit as $t \to \infty$.

e) Explain how to go about finding $\mathsf{f}_{Y(t)}(y)$ in general, assuming that f_X has a rational Laplace transform.

Exercise 4.38. Show that for a G/G/1 queue, the time-average wait in the system is the same as $\lim_{n\to\infty} \mathsf{E}[W_n]$. Hint: Consider an integer renewal counting process $\{M(n); n \geq 0\}$ where $M(n)$ is the number of renewals in the G/G/1 process of Section 4.5.3 that have occurred by the nth arrival. Show that this renewal process has a span of 1. Then consider $\{W_n; n \geq 1\}$ as a reward within this renewal process.

Exercise 4.39. If one extends the definition of renewal processes to include inter-renewal intervals of duration 0, with $\Pr\{X{=}0\} = \alpha$, show that the expected number of simultaneous renewals at a renewal epoch is $1/(1-\alpha)$, and that, for a non-arithmetic process, the probability of 1 or more renewals in the interval $(t, t+\delta]$ tends to $(1-\alpha)\delta/\mathsf{E}[X] + o(\delta)$ as $t \to \infty$.

Exercise 4.40. The purpose of this exercise is to show why the interchange of expectation and sum in the proof of Wald's equality is justified when $\mathsf{E}[J] < \infty$ but not otherwise. Let $X_1, X_2, \dots,$ be a sequence of IID rv's, each with the distribution F_X. Assume that $\mathsf{E}[|X|] < \infty$.

a) Show that $S_n = X_1 + \cdots + X_n$ is a rv for each integer $n > 0$. Note: S_n is obviously a mapping from the sample space to the real numbers, but you must show that it is finite with probability 1. Hint: Recall the additivity axiom for the real numbers.

b) Let J be a stopping trial for $X_1, X_2, \dots$. Show that $S_J = X_1 + \cdots X_J$ is a rv. Hint: Represent $\Pr\{S_J\}$ as $\sum_{n=1}^{\infty} \Pr\{J = n\} S_n$.

c) For the stopping trial J above, let $J^{(k)} = \min(J, k)$ be the stopping trial J truncated to integer k. Explain why the interchange of sum and expectation in the proof of Wald's equality is justified in this case, so $\mathsf{E}[S_{J^{(k)}}] = \overline{X}\mathsf{E}\left[J^{(k)}\right]$.

d) In parts d), e), and f), assume, in addition to the assumptions above, that $\mathsf{F}_X(0) = 0$, *i.e.*, that the X_i are positive rv's. Show that $\lim_{k\to\infty} \mathsf{E}[S_{J^{(k)}}] < \infty$ if $\mathsf{E}[J] < \infty$ and $\lim_{k\to\infty} \mathsf{E}[S_{J^{(k)}}] = \infty$ if $\mathsf{E}[J] = \infty$.

e) Show that

$$\Pr\{S_{J^{(k)}} > x\} \leq \Pr\{S_J > x\}$$

for all k, x.

f) Show that $\mathsf{E}[S_J] = \overline{X}\mathsf{E}[J]$ if $\mathsf{E}[J] < \infty$ and $\mathsf{E}[S_J] = \infty$ if $\mathsf{E}[J] = \infty$.

g) Now assume that X has both negative and positive values with nonzero probability and let $X^+ = \max(0, X)$ and $X^- = \min(X, 0)$. Express S_J as $S_J^+ + S_J^-$ where $S_J^+ = \sum_{i=1}^{J} X_i^+$ and $S_J^- = \sum_{i=1}^{J} X_i^-$. Show that $\mathsf{E}[S_J] = \overline{X}\mathsf{E}[J]$ if $\mathsf{E}[J] < \infty$ and that $\mathsf{E}[S_j]$ is undefined otherwise.

