Chapter 3

Renewal Theory

Consider an irreducible recurrent DTMC $Y_n, n \geq 0$, with $Y_0 = i$, and consider visits to the state j. Let X_1 denote the time until the first visit, and let $X_k, k \geq 2$, denote the subsequent intervisit times. We recall from Chapter 2 that $P(X_1 = m|Y_0 = i) = f_{ij}^{(m)}$, and, for $k \geq 2$, $P(X_k = m) = f_{jj}^{(m)}$. Also, let $Z_k, k \geq 1$, denote the time of the kth visit to state j. Now for any $k \geq 1$, Z_k is a stopping time for the process Y_n. It then easily follows that (see Exercise 2.5)

$$P(X_{k_1} = m_1, X_{k_2} = m_2, \cdots, X_{k_n} = m_n|Y_0 = i) \quad = \quad \begin{cases} f_{ij}^{(m_1)} f_{jj}^{(m_2)} \cdots f_{jj}^{(m_n)} & \text{for } k_1 = 1 \\ f_{jj}^{(m_1)} f_{jj}^{(m_2)} \cdots f_{jj}^{(m_n)} & \text{for } k_1 > 1 \end{cases}$$

Thus, we see that the sequence of intervisit times, $X_k, k \geq 1$, to the state j in an irreducible recurrent DTMC is a sequence of mutually independent random variables, with the $X_k, k \geq 2$, being a sequence of i.i.d. random variables. Such a sequence of times often arises in discrete event processes, and we call them *renewal life-times*. The associated instants Z_k are called *renewal instants.* This terminology is motivated by the analogy of a component in a system being repeatedly replaced after it gets worn out (for example, the batteries in a portable electronic device).

3.1 Definition and Some Related Processes

In general, a renewal process can be defined as follows.

Definition 3.1. *Given a sequence of mutually independent nonnegative real valued random variables, $X_k, k \geq 1$, with the random variables $X_k, k \geq 2$, also being identically distributed,*

(a.) this sequence of random variables is called the sequence of life-times *of the renewal process,*

(b.) for $k \geq 1$, $Z_k = \sum_{i=1}^{k} X_i$ *is the kth* renewal instant*; define* $Z_0 = 0$.

(c.) For $t \geq 0$, $M(t) = \sup\{k \geq 0 : Z_k \leq t\}$*, the number of renewals in* $[0, t]$*, is called the* renewal process*, and,*

(d.) for $t \geq 0$, $m(t) = \mathsf{E}(M(t))$ *is called the* renewal function.

■

Remarks 3.1.

(a.) In general, Z_k is a discrete parameter random process that takes nonnegative real values. Also, $M(t)$ is a continuous time random process that takes values in $\{0, 1, 2, 3, \cdots\}$. On the other hand $m(t)$ is just a nonnegative, real valued, nondecreasing, (deterministic) function of time.

(b.) Clearly, the processes Z_k and $M(t)$ are determined when a sample path of the life-times $X_k, k \geq 1$, is given. Thus we can think of each such sample path of the life-time process as an elementary outcome ω.

(c.) Note, from this definition, that $M(t)$ stays flat between renewal instants and increases in jumps at renewal instants. A renewal at time t is included in the process $M(t)$. Thus the sample paths of $M(t)$ are nondecreasing step functions that are right continuous at the jumps. It is possible to have multiple jumps at an instant; for example, it is possible that $Z_k = Z_{k+1}$ in which case $M(Z_{k+1}) = M(Z_k)$, and $M(Z_{k+1}) - M(Z_k-) = 2$.

■

Several interesting related processes can be defined, and some relationships between the processes can be observed.

(a.) $Z_{M(t)}$ is the instant of the last renewal in the interval $[0, t]$. Notice $M(t)$ is a random time (or random index) for the process Z_k, and $Z_{M(t)}$ is the value of the process Z_k at the random time $M(t)$. Hence, for a sample path ω, $Z_{M(t)}$ means $Z_{M(t,\omega)}(\omega)$. By definition, $Z_{M(t)} \leq t$.

(b.) $M(t) + 1$ is the index of the first renewal after t. Hence, $Z_{M(t)+1}$ is the first renewal instant (strictly) after t, i.e., the first renewal instant in (t, ∞). We have $Z_{M(t)+1} > t$.

(c.) For $t \geq 0$, $Y(t) = Z_{M(t)+1} - t$ is the *residual life-time* at time t. Thus $Y(0) = X_1$, $Y(t)$ decreases at 1 unit per unit time until Z_1, then $Y(t)$ jumps up by X_2, and so on. The sample paths of $Y(t)$ are right continuous at the jump instants $Z_k, k \geq 1$, and $Y(Z_k-) = 0, k \geq 1$.

(d.) For $t \geq 0$, define the *age* process $U(t)$ by $U(t) = t - Z_{M(t)}$. If $M(t) = 0$, since, by definition, $Z_0 = 0$, we have $U(t) = t$.

(e.) Notice that, for all $n \geq 1$, and $t \geq 0$, $(Z_n \leq t) = (M(t) \geq n)$, i.e., the event that the nth renewal occurs at or before t is the same as the event that there are at least n renewals in $[0, t]$. This is easily seen by checking that each ω that is in the event on the left hand side is in the event on the right hand side, and vice versa.

We now see how the renewal function, $m(t)$, can be expressed in terms of the life-time distributions. Let $A(\cdot)$ be the distribution of X_1 and $F(\cdot)$ that of $\{X_2, X_3, \ldots,\}$. It then follows that

$$\begin{aligned} P(Z_1 \leq t) &= A(t) \\ P(Z_2 \leq t) &= \int_0^t F(t-u) dA(u) \\ &= (A \star F)(t) \end{aligned}$$

where, as usual, $\star$ denotes the convolution of the distributions A and F[1]. Continuing, we have

$$\begin{aligned} P(Z_3 \leq t) &= \int_0^t F(t-u)\, d(A \star F)(u) \\ &= (A \star F \star F)(t) = (A \star F^{(2)})(t) \end{aligned}$$

and, for $n \geq 1$,

$$P(Z_n \leq t) = \left(A \star F^{(n-1)}\right)(t)$$

Thus we conclude that

$$\begin{aligned} P(M(t) \geq n) &= P(Z_n \leq t) \\ &= \left(A \star F^{(n-1)}\right)(t) \end{aligned}$$

Hence we can write the renewal function as follows

$$\begin{aligned} m(t) &= EM(t) \\ &= \sum_{n=1}^{\infty} P(M(t) \geq n) \end{aligned}$$

[1]If X and Y are independent random variables, with c.d.f.s $F(x)$ and $G(y)$, then the c.d.f. of $Z := X+Y$ is obtained as

$$P(Z \leq z) = P(X + Y \leq z) = \int_{u=0}^{\infty} F(z-u)\, dG(u) =: (F \star G)(z),$$

where $\star$ denotes convolution.

Thus we obtain the following expression for $m(t)$ in terms of the life-time distributions.

$$m(t) = \sum_{n=1}^{\infty} \left(A \star F^{(n-1)}\right)(t) \tag{3.1}$$

Based on this observation we can establish the following lemma that will be used in several results to follow.

Lemma 3.1. *If* $\mathsf{E}(X_j) > 0, j \geq 2$, *then* $m(t) < \infty$ *for all* t.

Remark: Note that for the nonnegative random variables X_j, $\mathsf{E}(X_j) > 0$ is equivalent to the statement $P(X_j > 0) > 0$, or $F_j(0^+) < 1$. Observe that this result implies that, under the hypotheses, $M(t) < \infty$ with probability 1.
Proof: We first observe that, for any $n \geq 1$, and $0 \leq m \leq n$,

$$\begin{aligned} F^{(n)}(t) &= \int_0^t F^{(n-m)}(t-u)\, dF^{(m)}(u) \\ &\leq F^{(n-m)}(t) \int_0^t dF^m(u) \\ &\leq F^{(n-m)}(t) F^{(m)}(t) \end{aligned} \tag{3.2}$$

where in the two inequalities we have used the fact that a distribution function is nondecreasing. Hence, applying the previous inequality recursively, for any $n \geq 1, r \geq 1, k \geq 0$,

$$\begin{aligned} F^{(nr+k)}(t) &\leq F^{((n-1)r+k)}(t) \cdot F^{(r)}(t) \\ &\leq \left(F^{(r)}(t)\right)^n F^{(k)}(t) \end{aligned} \tag{3.3}$$

Now, from Equation 3.1, we have, for any $r \geq 1$,

$$\begin{aligned} m(t) &= \sum_{n=1}^{\infty} \left(A \star F^{(n-1)}\right)(t) \\ &\leq \sum_{n=0}^{\infty} F^{(n)}(t) \\ &= \sum_{m=0}^{\infty} \underbrace{\sum_{k=0}^{r-1} F^{(mr+k)}(t)}_{\leq rF^{(mr)}(t)} \\ &\leq r \sum_{m=0}^{\infty} \left(F^{(r)}(t)\right)^m \end{aligned}$$

where the first inequality uses the same calculation that let to (3.2) and the second inequality uses (3.3). Now, for each t choose r such that $F^{(r)}(t) < 1$; then we see, from the last expression, that $m(t) < \infty$. Such a choice of t is possible by virtue of the hypothesis that $\mathsf{E}(X_2) > 0$. For then there exists $\epsilon > 0$ such that $F(\epsilon) < 1$. Now, for every n, $(1 - F^{(n)}(n\epsilon)) \geq (1 - F(\epsilon))^n > 0$; i.e., $F^{(n)}(n\epsilon) < 1$. It follows that $F^{\lceil \frac{t}{\epsilon} \rceil}(\epsilon \lceil \frac{t}{\epsilon} \rceil) < 1$. Hence, $F^{\lceil \frac{t}{\epsilon} \rceil}(t) < 1$. ■

3.2 The Elementary Renewal Theorem (ERT)

Theorem 3.1. *Given a sequence of mutually independent nonnegative random variables (life-times) $X_k, k \geq 1$, with $X_k, k \geq 2$, being identically distributed, such that,*

(i) for $k \geq 1$, $P(X_k < \infty) = 1$ (all the life-time random variables are proper),

(ii) $0 \leq \mathsf{E}(X_1) \leq \infty$ (we allow the mean of the first life-time to be 0 and also ∞), and,

(iii) for $k \geq 2$, $0 < \mathsf{E}(X_k) \leq \infty$ (the mean life-times after the first are positive, and possibly infinite, and, of course, identical to $\mathsf{E}(X_2)$). Defining $\mathsf{E}(X_2) = \frac{1}{\mu}$, this hypothesis is equivalent to $0 \leq \mu < \infty$.

Then the following conclusions hold

(a) $\lim_{t \to \infty} \frac{M(t)}{t} = \mu$ *almost surely*

(b) $\lim_{t \to \infty} \frac{m(t)}{t} = \mu$

■

Remark:

(a.) Conclusion (a) of the theorem states that the rate of renewals converges almost surely to μ. This is intuitive since, after the first renewal, which occurs in finite time with probability one, the subsequent interrenewal times have mean $\frac{1}{\mu}$. Since $t \to \infty$ the effect of the first life-time eventually vanishes.

(b.) In the second conclusion we simply have a limit of numbers (unlike in the first, where there is a sequence of random variables). Note that we cannot say that Conclusion (a) implies Conclusion (b) since this would involve the interchange of expectation and limit which is not always legal; see Section 1.4.1

Proof: *Part (a)*

We first observe that, since, for all $j \geq 1$, $P(X_j < \infty) = 1$, hence, for all $k \geq 1$, $P(\cap_{j=1}^{k}\{X_j < \infty\}) = 1$. Further, we observe that

$$\{Z_k < \infty\} = \cap_{j=1}^{k}\{X_j < \infty\}$$

which can be verified by checking that a sample point ω is in the event on the left if and only if it is in the event on the right. Hence it follows that, for all $k \geq 1$,

$$P(Z_k < \infty) = 1$$

i.e., every renewal occurs in finite time with probability one. From this observation it follows that

$$P(\lim_{t\to\infty} M(t) = \infty) = 1$$

for if this were not the case we would have positive probability that a renewal occurs at infinity. Let us now consider the case $\mu > 0$, i.e., $\mathsf{E}(X_2) < \infty$. Observe that

$$\begin{aligned} \frac{Z_n}{n} &= \frac{1}{n}\sum_{i=1}^{n} X_i \\ &= \frac{X_1}{n} + \frac{1}{n}\sum_{i=2}^{n} X_i \end{aligned}$$

Now since $P(X_1 < \infty) = 1$, it follows that $P(\frac{X_1}{n} \to 0) = 1$. Also, since $X_j, j \geq 2$, are i.i.d. with finite mean, by Theorem 1.10, $\frac{1}{n}\sum_{i=2}^{n} X_i \to \frac{1}{\mu}$ with probability 1. Since the intersection of two sets with probability 1 also has probability 1, it follows that, with probability 1,

$$\frac{Z_n}{n} \to \frac{1}{\mu}$$

Also, we saw that, $P(M(t) \to \infty) = 1$. Hence it further follows that

$$P\left(\frac{Z_{M(t)}}{M(t)} \to \frac{1}{\mu}\right) = 1$$

We have seen earlier that

$$Z_{M(t)} \leq t < Z_{M(t)+1}$$

Dividing across by $M(t)$ (for t large enough so that $M(t) > 0$), we have

$$\frac{Z_{M(t)}}{M(t)} \leq \frac{t}{M(t)} < \frac{Z_{M(t)+1}}{M(t)}$$

But $\frac{Z_{M(t)}}{M(t)} \to \frac{1}{\mu}$ with probability 1. Hence $\frac{t}{M(t)} \to \frac{1}{\mu}$ almost surely, from which the desired conclusion follows. We have proved the first conclusion of the theorem for the case $\mathsf{E}(X_2) < \infty$.

For the case $\mathsf{E}(X_2) = \infty$, i.e., $\mu = 0$, Theorem 1.10 cannot be directly used, and so we need to proceed as in Exercise 3.1.

To prove the second part of the theorem we need Wald's Lemma (Lemma 3.2). ■

Exercise 3.1.

Complete the proof of the first part of Theorem 3.1 for the case in which $\mathsf{E}(X_2) = \infty$. (Hint: use the truncated lifetime sequence $X_k^{(c)}, k \geq 1$, as defined in the proof of Part (b) of Theorem 3.1, provided later.) ■

Lemma 3.2 (Wald's Lemma)**.** *Let N be a stopping time for an infinite sequence of mutually independent random variables $X_i, i \geq 1$. If*

(i) $\mathsf{E}(N) < \infty$,

(ii) $\mathsf{E}(|X_n|) < B$*, a constant, for all* $n \geq 1$*, and*

(iii) $\mathsf{E}(X_n) = \mathsf{E}(X_1)$ *for all* $n \geq 1$,

then

$$\mathsf{E}\left(\sum_{n=1}^{N} X_n\right) = \mathsf{E}(X_1) \cdot \mathsf{E}(N)$$

Remarks 3.2.

(a.) We emphasise that the sequence of random variables, $X_i, i \geq 1$, in the statement of the lemma are not necessarily nonnegative. Further, the random variables need not have the same distribution. They, however, have the same mean. Note that a sequence of i.i.d. nonnegative random variables with finite common expectation (as would arise in a renewal theory application) satisfies conditions (ii) and (iii) of the lemma.

(b.) Let us also observe that a seemingly obvious calculation does not always work. Let $p_k = P(N = k), k \geq 1$, and define $S_N := \sum_{n=1}^{N} X_n$. Now

$$\begin{aligned} \mathsf{E}(S_N) &= \mathsf{E}\left(\mathsf{E}\left(\sum_{n=1}^{N} X_n | N\right)\right) \\ &= \sum_{k=1}^{\infty} p_k \, \mathsf{E}\left(\sum_{n=1}^{N} X_n | N = k\right) \end{aligned}$$

Now if we could write $\mathsf{E}\left(\sum_{n=1}^{N} X_n | N = k\right) = k\mathsf{E}(X_1)$ then it would immediately follow that $\mathsf{E}(S_N) = \mathsf{E}(X_1)\,\mathsf{E}(N)$. However, in general, this step is incorrect, because conditioning on $N = k$ may change the joint distribution of $X_i, 1 \leq i \leq k$. In fact, given that $N = k$ the random variables $X_i, 1 \leq i \leq k$, may also have become dependent.

(c.) As an illustration of the previous remark, consider a renewal process with i.i.d. lifetimes $X_i, i \geq 1$, and, for given $t > 0$, define the random time $N = M(t)$. Now look at $\sum_{n=1}^{M(t)} X_n$, and observe that, given $M(t) = k$, it must be that $\sum_{n=1}^{k} X_n \leq t$, so that the random variables $X_1, X_2, \cdots, X_k$ are conditionally dependent and also are conditionally bounded between 0 and t.

Proof: We can write

$$\begin{aligned} S_N &= \sum_{n=1}^{N} X_n \\ &= \sum_{n=1}^{\infty} X_n \, I_{\{n \leq N\}} \end{aligned}$$

Hence

$$\mathsf{E}(S_N) = \mathsf{E}\left(\sum_{n=1}^{\infty} X_n \, I_{\{n \leq N\}} \right) \tag{3.4}$$

Suppose we could exchange $\mathsf{E}(\cdot)$ and $\sum_{n=1}^{\infty}$ in the right hand side (we will justify this before ending the proof). This will yield

$$\mathsf{E}(S_N) = \sum_{n=1}^{\infty} \mathsf{E}\big(X_n I_{\{n \leq N\}}\big)$$

Now observe that

$$\begin{aligned} I_{\{n \leq N\}} &= 1 - I_{\{N \leq (n-1)\}} \\ &= f(X_1, X_2, \cdots, X_{n-1}) \end{aligned}$$

for some function $f(\cdots)$, since N is a stopping time for the sequence $X_i, i \geq 1$. But the $X_i, i \geq 1$, are mutually independent. It therefore follows that X_n is independent of $I_{\{n \leq N\}}$, and we obtain

$$\mathsf{E}(S_N) = \sum_{n=1}^{\infty} \mathsf{E}(X_n) \cdot P(N \geq n)$$

Using the fact that $\mathsf{E}(X_n) = \mathsf{E}(X_1)$, we obtain

$$\mathsf{E}(S_N) = \mathsf{E}(X_1)\, \mathsf{E}(N)$$

and the result is proved.

We finally turn to the justification of the exchange of $\mathsf{E}(\cdot)$ and $\sum_{n=1}^{\infty}$ in Equation 3.4. Define $Y = \sum_{n=1}^{\infty} |X_n| \; I_{\{n \leq N\}}$. It can be seen that, for all $m \geq 1$,

$|\sum_{n=1}^{m} X_n \, I_{\{n \leq N\}}| \leq Y$. Then using the facts that $\mathsf{E}(|X_n|) < B$ and $\mathsf{E}(N) < \infty$, and that N is a stopping time, it follows that $\mathsf{E}(Y) < \infty$. Dominated convergence theorem then applies and it follows that we can interchange $\mathsf{E}(\cdot)$ and $\sum_{n=1}^{\infty}$ in the right hand side of Equation 3.4.

■

Example 3.1 (The Poisson Process)**.**

Later in this chapter we will study an important renewal process called the Poisson process. This is a renewal process with i.i.d. life-times that are exponentially distributed with mean $\frac{1}{\lambda}$. Since $\sum_{i=1}^{M(t)} X_i \leq t$ and there is a positive probability that $\sum_{i=1}^{M(t)} X_i \leq t_1 < t$ (e.g., take $t_1 = \frac{t}{2}$), we can see that $\mathsf{E}\left(\sum_{i=1}^{M(t)} X_i\right) < t$. On the other hand, we will see that $M(t)$ is a Poisson distributed random variable with mean λt. Hence $\mathsf{E}(M(t)) \cdot \mathsf{E}(X_1) = \lambda t \, \frac{1}{\lambda} = t$, and the conclusion of Wald's Lemma does not hold. The reason is that $M(t)$ is not a stopping time for the renewal process (while all the other hypotheses of the lemma hold). To see this note that, for any $n \geq 1$,

$$\begin{aligned} I_{\{M(t) \leq n\}} &= 1 - I_{\{M(t) \geq n+1\}} \\ &= 1 - I_{\{Z_{n+1} \leq t\}} \\ &= 1 - I_{\{\sum_{i=1}^{n+1} X_i \leq t\}} \\ &= I_{\{\sum_{i=1}^{n+1} X_i > t\}} \end{aligned}$$

Hence to determine if $M(t) \leq n$ we need to look at X_{n+1} as well. Hence $M(t)$ is not a stopping time. This becomes evident when we observe that $M(t)$ is the index of the *last* complete life-time before t. ■

Corollary 3.1. *Let $M(t)$ be a renewal process with i.i.d. life-times $X_i, i \geq 1$, and with $0 < \mathsf{E}(X_1) < \infty$. Then $\mathsf{E}\big(Z_{M(t)+1}\big) = \mathsf{E}(X_1)\,(m(t) + 1)$.*

Proof: Define the random time $N = M(t) + 1$. Now observe that

$$\begin{aligned} I_{\{N \leq n\}} &= 1 - I_{\{N \geq n+1\}} \\ &= 1 - I_{\{M(t) \geq n\}} \\ &= 1 - I_{\{Z_n \leq t\}} \\ &= I_{\{Z_n > t\}} \\ &= I_{\{\sum_{i=1}^{n} X_i > t\}} \end{aligned}$$

Hence N is a stopping time for the life-times $X_i, i \geq 1$. Since $0 < \mathsf{E}(X_1)$, from Lemma 3.1 it follows that $m(t) < \infty$ for every t. Applying Wald's Lemma 3.2 we obtain the desired result, as follows

$$\mathsf{E}\big(Z_{M(t)+1}\big) = \mathsf{E}\left(\sum_{i=1}^{M(t)+1} X_i\right) = \mathsf{E}(X_1)\,(m(t)+1)$$

■

Example 3.2 (The Poisson Process (continued))**.**

Let us again consider the Poisson process. We will see later in this chapter that in a Poisson process at any time t the remaining time until the next renewal is exponentially distributed with mean $\frac{1}{\lambda}$. This is simply a consequence of the memoryless property of the exponential distribution. It follows that

$$\mathsf{E}\left(\sum_{i=1}^{M(t)+1} X_i\right) = \mathsf{E}(t + Y(t)) = t + \frac{1}{\lambda}$$

Further, since $M(t)$ is Poisson distributed with mean λt, we get

$$\mathsf{E}(M(t)+1)\,\mathsf{E}(X_1) = (\lambda t + 1)\frac{1}{\lambda} = t + \frac{1}{\lambda}$$

thus verifying that Wald's Lemma holds in this case. ■

We now continue the proof of Theorem 3.1.
Proof: Theorem 3.1, *Part (b)*

We first take the case: $\mathsf{E}(X_1) < \infty$ and $\mathsf{E}(X_2) < \infty$ (i.e., $\mu > 0$), and consider

$$Z_{M(t)+1} = \sum_{j=1}^{M(t)+1} X_j$$

Following arguments identical to the ones in the proof of Wald's Lemma, we can write (see the proof of Lemma 3.2)

$$\begin{aligned}
\mathsf{E}\big(Z_{M(t)+1}\big) &= \sum_{j=1}^{\infty} \mathsf{E}(X_j) \cdot P((M(t)+1) \geq j) \\
&= \mathsf{E}(X_1) \cdot P(M(t) \geq 0) + \sum_{j=2}^{\infty} \mathsf{E}(X_j) \cdot P(M(t)+1 \geq j) \\
&= \mathsf{E}(X_1) + \mathsf{E}(X_2)\,\mathsf{E}(M(t)) \\
&= \mathsf{E}(X_1) + m(t)\mathsf{E}(X_2) \\
&= \mathsf{E}(X_1) + \frac{m(t)}{\mu}
\end{aligned}$$

where in the first equality we have already used the fact that $M(t) + 1$ is a stopping time for the i.i.d. sequence $X_j, j \geq 2$. We have also used the fact that $\mathsf{E}(X_2) < \infty$, and that $\mathsf{E}(X_2) > 0$, which in turn implies (using Lemma 3.1) that $\mathsf{E}(M(t)) < \infty$.