Exercise 4.41. This is a very simple exercise designed to clarify confusion about the roles of past, present, and future in stopping rules. Let $\{X_n;\ n \geq 1\}$ be a sequence of IID binary rv's , each with the pmf $\mathsf{p}_X(1) = 1/2$, $\mathsf{p}_X(0) = 1/2$. Let J be a positive integer-valued rv that takes on the sample value n of the first trial for which $X_n = 1$. That is, for each $n \geq 1$,

$$\{J = n\} = \{X_1{=}0,\ X_2{=}0, \dots, X_{n-1}{=}0,\ X_n{=}1\}.$$

a) Use the definition of stopping trial, Definition 4.5.1 in the text, to show that J is a stopping trial for $\{X_n;\ n \geq 1\}$.

b) Show that for any given n, the rv's X_n and $\mathbb{I}_{J=n}$ are *statistically dependent*.

c) Show that for every $m > n$, X_n and $\mathbb{I}_{J=m}$ are *statistically dependent*.

d) Show that for every $m < n$, X_n and $\mathbb{I}_{J=m}$ are *statistically independent*.

e) Show that X_n and $\mathbb{I}_{J\geq n}$ are *statistically independent*. Give the simplest characterization you can of the event $\{J \geq n\}$.

f) Show that X_n and $\mathbb{I}_{J>n}$ are *statistically dependent*.

Note: The results here are characteristic of most sequences of IID rv's. For most people, this requires some realignment of intuition, since $\{J{\geq}n\}$ is the union of $\{J = m\}$ for all $m \geq n$, and all of these events are highly dependent on X_n. The right way to think of this is that $\{J{\geq}n\}$ is the complement of $\{J{<}n\}$, which is determined by $X_1, \dots, X_{n-1}$. Thus $\{J{\geq}n\}$ is also determined by $X_1, \dots, X_{n-1}$ and is thus independent of X_n. The moral of the story is that thinking of stopping rules as rv's independent of the future is very tricky, even in totally obvious cases such as this.

Exercise 4.42. Assume a friend has developed an excellent program for finding the steady-state probabilities for finite-state Markov chains. More precisely, given the transition matrix [P], the program returns $\lim_{n\to\infty} P_{ii}^n$ for each i. Assume all chains are aperiodic.

a) You want to find the expected time to first reach a given state k starting from a different state m for a Markov chain with transition matrix $[P]$. You modify the matrix to $[P']$ where $P'_{km} = 1$, $P'_{kj} = 0$ for $j \neq m$, and $P'_{ij} = P_{ij}$ otherwise. How do you find the desired first-passage time from the program output given $[P']$ as an input? (Hint: The times at which a Markov chain enters any given state can be considered as renewals in a (perhaps delayed) renewal process).

b) Using the same $[P']$ as the program input, how can you find the expected number of returns to state m before the first passage to state k?

c) Suppose, for the same Markov chain $[P]$ and the same starting state m, you want to find the probability of reaching some given state n before the first-passage to k. Modify $[P]$ to some $[P'']$ so that the above program with P'' as an input allows you to easily find the desired probability.

d) Let $\Pr\{X(0) = i\} = Q_i$, $1 \leq i \leq \mathsf{M}$ be an arbitrary set of initial probabilities for the same Markov chain $[P]$ as above. Show how to modify $[P]$ to some $[P''']$ for which the steady-state probabilities allow you to easily find the expected time of the first-passage to state k.

Exercise 4.43. Consider a ferry that carries cars across a river. The ferry holds an integer number k of cars and departs the dock when full. At that time, a new ferry immediately appears and begins loading newly arriving cars ad infinitum. The ferry business has been good, but customers complain about the long wait for the ferry to fill up.

a) Assume that cars arrive according to a renewal process. The IID interarrival times have mean $\overline{X}$, variance σ^2 and moment generating function $g_X(r)$. Does the sequence of departure times of the ferries form a renewal process? Explain carefully.

b) Find the expected time that a customer waits, starting from its arrival at the ferry terminal and ending at the departure of its ferry. Note 1: Part of the problem here is to give a reasonable definition of the expected customer waiting time. Note 2: It might be useful to consider $k = 1$ and $k = 2$ first.

c Is there a 'slow truck' phenomenon (a dependence on $\mathsf{E}\left[X^2\right]$) here? Give an intuitive explanation. Hint: Look at $k = 1$ and $k = 2$ again.

d) In an effort to decrease waiting, the ferry managers institute a policy where no customer ever has to wait more than one hour. Thus, the first customer to arrive after a ferry departure waits for either one hour or the time at which the ferry is full, whichever comes first, and then the ferry leaves and a new ferry starts to accumulate new customers. Does the sequence of ferry departures form a renewal process under this new system? Does the sequence of times at which each successive empty ferry is entered by its first customer form a renewal process? You can assume here that $t = 0$ is the time of the first arrival to the first ferry. Explain carefully.

e) Give an expression for the expected waiting time of the first new customer to enter an empty ferry under this new strategy.