Now we observe that $Z_{M(t)+1} > t$, by definition. Hence

$$\frac{m(t)}{\mu} > t - \mathsf{E}(X_1)$$

i.e.,

$$\frac{m(t)}{t} > \mu - \mu\frac{\mathsf{E}(X_1)}{t}$$

We conclude that

$$\liminf_{t\to\infty} \frac{m(t)}{t} \geq \mu$$

We will be done if we show that $\limsup_{t\to\infty} \frac{m(t)}{t} \leq \mu$. For this purpose, for each $c > 0$, define $X_j^{(c)}$ as

$$X_j^{(c)} = \begin{cases} X_j & \text{if } X_j \leq c \\ c & \text{if } X_j > c \end{cases}$$

Further, define $\mu^{(c)}$ by $\mathsf{E}\left(X_2^{(c)}\right) = \frac{1}{\mu^{(c)}}$. Now consider the renewal process generated by $\{X_j^{(c)},\ j \geq 1\}$; i.e., for any realisation of the life-times $X_j, j \geq 1$, this new renewal process has life-times obtained by truncating the life-times $X_j, j \geq 1$, as shown above. Use the superscript $^{(c)}$ to denote any process associated with this new renewal process. Then, clearly, $Z^{(c)}_{M^{(c)}(t)+1} \leq t + c$ since $X^{(c)}_{M^{(c)}(t)+1} \leq c$. Proceeding as above, we now obtain

$$\begin{aligned} \mathsf{E}\left(X_1^{(c)}\right) + \frac{m^{(c)}(t)}{\mu^{(c)}} &\leq t + c \\ \frac{m^{(c)}(t)}{t} &\leq \mu^{(c)} + \underbrace{\left(\frac{c - \mathsf{E}\left(X_1^{(c)}\right)}{t}\right)}_{\geq 0}\mu^{(c)} \end{aligned}$$

hence $\limsup_{t\to\infty} \frac{m^{(c)}(t)}{t} \leq \mu^{(c)}$.

But, for all $j \geq 1$, $X_j^{(c)} \leq X_j$, hence $Z_n^{(c)} \leq Z_n$, which implies that $M^{(c)}(t) \geq M(t)$ and hence that $m^{(c)}(t) \geq m(t)$. Thus we have

$$\limsup_{t\to\infty} \frac{m(t)}{t} \leq \limsup_{t\to\infty} \frac{m^{(c)}(t)}{t} \leq \mu^{(c)}$$

But $\mu^{(c)} \to \mu$ as $c \to \infty$, hence

$$\limsup_{t\to\infty} \frac{m(t)}{t} \leq \mu$$

now

$$\mu \leq \liminf_{t\to\infty} \frac{m(t)}{t} \leq \limsup_{t\to\infty} \frac{m(t)}{t} \leq \mu$$

hence

$$\lim_{t\to\infty} \frac{m(t)}{t} = \mu$$

If $\mathsf{E}(X_1) < \infty$ but $\mathsf{E}(X_2) = \infty$ then $\mu = 0$ and the last part of the proof holds, i.e., we can still show $0 \leq \limsup_{t\to\infty} \frac{m(t)}{t} \leq 0$ hence

$$0 \leq \liminf_{t\to\infty} \frac{m(t)}{t} \leq \limsup_{t\to\infty} \frac{m(t)}{t} \leq 0$$

We skip the proof for $\mathsf{E}(X_1) = \infty$. ■

3.2.1 Application to DTMCs

We now turn to the proof of Theorem 2.5. Consider a DTMC $Y_n, n \geq 0$, on $\mathcal{S} = 0, 1, 2, \cdots$, with $Y_0 = i$. In the theorem statement, state j is given to be recurrent. As discussed in the beginning of this chapter, the visits to the state j will define a renewal process, with life-times $X_j, j \geq 1$. Note that here all the lifetimes are integer valued random variables. Since f_{ij} is given to be 1, and j is recurrent, we have $P(X_j < \infty) = 1$, for all $j \geq 1$. Further, $\mathsf{E}(X_2) = \nu_j$, the mean return time to the state j, with $\nu_j < \infty$ if j is positive recurrent, and $\nu_j = \infty$ if j is null recurrent.

Proof: Theorem 2.5

We apply the expectation version of ERT, Theorem 3.1, to the renewal process of visits to state j. For this renewal process, we see that, for $n \geq 1$,

$$\begin{aligned} \frac{m(n)}{n} &= \mathsf{E}\left(\left(\frac{M(n)}{n}\right) \mid Y_0 = i\right) \\ &= \mathsf{E}\left(\left(\frac{1}{n}\sum_{k=1}^{n} I_{\{Y_k=j\}}\right) \mid Y_0 = i\right) \\ &= \frac{1}{n}\sum_{k=1}^{n} p_{ij}^{(k)} \end{aligned}$$

Now, $P(X_j < \infty) = 1$, $0 \leq \mathsf{E}(X_1) \leq \infty$, and $1 \leq \mathsf{E}(X_2) = \nu_j$ (it takes at least one transition to return to j). Applying Theorem 3.1, we obtain

$$\frac{1}{n}\sum_{k=1}^{n} p_{ij}^{(k)} = \frac{m(n)}{n} \to_{t\to\infty} \frac{1}{\nu_j} \begin{cases} = 0 & \text{if } \nu_j = \infty \text{ (i.e., } j \text{ is null recurrent)} \\ > 0 & \text{if } \nu_j < \infty \text{ (i.e., } j \text{ is positive recurrent)} \end{cases}$$

■

Remark: With the above notation in mind, recall from Theorem 2.9 that when a DTMC is positive recurrent then the invariant measure $\pi_j, j \in \mathcal{S}$, will have the values $\pi_j = \frac{1}{\nu_j}$. Thus, for a positive recurrent class, the mean time to return to a state is the reciprocal of the invariant probability of that state. Also, we see that π_j has the interpretation of the mean rate of visiting the state j.

3.3 Renewal Reward Processes

Consider a renewal process with i.i.d. life times $X_j, j \geq 1$. Associated with each X_j is a "reward" R_j, such that $R_j, j \geq 1$, is also an i.i.d. sequence, however, R_j may depend on X_j (and, in general, *will* be dependent). We say R_j is the reward during cycle j. Note that we can think of the life-times or cycle times, together with the rewards, as the i.i.d. sequence of random vectors $(X_j, R_j), j \geq 1$.

Example 3.3.

Consider a DTMC $Y_k, k \geq 0$, taking values in $\{0, 1, 2, \cdots\}$. Let $Y_0 = j$ and consider returns to the state j. This defines a renewal process with life-times (or cycle times) $X_k, k \geq 1$, the times between the successive returns to j. Now let $i \neq j$ be another state, and, for $k \geq 1$, let R_k be the number of visits to state i during the duration between the $k-1$th visit to j and the kth visit to j, i.e., in the kth cycle time. It is then easily seen that $R_k, k \geq 1$, are i.i.d. Of course, R_j and X_j are dependent random variables, as one can expect that the longer is the time between two visits to j, the more often would the DTMC visit i between these two visits. ■

With the reward R_j being "earned" in cycle j, let, for $t \geq 0$, $C(t)$ be the net reward until time t (including any reward at t; i.e., like our other processes, $C(t)$ is also right continuous). Now several cases can arise, depending on how the rewards accrue.

- R_j may be obtained at the end of the cycle j. Then

$$C(t) = \sum_{j=1}^{M(t)} R_j$$

  i.e., the total reward until time t is the sum of the rewards in cycles *completed* until time t.

- R_j may be obtained at the beginning of cycle j. Then

$$C(t) = \sum_{j=1}^{M(t)+1} R_j$$

i.e., the total reward until time t is the sum of the rewards in cycles *begun* until time t.

- R_j may be earned over the cycle j (continuously or in discrete parts). Suppose at time instant t, $R(t)$ is defined as the partial reward earned until t in the current cycle. Then, clearly,

$$C(t) = \sum_{j=1}^{M(t)} R_j + R(t)$$

Theorem 3.2 (Renewal-Reward Theorem (RRT))**.** *If* $(X_j, R_j), j \geq 1$, *constitute an i.i.d. renewal reward process, with* $\mathsf{E}(R_1) < \infty$*, and* $0 < \mathsf{E}(X_1) < \infty$*, and* $C(t)$ *is the total reward accumulated during* $[0, t]$*, then*

(a) $\lim_{t\to\infty} \frac{C(t)}{t} = \frac{\mathsf{E}(R_1)}{\mathsf{E}(X_1)}$ *w.p.1*

(b) $\lim_{t\to\infty} \frac{\mathsf{E}(C(t))}{t} = \frac{\mathsf{E}(R_1)}{\mathsf{E}(X_1)}$

Remark: Note that if at the end of each cycle we obtain the reward $R_j = 1$ then the conclusions of this theorem are the same as those of Theorem 3.1, i.e., ERT.
Proof: *of Part (a).*
Consider the case

$$C(t) = \sum_{j=1}^{M(t)} R_j$$

Then we can write

$$\frac{C(t)}{t} = \frac{\sum_{j=1}^{M(t)} R_j}{M(t)} \cdot \frac{M(t)}{t}$$

Now, $\mathsf{E}(X_1) < \infty$ implies that $P(X_1 < \infty) = 1$. Hence, as we saw in the proof of Theorem 3.1, $M(t) \to \infty$, w.p. 1. Then, using the fact that $\mathsf{E}(R_1) < \infty$, Theorem 1.10 implies that $\frac{\sum_{j=1}^{M(t)} R_j}{M(t)} \to \mathsf{E}(R_1)$, w.p. 1. Further, by Theorem 3.1, $\frac{M(t)}{t} \to \frac{1}{\mathsf{E}(X_1)}$, w.p. 1. It follows that, w.p. 1,

$$\lim_{t\to\infty} \frac{C(t)}{t} = \frac{\mathsf{E}(R_1)}{\mathsf{E}(X_1)}$$

An identical argument works for

$$C(t) = \sum_{j=1}^{M(t)+1} R_j$$

after writing this as

$$\frac{C(t)}{t} = \frac{\sum_{j=1}^{M(t)+1} R_j}{M(t)+1} \frac{M(t)+1}{t}$$

Turning to the case where the rewards in a cycle accrue gradually over the cycle, write the reward $R_j = R_j^+ - R_j^-$, where $x^+ = \max\{x, 0\}$, and $x^- = \max\{-x, 0\}$, i.e., we split the reward as the net "gain" minus the net "loss." Also write

$$C(t) = C^+(t) - C^-(t)$$

i.e., even the net reward until t is split as the net gain minus the net loss. Then, we can write

$$\sum_{j=1}^{M(t)} R_j^+ \leq C^+(t) \leq \sum_{j=1}^{M(t)+1} R_j^+$$

Note that this inequalty does not hold if we replace $C^+(t)$ with $C(t)$ and R_j^+ with R_j, since, in general, rewards need not be positive. It follows then, from the cases already proved, that, w.p. 1,

$$\lim_{t\to\infty} \frac{C^+(t)}{t} = \frac{\mathsf{E}\left(R_1^+\right)}{\mathsf{E}(X_1)}$$

In an identical fashion, we also obtain

$$\lim_{t\to\infty} \frac{C^-(t)}{t} = \frac{\mathsf{E}\left(R_1^-\right)}{\mathsf{E}(X_1)}$$

It follows that, w.p. 1,

$$\lim_{t\to\infty} \frac{C(t)}{t} = \frac{\mathsf{E}\left(R_1^+\right) - \mathsf{E}\left(R_1^-\right)}{\mathsf{E}(X_1)} = \frac{\mathsf{E}(R_1)}{\mathsf{E}(X_1)}$$

Proof of Part (b).
Consider the case

$$C(t) = \sum_{j=1}^{M(t)+1} R_j$$

We know that $M(t) + 1$ is a stopping time for $X_j,\ j \geq 1$, i.e., $I_{\{M(t)+1 \geq n\}}$, or, equivalently, $I_{\{M(t)+1 \leq (n-1)\}}$, is determined by $(X_1, X_2, \cdots, X_{n-1})$. Since $(R_j,\ j \geq n)$ is independent of $(X_1, X_2, \cdots, X_{n-1})$, it follows that $I_{\{M(t)+1 \geq n\}}$ is independent of R_n. Then, following an argument exactly as in the proof of Lemma 3.2 (Wald's Lemma), we obtain

$$\mathsf{E}(C(t)) = \mathsf{E}(R_1)\,\mathsf{E}(M(t) + 1)$$

i.e.,

$$\mathsf{E}(C(t)) = \mathsf{E}(R_1)\,(m(t) + 1)$$

Now, dividing by t, and using the expectation part of ERT (Theorem 3.1), it follows that

$$\lim_{t\to\infty} \frac{\mathsf{E}(C(t))}{t} = \frac{\mathsf{E}(R_1)}{\mathsf{E}(X_1)}$$

We omit the proofs of the remaining cases. ■

3.3.1 Application to Time Averages

Let $Y(t), t \geq 0$, be the residual life process of a renewal process with i.i.d. life-times. The common distribution of the life-times is denoted by $F(\cdot)$. Then consider, for a fixed $y \geq 0$,

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t P(Y(u) \leq y)\, du$$

Remark: This expression can be interpreted in two ways. Consider a random observer "arriving" uniformly over the interval $[0, t]$. Then the expression inside the limit, i.e,

$$\int_0^t P(Y(u) \leq y)\, \frac{1}{t}\, du$$

can be interpreted as the distribution of the residual life seen by the random observer; since $\frac{1}{t}du$ is the probability that the observer arrives in the infinitesimal interval $(u, u+du)$, and, conditional on this happening, $P(Y(u) \leq y)$ is the probability of the observer seeing a residual life-time $\leq y$. On the other hand we can write the expression (inside the $\lim_{t\to\infty}$) as

$$\mathsf{E}\left(\frac{1}{t}\int_0^t I_{\{Y(u)\leq y\}} du\right)$$

i.e., the expected fraction of time over $[0, t]$ during which the residual life is $\leq y$. In either case we are asking for the limit of the expression as $t \to \infty$.

Theorem 3.3. *For a renewal process with i.i.d. life-times, such that* $0 < \mathsf{E}\big(X_1^2\big) < \infty$*, the following hold:*

(i) For fixed $y \geq 0$, $\lim_{t\to\infty} \frac{1}{t}\int_0^t P(Y(u) \leq y)du = \frac{1}{\mathsf{E}(X_1)}\int_0^y (1 - F(x))dx$

(ii) With probability 1, $\lim_{t\to\infty} \frac{1}{t}\int_0^t Y(u)du = \frac{\mathsf{E}\left(X_1^2\right)}{2\mathsf{E}(X_1)}$.

Remarks 3.3.

Before we prove Theorem 3.3, we make some observations about its conclusions. Note that, since X_1 is a non-negative random variable, $\mathsf{E}\big(X_1^2\big) > 0$ implies that $\mathsf{E}(X_1) > 0$.

(a.) We first observe that the right hand side of Conclusion (i) in this theorem (i.e., $\frac{1}{\mathsf{E}(X_1)}\int_0^y (1 - F(x))dx$), as a function of y, is a distribution. To see this, note that this expression is nonnegative, nondecreasing with y, and also recall that $\int_0^\infty (1 - F(x))dx = \mathsf{E}(X_1)$. Given a distribution $F(\cdot)$ of a nonnegative random variable, with finite mean, define, for all y,

$$F_e(y) = \frac{1}{\mathsf{E}(X_1)} \int_0^y (1 - F(x))dx$$

$F_e(\cdot)$ is called the *excess* distribution corresponding to the (life-time) distribution $F(\cdot)$.

(b.) Notice that, for each t, the expression $\frac{1}{t}\int_0^t P(Y(u) \le y)du$, as a function of y, is also a distribution. Thus, the first conclusion of the theorem states that the time average distribution of the residual life process converges to the excess distribution of the life-time distribution $F(\cdot)$.

(c.) The second conclusion of the theorem states that the time-average of the residual life process $Y(t)$ converges almost surely to the number $\frac{\mathsf{E}(X_1^2)}{2\mathsf{E}(X_1)}$. It can easily be verified that

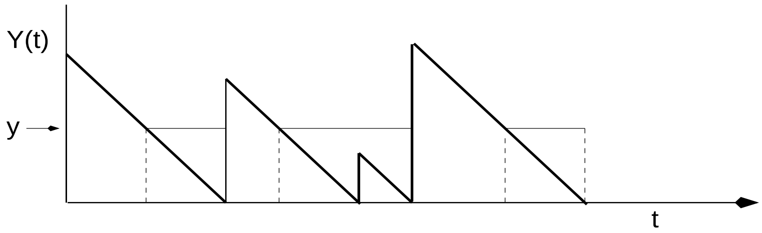
$$\int_0^\infty (1 - F_e(y))dy = \frac{\mathsf{E}(X_1^2)}{2\mathsf{E}(X_1)}$$

To show this, substitute the expression for $F_e(y)$, interchange the order of integration, and use the fact that $\int_0^\infty 2x(1 - F(x))dx = \mathsf{E}(X_1^2))$. Thus, we see that the limit of the time average of the residual life process is the expectation of its time average limiting distribution. A little later in this chapter, we will see that this kind of a result holds more generally. An application of such a result can be that if we are able to obtain the limiting distribution, and we know that such a result holds, then taking the expectation of the limit distribution provides us with the limit of the time average of the process.

Proof: Theorem 3.3, *Part (i).*
Define, for fixed $y \ge 0$, the reward process

$$C(t) = \int_0^t I_{\{Y(u) \le y\}}\, du$$

Thus $C(t)$ is the total time in $[0, t]$ that the proces $Y(t)$ is below y (the integral is the area under the thin horizontal lines in the figure). We see that we have a renewal reward process



with the reward R_j in the jth cycle being given by

$$R_j = \int_{Z_{j-1}}^{Z_j} I_{\{Y(u) \leq y\}} du$$

where, as before, Z_j is the jth renewal instant. Since the life-times are i.i.d. it follows that $\{R_j,\ j \geq 1\}$ are also i.i.d. Further, observe that

$$R_j = \left\{ \begin{array}{lll} X_j & \text{if} & X_j < y \\ y & \text{if} & X_j \geq y \end{array} \right\} = \min(X_j, y)$$

Thus $0 \leq R_j \leq X_j$. Further, $0 < \mathsf{E}\big(X_1^2\big) < \infty$ implies that $0 < \mathsf{E}(X_1) < \infty$, and, hence, that $\mathsf{E}(R_1) < \infty$. Thus the conditions of RRT (Theorem 3.2) are met. Hence

$$\lim_{t\to\infty} \mathsf{E}\bigg(\frac{C(t)}{t}\bigg) \quad = \quad \frac{\mathsf{E}(R_1)}{\mathsf{E}(X_1)}$$

In order to obtain $\mathsf{E}(R_1)$ let us obtain the complementary c.d.f. of R_1. For $r \geq 0$,

$$\begin{aligned} P(R_1 > r) &= P(\min\{X_1, y\} > r) \\ &= P(X_1 > r, y > r) \\ &= \left\{ \begin{array}{lll} P(X_1 > r)(= 1 - F(r)) & \text{for} & r < y \\ 0 & \text{for} & r \geq y \end{array} \right. \end{aligned}$$

It follows that

$$\mathsf{E}(R_1) = \int_0^y (1 - F(r))dr$$

We conclude that

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t P(Y(u) \leq y)du = \frac{1}{\mathsf{E}(X_1)} \int_0^y (1 - F(u))du$$

Proof of Part (ii). Let us take the cumulative reward over $[0, t]$ to be

$$C(t) = \int_0^t Y(u)du$$

i.e., cumulative "area" under $Y(t)$ over $[0, t]$. Then the reward in the jth cycle becomes

$$R_j = \int_{Z_{j-1}}^{Z_j} Y(u)du$$

Therefore,

$$R_j = \frac{1}{2}X_j^2$$

It is evident that $R_j, j \geq 1$, is a sequence of i.i.d. random variables. Further, since we are given that $\mathsf{E}(X_1^2) < \infty$, it follows that $\mathsf{E}(R_1) < \infty$. Thus the conditions of RRT (Theorem 3.2) are met, and we conclude that, with probability 1,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t Y(u)du = \frac{\mathsf{E}(X_1^2)}{2\mathsf{E}(X_1)}$$

■

3.3.2 Length and Batch Biasing

We can interpret the second part of Theorem 3.3 as asserting that in a renewal process with i.i.d. life-times the mean residual time seen by a random observer is given by $\frac{\mathsf{E}(X_1^2)}{2\mathsf{E}(X_1)}$. The following (fallacious) argument gives a different (and wrong) answer. "If a randomly arriving observer arrives in a life-time of duration x, then (since he arrives uniformly over this interval) he sees a mean residual time of $\frac{x}{2}$. The distribution of life-times is $F(\cdot)$. Hence, unconditioning on the length of the life-time into which the observer arrives, the mean residual life seen by a random observer should be given by $\int_0^\infty \frac{x}{2}\, dF(x) = \frac{1}{2}\mathsf{E}(X_1)$."

In fact, we notice that, since $\mathsf{E}(X_1^2) \geq (\mathsf{E}(X_1))^2$ (with equality only if the random variable X_1 is constant with probability 1),

$$\frac{\mathsf{E}(X_1^2)}{2\mathsf{E}(X_1)} \geq \frac{1}{2}\mathsf{E}(X_1)$$

with strict inequality if the variance of X_1 is positive. Thus, the correct mean residual life will typically be larger than half the mean life-time. What is the flaw in the argument (in quotes) above? The error is in the use of the distribution $F(\cdot)$ when unconditioning on the distribution of the life-time into which the random observer arrives. In fact, we should use a *length biased* distribution as we shall see next.

Theorem 3.4. *If the life-time distribution has finite mean* $\mathsf{E}(X_1)$*, then, for given* $x \geq 0$*,*

$$\lim_{t\to\infty} \frac{1}{t} \int\limits_0^t P(X(u) \leq x) du$$
$$= \frac{1}{EX_1} \left(xF(x) - \int\limits_0^x F(u) du \right) := F_s(x)$$

Proof: Consider, for given $x \geq 0$, the cumulative reward process

$$C(t) = \int\limits_0^t I_{\{X(u)\leq x\}} du$$

and the reward sequence R_j defined by

$$R_j = \begin{cases} 0 & \text{for} \quad X_j > x \\ X_j & \text{for} \quad X_j \leq x \end{cases}$$

Since $R_j \leq X_j$, $\mathsf{E}(R_1) \leq \mathsf{E}(X_1) < \infty$. Hence, applying RRT (Theorem 3.2) we conclude that

$$\lim_{t\to\infty} \frac{1}{t} \int\limits_0^t P(X(u) \leq x) du = \frac{\mathsf{E}(R_1)}{\mathsf{E}(X_1)}$$

In order to evaluate $\mathsf{E}(R_1)$ we first obtain the complementary c.d.f. of R_j. Clearly, $P(R_j > x) = 0$. Further, for $0 \leq u \leq x$,

$$\begin{aligned} P(R_j > u) &= P(X_j \leq x, X_j > u) \\ &= P(X_j \leq x) - P(X_j \leq u) \\ &= F(x) - F(u) \end{aligned}$$

Hence

$$\begin{aligned} \mathsf{E}(R_1) &= \int\limits_0^\infty P(R_1 > u)\, du \\ &= \int\limits_0^x (F(x) - F(u))\, du \\ &= xF(x) - \int\limits_0^x F(u)\, du \end{aligned}$$

from which the result follows. ∎

Remarks 3.4.

(a.) Given a life-time distribution $F(\cdot)$, the distribution $F_s(\cdot)$ obtained in Theorem 3.4 is called the *spread distribution* of $F(\cdot)$. Compare this with the earlier definition of the excess distribution $F_e(\cdot)$.

(b.) As in the case of $F_e(\cdot)$, the distribution $F_s(\cdot)$ can be viewed as the distribution of the *total* life-time seen by a random observer. When $F(\cdot)$ has a density, then it is easily seen, by differentiating the distribution $F_s(\cdot)$, that the density of the spread distribution is given by

$$f_s(x) = \frac{1}{\mathsf{E}(X_1)}(xf(x))$$

Thus we see that the density is different from $f(\cdot)$. Now following the discussion at the beginning of this section, if we obtain the mean residual life by taking $f_s(\cdot)$ to be the distribution of life-time seen by a random observer, then we will obtain

$$\int_0^\infty \frac{x}{2}\frac{1}{\mathsf{E}(X_1)}(xf(x))dx = \frac{\mathsf{E}(X_1^2)}{2\mathsf{E}(X_1)}$$

the desired answer. What is the intuition behind $f_s(\cdot)$ being the correct density to use, and not $f(\cdot)$? The following rough argument will make this clear. Consider the time interval $[0, T]$, where T is suitably large. Over this time there are roughly $\frac{T}{\mathsf{E}(X_1)}$ renewal cycles; this follows from ERT. Out of these cycles the number that have durations in the interval $(x, x + dx)$ is approximately $\frac{T}{\mathsf{E}(X_1)}f(x)dx$. The amount of time in the interval $[0, T]$ that is covered by such intervals is $x\frac{T}{\mathsf{E}(X_1)}f(x)dx$, where we have ignored terms of order smaller than dx. Hence a random arrival over $[0, T]$ arrives in an interval of length $(x, x + dx)$ with probability $\frac{x\frac{T}{\mathsf{E}(X_1)}f(x)dx}{T} = f_s(x)dx$. Hence $f_s(\cdot)$ is the density of the life-time seen by a random observer. It should now be clear why $f_s(\cdot)$ is called the *length biased* density.

(c.) The above discussion illustrates the phenomenon of *length biased sampling*. Let $X_k, k \geq 1$, be i.i.d. life-time samples from the c.d.f. $F(\cdot)$ (with nonnegative support). Suppose we take a large number n of samples and place them in a bin and then draw a sample, the distribution of the sample will be $F(\cdot)$. On the other hand if we place the life-time samples “side-by-side” (thinking of them as line segments) on the positive real line (starting at the origin), and then “draw” a sample by randomly picking a point of the positive real line, yielding the sample in which the point falls, then evidently our sampling will be biased towards the larger values, as these cover more of the line.

■

The same phenomenon obviously occurs when the random variables $X_k, k \geq 1$, take only nonnegative integer values, but in this context it is called *batch size biasing*. Suppose families either have 1 child or 2 children, with probability 0.5 for each case. If mothers are asked how many children they have then roughly half will reply "one", and other half will reply "two," yielding an average number of children equal to 1.5. On the other hand if children are asked how many siblings they are, roughly $\frac{1}{3}$rd will answer "one", and the rest will answer "two," yielding the batch biased average: $\frac{5}{3}$. The biasing occurs since in the population of children more children come from the larger families and hence we more often get the larger answer. In some queueing systems, customers arrive in batches, and if a customer is picked at random and asked the size of its batch a batch biased answer will result[2]. We can again study this using the renewal reward theorem.

Let the batch sizes be $\{X_j,\ j \geq 1\}$; these are i.i.d. nonnegative integer random variables representing the number of customers in the batches indexed by $j \geq 1$. Let $P(X_1 = k) = p_k, k \geq 1$. Let us index the customers by $n \in \{1, 2, 3, \cdots\}$, so that customers 1 to X_1 are in the first batch, from $X_1 + 1$ to $X_1 + X_2$ are in the second batch, and so on. Note that we can view this as a renewal process in discrete "time," with the batches corresponding to "life-times." Let $X(n)$ denote the batch of the nth customer; this notation is similar to the notation $X(t)$ in continuous time renewal theory introduced earlier. Now, for fixed $k \geq 1$, consider

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} I_{\{X(i)=k\}}$$

i.e., each customer is asked if its batch size is k, and we seek the fraction of customers who respond with a "yes." View the cumulative "reward" until the nth customer as $\sum_{i=1}^{n} I_{\{X(i)=k\}}$, and for $j \geq 1$, define

$$\begin{aligned} R_j &= 0 \text{ if } X_j \neq k \\ &= k \text{ if } X_j = k \end{aligned}$$

Evidently $(X_j, R_j), j \geq 1$, are i.i.d. and we have a renewal reward process. We have

$$\mathsf{E}(R_j) = kp_k < \infty$$

Hence, using Theorem 3.2, w.p. 1

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} I_{\{X(j)=k\}} = \frac{kp_k}{\mathsf{E}(X_1)}$$

[2]One consequence of this is the following. Consider a single server queue with batch arrivals. If the batches of customers are served in first-come-first-served order, then the delay of a typical customer is the sum of the delay of the batch in which it arrives, and the total service time of the customers in its own batch that are served before it. This latter number will need to be obtained by using the batch biased distribution.

which is the biased distribution of the batch size. It can similarly be seen that

$$\lim_{t\to\infty} \frac{1}{n} \sum_{j=1}^{n} P(X(j) = k) = \frac{kp_k}{\mathsf{E}(X_1)}$$

and, w.p. 1,

$$\lim_{t\to\infty} \frac{1}{n} \sum_{j=1}^{n} X(j) = \frac{\sum_{k=1}^{\infty} k^2 p_k}{\mathsf{E}(X_1)} = \frac{\mathsf{E}\left(X_1^2\right)}{\mathsf{E}(X_1)}$$

which is the average of replies received if each customer is asked "What is your batch size?"

Remarks 3.5.

Finally, in the context of the renewal reward theorem, it is important to observe that the form of the result obtained, i.e., that the limit of the average reward rate is $\frac{\mathsf{E}(R_1)}{\mathsf{E}(X_1)}$, incorporates the effect of length biasing. In particular, one could ask why the limit was not $\mathsf{E}\left(\frac{R_1}{X_1}\right)$ (in fact, this expression has sometimes been erroneously used). For example, suppose $(X_k, R_k), k \geq 1$, are i.i.d.. Suppose that $(X_k, R_k) = (1, 10)$ with probability 0.5, and $(X_k, R_k) = (10, 1000)$ with probability 0.5. So, on the average, in half the intervals the reward rate is 10, and in half the intervals the reward rate is 100. One might want to say that the average reward rate is $0.5 \times 10 + 0.5 \times 100 = 55$; this would be the answer one would get if the formula $\mathsf{E}\left(\frac{R_1}{X_1}\right)$ is used. Yet, the theorem declares that the answer is $\frac{\mathsf{E}(R_1)}{\mathsf{E}(X_1)} = \frac{0.5\times10+0.5\times1000}{0.5\times1+0.5\times10} = 91.82$. It is easily checked that this is the answer that we will get if we use the length biased distribution for the cycle time.

■

3.4 The Poisson Process

A point process $N(t), t \geq 0$, is a random process taking values in $0, 1, 2, \cdots$, such that $N(0) = 0$, and, for each ω, $N(t, \omega)$ is a nondecreasing, right continuous step function. A point process is essentially characterised by a random distribution of points on $\mathbb{R}^+$ (i.e., the jump instants), and a sequence of integer random variables corresponding to the jump at each point. If we think of a batch of arrivals at each jump, then $N(t)$ is the number of arrivals in the interval $[0, t]$.

Definition 3.2 (Poisson Process)**.** *A point process $N(t), t \geq 0$, is called a Poisson process if*

(i) *all jumps of $N(t)$ are of unit size, with probability 1,*

(ii) *for all $t, s \geq 0$, $(N(t+s) - N(t)) \amalg \{N(u) : u \leq t\}$, and*

(iii) *for all* $t, s \geq 0$*, distribution of* $N(t+s) - N(t)$ *does not depend on* t*.*

■

Remarks 3.6.

(a.) For a point process $N(t)$, given $t \geq 0$ and $\tau \geq 0$, $N(t+\tau) - N(t)$ is called the *increment* of $N(t)$ over the interval $(t, t+\tau]$, i.e., $N(t+\tau) - N(t)$ is the cumulative value of the jumps in the interval $(t, t+\tau]$.

(b.) Thus Definition 3.2 (ii) asserts that any increment of a Poisson process is independent of the past of the process. We can conclude more. Consider $t_1 < t_2 < t_3 \cdots < t_n$ then, by property (ii) in the definition,

$$N(t_1) \perp\!\!\!\perp (N(t_2) - N(t_1))$$

Further, again using the same property,

$$N(t_3) - N(t_2) \perp\!\!\!\perp (N(t_1), N(t_2))$$

or, equivalently,

$$(N(t_3) - N(t_2)) \perp\!\!\!\perp (N(t_1), N(t_2) - N(t_1))$$

From the above, it easily follows that

$$(N(t_3) - N(t_2)), (N(t_2) - N(t_1)), N(t_1) \text{ are mutually independent.}$$

Continuing in this way, for every $k, 3 \leq k \leq n$, by Definition 3.2 (ii) we can assert

$$(N(t_k) - N(t_{k-1})) \perp\!\!\!\perp ((N(t_{k-1}) - N(t_{k-2})), (N(t_{k-2}) - N(t_{k-3})), \ldots, N(t_1))$$

from which, given that $(N(t_{k-1}) - N(t_{k-2})), (N(t_{k-2}) - N(t_{k-3})), \ldots, N(t_1)$ are mutually independent, it follows that

$$(N(t_k) - N(t_{k-1})), (N(t_{k-1}) - N(t_{k-2})), (N(t_{k-2}) - N(t_{k-3})), \ldots, N(t_1)$$

are mutually independent.

We conclude that

$$(N(t_n) - N(t_{n-1})), (N(t_{n-1}) - N(t_{n-2})), (N(t_{n-2}) - N(t_{n-3})), \ldots, N(t_1)$$

are mutually independent. Thus, the increments of a Poisson process (over disjoint intervals) are independent. This is known as the *independent increment* property.

(c.) Definition 3.2 (iii) states that the distribution of an increment depends only on the width of the interval over which it is taken, not the location of the interval in time. This is called the *stationary increment* property.

(d.) Thus we can say that a Poisson process is a point process with stationary and independent increments. In addition, at each point of a Poisson process there is a unit jump.

■

Definition 3.2 defines a stochastic process. We have seen that a stochastic process is characterised in terms of its finite dimensional distributions (recall Section 1.3.1). Are the properties specified in the definition of a Poisson process sufficient to specify its finite dimensional distributions? The answer is indeed "Yes" as we now proceed to show. For $t_1 < t_2 < t_3 < \cdots < t_n$, and $i_1 \leq i_2 \leq i_3 \leq \cdots \leq i_n$, using the properties of the Poisson process, we find that the finite dimensional distribution can be computed as follows

$$\begin{aligned} & P(N(t_1) = i_1, N(t_2) = i_2, \cdots, N(t_n) = i_n) \\ & = P(N(t_1) = i_1, N(t_2) - N(t_1) = i_2 - i_1, \cdots, N(t_n) - N(t_{n-1}) = i_n - i_{n-1}) \\ & = P(N(t_1) = i_1) \cdot P(N(t_2 - t_1) = i_2 - i_1) \cdots P(N(t_n - t_{n-1}) = i_n - i_{n-1}) \end{aligned}$$

where, in writing the second equality, we have used the stationary and independent increment properties. Thus we would have the finite dimensional distributions if we could obtain $P(N(t) = k)$ for each $t \geq 0$ and $k \in \{0, 1, 2, \cdots\}$. This distribution is obtained via the following lemmas.

Lemma 3.3. *There exists a* $\lambda, 0 \leq \lambda < \infty$, *such that, for all* $t \geq 0$, $P\{N(t) = 0\} = e^{-\lambda t}$.

Remark: This is equivalent to the assertion that the time until the first jump in a Poisson process is exponentially distributed with mean $\frac{1}{\lambda}$.
Proof: Using the stationary and independent increment property, we can write

$$\begin{aligned} P(N(t+s) = 0) &= P(N(t) = 0, N(t+s) - N(t) = 0) \\ &= P(N(t) = 0) P(N(t+s) - N(t) = 0) \end{aligned}$$

Let us write, for $t \geq 0$, $f(t) = P(N(t) = 0)$. Thus, we have established that $f(\cdot)$ satisfies the *functional equation:* for all $s, t \geq 0$,

$$f(t+s) = f(t)\, f(s)$$

Since $N(0) = 0$, we have $f(0) = 1$. Define $T_1 := \inf\{t > 0 : N(t) \neq 0\}$, the first "jump" time of $N(t)$. Clearly, $f(t) = P(T_1 > t)$. Thus, $f(t)$ is the complementary c.d.f. of a nonnegative random variable; hence, $f(t), t \geq 0$, is right continuous and nonincreasing. It then follows that (see Theorem 3.22 in the Appendix of this chapter) the only nonzero solution of the functional equation is

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

for some $\lambda, 0 \leq \lambda < \infty$.

■

Lemma 3.4.

$$\lim_{t \to 0} \frac{1}{t} P(N(t) \geq 2) = 0$$

Remark: This result states that $P(N(t) \geq 2) = o(t)$, i.e., that the probability of there being 2 or more points of the process in an interval of length t decreases to 0 faster than t, as t decreases to 0.
Proof: We skip the proof of this result, and only note that the proof utilises the property (i) in Definition 3.2. ■

Lemma 3.5.

$$\lim_{t \to 0} \frac{1}{t} P(N(t) = 1) = \lambda,$$

where λ is as obtained in Lemma 3.3.

Remark: In other words, $P(N(t) = 1)$ can be approximated as $\lambda t + o(t)$ as $t \to 0$.
Proof:

$$P(N(t) = 1) = 1 - P(N(t) = 0) - P(N(t) \geq 2)$$

Using Lemma 3.3, we can write

$$\lim_{t \to 0} \frac{1}{t} P(N(t) = 1) = \lim_{t \to 0} \left(\frac{1 - e^{-\lambda t}}{t} - \frac{P(N(t) \geq 2)}{t} \right)$$

From which the result is obtained after using Lemma 3.4. ■

Theorem 3.5. *If $N(t), t \geq 0$, is a Poisson process then, for all $t \geq 0$, and $k \in \{0, 1, 2, \cdots\}$,*

$$P(N(t) = k) = \frac{e^{-\lambda t} (\lambda t)^k}{k!},$$

where λ is as obtained in Lemma 3.3.

Remark: Thus an increment of a Poisson process over an interval of length t is Poisson distributed with parameter λt.
Proof: For $0 < \alpha < 1$, define $G(t) = \mathsf{E}\big(\alpha^{N(t)}\big)$, i.e., $G(t)$ is the moment generating function of the random variable $N(t)$. For notational simplicity, we have not retained

α as an argument of the moment generating function. Now, using the stationary and independent increment property, we obtain a functional equation for $G(t)$ as follows

$$\begin{aligned} G(t+s) &= \mathsf{E}\Big(\alpha^{N(t+s)}\Big) \\ &= \mathsf{E}\Big(\alpha^{N(t)} \cdot \alpha^{(N(t+s)-N(t))}\Big) \\ &= G(t)\, G(s) \end{aligned}$$

Since $N(0) = 0$, $G(0) = 1$; since $N(t)$ increases with t, and $0 < \alpha < 1$, we also conclude that $G(t)$ is nonincreasing with t. Also, by Lemmas 3.3, 3.5, and 3.4, we can write

$$\begin{aligned} \lim_{t\to 0} G(t) &= \lim_{t\to 0} \mathsf{E}\Big(\alpha^{N(t)}\Big) \\ &= \lim_{t\to 0} \left(1 \cdot e^{-\lambda t} + \alpha \cdot (\lambda t + o(t)) + \sum_{k=2}^{\infty} \alpha^k P(N(t) = k)\right) \\ &= 1 + 0 + \lim_{t\to 0} o(t) \\ &= 1 \end{aligned}$$

establishing that $G(t)$ is continuous from the right at $t = 0$. Now using Theorem 3.22 (in the Appendix of this chapter) we conclude that the unique solution to this functional equation is

$$G(t) = e^{g(\alpha)t}$$

for some constant $g(\alpha)$. To obtain $g(\alpha)$, we observe that

$$\begin{aligned} g(\alpha) &= \lim_{t\to 0} \frac{G(t) - G(0)}{t} \\ &= \lim_{t\to 0} \left(\frac{1}{t}\left[P(N(t) = 0) - 1\right] + \frac{\alpha \cdot P(N(t) = 1)}{t}\right. \\ &\qquad \left. + \frac{1}{t} \sum_{k=2}^{\infty} \alpha^k P(N(t) = k)\right) \end{aligned}$$

which, on using Lemmas 3.4 and 3.5, yields $g(\alpha) = -\lambda + \alpha\lambda$. Thus we find that

$$G(t) = e^{-\lambda t + \lambda t \alpha}$$

i.e.,

$$\begin{aligned} G(t) &= \sum_{k=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^k}{k!}\, \alpha^k \\ &= \sum_{k=0}^{\infty} P(N(t) = k)\alpha^k \end{aligned}$$

It follows that, for $k \in \{0, 1, 2, \cdots\}$,

$$P(N(t) = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}$$

■

Remarks 3.7.

(a.) Continuing the argument begun before Lemma 3.3, we see that the finite dimensional distributions of a Poisson process are completely characterised in terms of a single parameter λ. Thus we can now use the term: "a Poisson process with parameter λ". A little later we shall see that λ is the *rate* of the Poisson point process. We can then refer to a Poisson process with rate λ.

(b.) It can easily be verified that $\mathsf{E}(N(t)) = \lambda t$, and $\mathsf{Var}(N(t)) = \lambda t$. It is useful to remember that the variance to mean ratio of a Poisson process is 1. Thus the Poisson process is often used as a "benchmark" for the variability in arrival processes. An arrival process with a higher variance to mean ratio is said to be *burstier* than Poisson, and an arrival process with variance to mean ratio less than Poisson can be viewed as being smoother than Poisson.

■

Theorem 3.6. *$N(t)$ is a Poisson process with rate λ. Let $0 = t_0 < t_1 < t_2 < \cdots < t_n = t$. Then for all $k, k_1, k_2, \cdots, k_n$ nonnegative integers such that $\sum_{i=1}^n k_i = k$ we have*

$$P(N(t_1) = k_1, N(t_2) - N(t_1) = k_2, \cdots, N(t_n) - N(t_{n-1}) = k_n | N(t) = k) = \frac{k!}{\prod_{i=1}^n k_i!} \prod_{i=1}^n \left(\frac{t_i - t_{i-1}}{t}\right)^{k_i}$$

Remarks 3.8.

(a.) Note that the time points $t_i, 1 \leq i \leq n-1$, partition the interval $[0, t]$ into consecutive subintervals. The question being asked is that, given that exactly k Poisson points are known to have occurred in the interval $[0, t]$, what is the probability that k_i of them fell in the ith interval, $1 \leq i \leq n$.

(b.) This result states that conditioned on there being k Poisson arrivals in an interval $[0, t]$, these k points are distributed over the interval as if each one of them was independently and uniformly distributed over the interval. With this explanation, the form of the right hand side of the conclusion in the theorem becomes self evident; it is the multinomial probability distribution with k "trials" and n alternatives in each trial, the ith alternative being that a point falls in the ith subinterval, and the probability of this alternative is $\frac{t_i - t_{i-1}}{t}$.

(c.) Conversely, it can be seen that the Poisson process is obtained if we uniformly distributed points on $\mathbb{R}^+$ in the following way. Let us uniformly and independently distribute n points on the interval $[0, \frac{n}{\lambda}]$. Consider the interval $[0, t]$. Let n be large enough so that $t < \frac{n}{\lambda}$. Then the number of points that fall in the interval $[0, t]$ is distributed as Binomial$(n, \frac{t}{\frac{n}{\lambda}})$. As $n \to \infty$, it then follows that the distribution of the number of points in $[0, t]$ converges to Poisson(λt). Thus the Poisson process can be viewed as the limit of a uniform distribution of points (the reader should try to provide a complete proof; what remains is to show the independent increment property).

Proof: The proof is a simple matter of applying the conditional probability formula, using the stationary and independent increment property, and the Poisson distribution of each increment. The details are left as an exercise. ■

3.4.1 Stopping Times

Definition 3.3. *A random time T is called a stopping time for a Poisson process $N(t), t \geq 0$, if $\{T \leq u\} \amalg \{N(s) - N(u),\ s \geq u\}$.* ■

Remarks 3.9.

(a.) Thus a random time T is a stopping time for a Poisson process $N(t)$ if the question "Is $T \leq u$?" can be answered independently of the increments of $N(t)$ after the time u. Unlike the earlier Definition 2.5, here we do not require that the event $\{T \leq u\}$ be determined by $N(t), t \leq u$. Note that this property of independence of the future is all that was required in proving Wald's Lemma (Lemma 3.2).

(b.) From the point of view of applications, this definition of stopping time is what will be more useful. In queueing applications the Poisson process will typically model an arrival process, and various random times will be determined by the arrival process in conjunction with other random processes, such as the sequence of service requirements of customers. Thus the only reasonable requirement of a random time T would be that $\{T \leq u\}$ is independent of future arrivals.

■

Example 3.4.

Let us consider an M/G/1 queue. This is a queueing system in which customers arrive to an infinite buffer and wait for service by a single server (the trailing "1" in the notation). The customer arrival instants constitute a Poisson process (the "M" in the notation), and the service times of the customers is a sequence of i.i.d. random variables (the "G" in the notation, meaning generally distributed services times). Let $X(t), t \geq 0$, be the number of customers at time t. Let $A(t)$ denote the arrival process, and let $T_k, k \geq 1$, denote the successive arrival instants. Suppose that $X(t) = 0$, and, for a sample point ω, let $Z(\omega) := \inf\{t \geq 0 : t \geq T_1(\omega), X(t, \omega) = 0\}$; i.e., Z is the random time at which the queue becomes empty for the first time after once becoming nonempty. The random interval $[0, Z]$ is called an idle-busy period of the queue. It is easily seen that $\{Z \leq z\}$ is determined by the arrivals in the interval $(0, z]$ and the service times of these arrivals, and hence is independent of future increments of the arrival process, i.e., of $A(z + s) - A(z), s \geq 0$. ■

The following is an important property of the Poisson process in relation to stopping times.

Theorem 3.7. *Let T be a stopping time for the Poisson process $N(t)$, with $P(T < \infty) = 1$, then $N(T+s)-N(T),\ s \geq 0$, is a Poisson process independent of T and of $\{N(t),\ t \leq T\}$.*

Remark: This results asserts that if T is a proper stopping time for the Poisson process $N(t)$ then the increments of $N(t)$ starting from T form a Poisson process, and this process is independent of T as well as of the past of $N(t)$ prior to T. With reference to Example 3.4, we can now assert that if Z_1 is the end of the first idle-busy cycle then the process $X(t + Z_1), t \geq 0$, is again an M/G/1 queue length process that starts with an empty queue. This is because, by virtue of Z_1 being a stopping time for the arrival process, $A(t+Z_1) - A(Z_1), t \geq 0$, is again a Poisson process, and the successive service times are in any case i.i.d. Thus $X(t + Z_1), t \geq 0$, is a statistical replica of $X(t), t \geq 0$. Moreover, since $A(t + Z_1) - A(Z_1), t \geq 0$, is independent of Z_1 and of $A(t), t \leq Z_1$, we conclude that $X(t + Z_1), t \geq 0$, is independent of Z_1 and of $X(t), t \leq Z_1$.

Proof: This result follows from the strong Markov property of continuous time Markov chains; see Theorem 4.3. ■

Exercise 3.2.

Given a Poisson process $N(t), t \geq 0$, show that the following assertions hold

(a.) The jump instants $T_k, k \geq 1$, are stopping times.

(b.) If T_k is a jump instant and $\epsilon > 0$, then $T_k - \epsilon$ is not a stopping time.

(c.) For each $t \geq 0$, $T = t$ is a stopping time.

■

Corollary 3.2. *Given a Poisson process $N(t)$ with rate λ, and T a stopping time with $P(T < \infty) = 1$,*

$$P(N(T + s) - N(T) = 0 | N(s), s \leq T) = e^{-\lambda s}$$

Proof: It follows from Theorem 3.7 that

$$P(N(T+s) - N(T) = 0 | N(s), s \leq T) = P(N(s) = 0) = e^{-\lambda s}$$

■

Remarks 3.10.

(a.) We saw in Exercise 3.2 that all jump times of a Poisson process are stopping times. It follows from this corollary that the successive interjump times are i.i.d. exponential with parameter λ. Hence we learn that the Poisson process is a renewal process with i.i.d. exponential life-times. This is, in fact, an alternate characterisation of a Poisson process.

(b.) We also saw in Exercise 3.2 that any time $t \geq 0$ is a stopping time. It follows that the residual life process $Y(t)$ of this renewal process is exponentially distributed for every $t \geq 0$; i.e., $P(Y(t) > s) = e^{-\lambda s}$ for all $t \geq 0$. Also we note here that for any renewal process the residual life process $\{Y(t),\, t \geq 0\}$ is a continuous time continuous state Markov Process. This along with the just observed fact that for a Poisson process the distribution of $Y(t)$ is invariant with t, shows that, for a Poisson process, $\{Y(t),\, t \geq 0\}$ is a stationary process. Later in this chapter will explore this notion of stationarity for more general renewal processes.

(c.) It should be intuitively clear that these results are a consequence of the memoryless nature of the exponential distribution; i.e., if X has distribution Exp(λ) then $P(X > x + y \mid X > x) = e^{-\lambda y}$. Hence not only is the residual life-time exponentially distributed, it is also independent of the elapsed life.

■

3.4.2 Other Characterisations

In Definition 3.2 a Poisson process was defined as a point process with stationary and independent increments, and with unit jumps, with probability 1. In applications it is often useful to have alternate equivalent characterisations of the Poisson process. The following are two such characterisations.

Theorem 3.8. *A point process $N(t), t \geq 0$, is a Poisson process if and only if*

(a.) for all $t_0 < t_1 < t_2 < \ldots < t_n$, the increments $N(t_i) - N(t_{i-1}), 1 \leq i \leq n$, are independent random variables, and

(b.) there exists $\infty > \lambda \geq 0$ such that $P(N(t+s) - N(t) = k) = \frac{(\lambda s)^k\, e^{-\lambda s}}{k!}$.

Proof: The "only if" assertion follows since it has already been shown that the original definition of the Poisson process (i.e., Definition 3.2) implies these properties (see Theorem 3.5). As for the "if part" of the proof, note that the stationary and independent increment property follows immediately. Also the Poisson distribution of the increments implies that the time between successive jumps is 0 with zero probability, thus completing the proof. ■

Theorem 3.9. *A point process $N(t)$, $t \geq 0$, is a Poisson process with parameter λ if and only if the successive jump times $T_k, k \geq 0$, are renewal instants with i.i.d. exponentially distributed inter-renewal times with mean $\frac{1}{\lambda}$.*

Remark: It follows from ERT (Theorem 3.1) that, with probability 1, $\lim_{t\to\infty} \frac{N(t)}{t} = \lambda$. Hence λ is the *rate* of the Poisson point process.

Proof: The "only if" part has already been established in Remarks 3.10. For the "if part," (i) the almost surely unit jump property follows from the exponentially distributed inter-renewal times, (ii) the stationary increment property will follow when we study stationary renewal processes in Section 3.7, and (iii) the independent increment property follows from the memoryless property of the exponentially distributed inter-renewal times. ■

3.4.3 Splitting and Superposition

Consider a Poisson process $N(t)$ with rate λ, and denote its jump instants by $\{T_1, T_2, \ldots\}$. Consider an independent Bernoulli process $Z_k, k \geq 1$,; i.e., such that Z_k are i.i.d. with

$$Z_k = \begin{cases} 1 & \text{w.p.} \quad p \\ 0 & \text{w.p.} \quad 1-p \end{cases}$$

Now define two new point processes $N^{(1)}(t)$ and $N^{(2)}(t)$ as follows: each point $T_k, 1 \leq k \leq N(t)$, is a point of $N^{(1)}(t)$ if $Z_k = 1$, else T_k is a point of $N^{(2)}(t)$. Thus each point of the Poisson process $N(t)$ is assigned to either $N^{(1)}(t)$ or to $N^{(2)}(t)$, with probability p or $1-p$, respectively, and the assignment is independent from across the points of $N(t)$. If $N(t)$ is an arrival process into a queue, it as if Z_k is used to split the process into two arrival processes.

Theorem 3.10. *$N^{(1)}(t)$ and $N^{(2)}(t)$ are independent Poisson processes with rates $p\lambda$ and $(1-p)\lambda$, respectively.*

Remark: The hypothesis of Bernoulli sampling is crucial. If there is dependence between the selection of successive points of $N(t)$ then the resulting process will not be Poisson. As an elementary counterexample, consider splitting the points of $N(t)$ so that alternate points are assigned to $N^{(1)}(t)$ and $N^{(2)}(t)$. Now we see that the interarrival instants in each of $N^{(1)}(t)$ and $N^{(2)}(t)$ are convolutions of exponentially distributed random variables (i.e., they are Erlang distributed, or gamma distributed), and hence neither $N^{(1)}(t)$ nor

$N^{(2)}(t)$ can be Poisson. The Bernoulli nature of the splitting is essential for retaining the memoryless property of the life-times.

Proof: Invoking Theorem 3.8, we need to prove the following three assertions

(a.) $N^{(1)}(t)$ and $N^{(2)}(t)$ are independent processes.

(b.) $N^{(1)}(t)$ has independent increments, and so does $N^{(2)}(t)$.

(c.) $N^{(1)}(t)$ has Poisson distributed increments with parameter $p\lambda$, and $N^{(2)}(t)$ has Poisson distributed increments with parameter $(1-p)\lambda$.

Assertion 1 and 2 are established if we show that for $t_2 > t_1$

$$\begin{aligned} & P\left(N^{(1)}(t_2) - N^{(1)}(t_1) = k_1, N^{(2)}(t_2) - N^{(2)}(t_1) = k_2\right) \\ = \; & \frac{(\lambda p(t_2 - t_1))^{k_1} \; e^{-\lambda p(t_2-t_1)}}{k_1!} \cdot \\ & \frac{(\lambda(1-p)(t_2-t_1))^{k_2} \; e^{-\lambda(1-p)(t_2-t_1)}}{k_2!} \end{aligned}$$

This is easily seen as follows

$$\begin{aligned} & P\left(N^{(1)}(t_2) - N^{(1)}(t_1) = k_1, \; N^{(2)}(t_2) - N^{(2)}(t_1) = k_2\right) \\ = \; & P(N(t_2) - N(t_1) = k_1 + k_2) \cdot \frac{(k_1+k_2)!}{k_1!k_2!} \, p^{k_1}(1-p)^{k_2} \\ = \; & \frac{(\lambda(t_2-t_1))^{k_1+k_2} \; e^{-\lambda(t_2-t_1)}}{(k_1+k_2)!} \; \frac{(k_1+k_2)!}{k_1!k_2!} \, p^{k_1}(1-p)^{k_2} \\ = \; & \frac{(\lambda p(t_2 - t_1))^{k_1} \; e^{-\lambda p(t_2-t_1)}}{k_1!} \cdot \\ & \frac{(\lambda(1-p)(t_2-t_1))^{k_2} \; e^{-\lambda(1-p)(t_2-t_1)}}{k_2!} \end{aligned}$$

To establish the independent increment property for $N^{(1)}(t)$ (or for $N^{(2)}(t)$) (over disjoint intervals) we observe that such increments of $N(t)$ are independent, and each increment of $N^{(1)}(t)$ is obtained by independently selecting points from the increments of $N(t)$ (i.e., the selection of points from disjoint increments of $N(t)$ are done independently by virtue of the Bernoulli sampling). The elementary details are left as an exercise. ∎

Let us now consider the superposition (or the merging) of two Poisson processes. Let $N^{(1)}(t)$ and $N^{(2)}(t)$ be two independent Poisson processes with rates λ_1 and λ_2. Define the point process $N(t)$ by

$$N(t) := N^{(1)}(t) + N^{(2)}(t)$$

i.e., each point of $N^{(1)}(t)$ and of $N^{(2)}(t)$ is assigned to $N(t)$.

Theorem 3.11. *$N(t)$ is a Poisson process with rate $\lambda = \lambda_1 + \lambda_2$.*

Proof: Invoking Theorem 3.8, we need to prove that increments of $N(t)$ over disjoint intervals are independent, and an increment over the interval $(t_1, t_2]$ is Poisson distributed with mean $\lambda(t_2 - t_1)$, where $\lambda = \lambda_1 + \lambda_2$. The independent increment property follows easily from the corresponding property of the Poisson processes $N^{(1)}(t)$ and $N^{(2)}(t)$, and we leave the details as an exercise. Turning to the distribution of the increments, observe that

$$
\begin{aligned}
&P(N(t_2) - N(t_1) = k) \\
&= \sum_{i=0}^{k} P(N^{(1)}(t_2) - N^{(1)}(t_1) = i) \cdot P(N^{(2)}(t_2) - N^{(2)}(t_1) = k - i) \\
&= \sum_{i=0}^{k} \frac{(\lambda_1(t_2 - t_1))^i \, e^{-\lambda_1(t_2 - t_1)}}{i!} \cdot \frac{(\lambda_2(t_2 - t_1))^{k-i} \, e^{-\lambda_2(t_2 - t_1)}}{(k-i)!} \\
&= \frac{((\lambda_1 + \lambda_2)(t_2 - t_1))^k \, e^{-(\lambda_1 + \lambda_2)(t_2 - t_1)}}{k!} \cdot \sum_{i=0}^{k} \frac{k!}{i!(k-i)!} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^i \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{k-i} \\
&= \frac{(\lambda(t_2 - t_1))^k \, e^{-\lambda(t_2 - t_1)}}{k!}
\end{aligned}
$$

i.e., the increment of $N(t)$ over the interval $(t_1, t_2]$ is Poisson distributed with mean $\lambda(t_2 - t_1)$. ■

It is opportune to state the following result, even though it depends on the concept of a stationary renewal process which we will introduce in Section 3.7. The "only if" part of this result (which we state without proof) can be viewed as the converse to Theorem 3.11.

Theorem 3.12. *The superposition of two independent stationary renewal processes is renewal iff they are both Poisson.*

Remarks 3.11.

(a.) For this result to hold, the necessity of the two independent renewal processes being Poisson can be seen intuitively as follows. Consider a point in the superposition. It belongs to one of the independent processes, and the time until the next renewal in this process is independent of the past. Therefore, for the time until the next renewal in the superposition to be independent of the past, it is necessary that the residual life-time of the other renewal process is independent of the past, which requires that the life-times of the component processes are exponential.

(b.) It is not difficult to construct an example that demonstrates that the superposition of two dependent renewal processes can be renewal. Take an ordinary renewal process with non-exponentially distributed life times, and split it using Bernoulli splitting (as in the beginning of this section). The resulting two point processes are each a renewal process, but are dependent.

■

3.5 Regenerative Processes

Let us recall Example 3.4 and the remark following Theorem 3.7. We had the queue length process, $X(t)$, of an M/G/1 system with $X(0) = 0$, and we saw that there exists a random time Z_1 such that $X(t + Z_1), t \geq 0$, (i.e., the evolution of the queue length process after Z_1) (i) is *statistically* identical to the evolution of the process $X(t), t \geq 0$, and (ii) is independent of Z_1 and the past of the queue length process up to Z_1. We say that the process $X(t)$ regenerates at Z_1, and call such a process a regenerative process.

Definition 3.4. *A process $X(t), t \geq 0$, is called a* regenerative process *if there exists a stopping time T_1 such that*

(a.) $X(t + T_1),\ t \geq 0$, *and* $X(t),\ t \geq 0$, *have the same probability law (i.e., are statistically identical), and*

(b.) $\{X(t + T_1),\ t \geq 0\} \amalg \{T_1, \text{and } X(u),\ u < T_1\}$.

■

Example 3.5.

(a.) Consider a DTMC $X_n, n \geq 0$, with $X_0 = j$, then the time of first return to j is a stopping time. By the strong Markov property, we see that X_n is a regenerative process.

(b.) As we saw in Example 3.4 and the remark following Theorem 3.7, the queue length process $X(t)$ of an M/G/1 queue (that starts off empty) is a regenerative process.

Remarks 3.12.

(a.) It is important to note that the property that $X(t + T_1), t \geq 0$, is independent of T_1 is a rather special property. In general, the evolution after a stopping time need not be independent of the stopping time. As an exercise, the reader is encouraged to construct an example that illustrates this lack of independence, in general.

(b.) With reference to Definition 3.4, we say that $X(t)$ *regenerates* at T_1. Since $X(t + T_1), t \geq 0$, is statistically identical to $X(t), t \geq 0$, there must be a stopping time $T_2 \geq T_1$, such that the process again regenerates at T_2. By the fact that the successive regenerations are independent, it follows that the sequence of regeneration points $T_1, T_2, T_3, \cdots$, are the renewal points in a renewal process. We call this the *embedded* renewal process. Also, the successive intervals into which these renewals divide time are called *regeneration cycles*.

(c.) In general, we have what are called *delayed regenerative processes*. For such a process, there exists a stopping time T_1 such that Property 2 in Definition 3.4 holds but $X(t+T_1), t \geq 0$, has a statistical behaviour different from $X(t), t \geq 0$. However, $X(t+T_1), t \geq 0$, is a regenerative process, as defined in Definition 3.4. Thus it as if the regenerations in the process $X(t)$ have been delayed by the time T_1. We observe now that the sequence of random times $T_1, T_2, \cdots$, are the points of a renewal process whose first life-time has a different distribution from the subsequent life-times. As an example, consider a DTMC $X_k, k \geq 0$, with $X_0 = i$, and consider visits to state $j \neq i$. The visits to j constitute regeneration times, but clearly this is a delayed regenerative process. Thus the process defined in Definition 3.4 can be called an *ordinary regenerative process*.

■

3.5.1 Time Averages of a Regenerative Process

Let us consider a regenerative process $X(t), t \geq 0$. We are interested in evaluating limits of the following type

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t X(u)du$$

i.e., the time average of the process $X(t)$, or, for $b \in \mathbb{R}$,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t I_{\{X(u)\leq b\}}du$$

i.e., the fraction of time the process value is in the interval $(-\infty, b]$, or

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t P\{X(u)\leq b\}du$$

i.e., the limit of the expectation of the fraction of time the process value is in the interval $(-\infty, b]$. Considering the first limit, we can proceed by thinking of $X(t)$ as a “reward” rate at time t. Then the cumulative reward until t is

$$C(t) = \int_0^t X(u)du$$

and we need, the time average reward rate, i.e.,

$$\lim_{t\to\infty} \frac{C(t)}{t}$$

As before, denote the regeneration times of $X(t)$ by $T_1, T_2, \cdots$, and let $T_0 = 0$. Define the reward in the jth cycle, $j \geq 1$, by

$$R_j = \int_{T_{j-1}}^{T_j} X(u)du$$

By the properties of the regenerative process, the sequence of random variable $R_j, j \geq 1$, are mutually independent, and for an ordinary regenerative process this random sequence is i.i.d. Thus, along with the renewal instants $T_k, k \geq 1$, we are now in the renewal-reward frame work. Let us consider the case of an ordinary regenerative process. If $\mathsf{E}\left(\int_0^{T_1} |X(u)|du\right) < \infty$, and $\mathsf{E}(T_1) < \infty$, Theorem 3.2 immediately applies and we conclude that, w.p. 1,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t X(u)du = \frac{\mathsf{E}\left(\int_0^{T_1} X(u)du\right)}{\mathsf{E}(T_1)}$$

and, w.p. 1,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t I_{\{X(u)\leq b\}}du = \frac{\mathsf{E}\left(\int_0^{T_1} I_{\{X(u)\leq b\}}du\right)}{\mathsf{E}(T_1)}$$

and

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t P(X(u)\leq b)du = \frac{\mathsf{E}\left(\int_0^{T_1} I_{\{X(u)\leq b\}}du\right)}{\mathsf{E}(T_1)}$$

Note that, in order to make the last two assertions, we only need that $\mathsf{E}(T_1) < \infty$, since

$$\mathsf{E}\left(\int_0^{T_1} I_{\{X(u)\leq b\}}du\right) \leq \mathsf{E}\left(\int_0^{T_1} du\right) = \mathsf{E}(T_1) < \infty$$

Thus for the limiting time average distribution to exist it suffices that the cycle times of the regenerative process have finite mean.

Let us observe that $\frac{\mathsf{E}\left(\int_0^{T_1} I_{\{X(u)\leq b\}}du\right)}{\mathsf{E}(T_1)}$ is a cumulative distribution function in the argument b, for it is nonnegative, nondecreasing in b, and the limit as $b \to \infty$ is 1. Note also that in the expression $\mathsf{E}\left(\int_0^{T_1} I_{\{X(u)\leq b\}}\right) du$ we cannot bring the expectation inside the integral, since the upper limit of integration is a random variable, i.e., T_1.

It is usually easier to obtain such limiting distributions. For example, for a positive recurrent DTMC we know that (recall the proof of Theorem 2.9)

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} p_{ij}^{(k)} \to \pi_j$$

Can we use this result to obtain the limiting time average of the process, i.e., $\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} X_k$? The following result asserts that, under certain conditions, the limiting time average of the process can be obtained by taking the expectation of the limiting time average distribution.

Theorem 3.13. *Consider a regenerative process $X(t)$ with $\mathsf{E}(T_1) < \infty$ and let X_∞ denote a random variable with the limiting time average distribution, i.e., $P(X_\infty \le b) = \frac{\mathsf{E}\left(\int_0^{T_1} I_{\{X(u)\le b\}}du\right)}{\mathsf{E}(T_1)}$.*

(a.) If $\mathsf{E}\left(\int_0^{T_1} |X(u)|du\right) < \infty$ then

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t X(u)du \overset{a.s.}{=} \mathsf{E}(X_\infty)$$

(b.) If $X(u) \ge 0$ then

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t X(u)du \overset{a.s.}{=} \mathsf{E}(X_\infty)$$

and this limit could be finite or infinite.

Remark: Thus, continuing the example that we were discussing just before the theorem, we can conclude that for a positive recurrent DTMC $X_k \in \mathcal{S}$ (that takes nonnegative values), w.p. 1,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = \sum_{j\in\mathcal{S}} j\pi_j$$

where π_j is the invariant distribution of the DTMC.

Proof: We omit the proof of this theorem. ■

3.6 The Renewal Equation

We have so far only examined the limiting time averages of various processes and functions. For example, for a renewal process we obtained the limiting time average of the renewal function $m(t)$, or of the distribution of the residual life process $P(Y(t) > y)$. We now study how to characterise the transient behaviour of these processes and functions, and hence to obtain their limiting behaviour. For example, if we can show that for a renewal process

$$\lim_{t \to \infty} P(Y(t) > y) = G^c(y)$$

where $G(\cdot)$ is a continuous distribution, then we have shown the convergence in distribution of the residual life process $Y(t)$.

Example 3.6.

Consider a renewal process with i.i.d. life-times with c.d.f. $F(\cdot)$. Define, for $y \geq 0$,

$$K(t) = P(Y(t) > y)$$

Now we can break up this event into three disjoint parts depending on whether the first renewal occurs in either of the three intervals $(t+y, \infty)$, $(t, t+y]$, $[0, t]$. This yields

$$\begin{aligned} K(t) &= P(Y(t) > y,\ X_1 > t+y) + P(Y(t) > y,\ t < X_1 \leq t+y) \\ &\quad + P(Y(t) > y,\ X_1 \leq t) \\ &= (1 - F(t+y)) + 0 + \int_0^t K(t-x)dF(x) \end{aligned}$$

where, in the second equality, the first term is $(1-F(t+y))$ because the event $\{X_1 > t+y\}$ implies the event $\{Y(t) > y\}$, and the second term is 0 because $\{t < X_1 \leq t+y\}$ implies that $Y(t) \leq y$. In the third term we integrate over all the possible points in $[0, t]$ at which the first renewal can occur. With the first renewal having occurred at x, the time remaining until t is $t - x$, and hence the desired probability is $K(t-x)$. This argument is called a *renewal argument*. Thus we find that, for fixed $y \geq 0$, the function $K(t)$ satisfies the equation

$$K(t) = F^c(t+y) + \int_0^t K(t-x)dF(x) \tag{3.5}$$

or, equivalently,

$$K(t) = a(t) + (K \star F)(t)$$

where $a(t) = F^c(t+y)$, and, as before, $\star$ denotes the Riemann-Stieltjes convolution. ■

The equation for $K(t)$ obtained in the previous example is of the form

$$H(t) = a(t) + \int_0^t H(t-x)dF(x)$$

or

$$H(t) = a(t) + (H \star F)(t)$$

where $a(\cdot)$ is a given function and $F(\cdot)$ is a cumulative distribution function. Such an equation is called a *renewal equation*[3]. More compactly, taking the time argument as being understood, we also write

$$H = a + H \star F$$

The following result provides a formal solution for the renewal equation, under some fairly general conditions.

Theorem 3.14. *If the c.d.f. $F(\cdot)$ has positive mean (i.e., $\int_0^\infty (1 - F(x))dx > 0$), and if $a(t)$ is a bounded function, then the unique solution of $H = a + H \star F$ that is bounded on finite intervals is*

$$H(t) = a(t) + \int_0^t a(t-u)dm(u)$$

i.e.,

$$H = a + a \star m$$

where $m(t)$ is the renewal function for an ordinary renewal process with life-time c.d.f. $F(\cdot)$, i.e., $m(t) = \sum_{k=1}^{\infty} F^{(k)}(t)$.

Proof: We need to show that the proposed solution

(i) is indeed a solution,

(ii) is bounded on finite intervals, and

(iii) is unique.

(i) To check that the proposed solution satisfies the renewal equation, we substitute it in the renewal equation to obtain

$$\begin{aligned} H &= a + (a + a \star m) \star F \\ &= a + (a \star F + a \star (m \star F)) \\ &= a + a \star (F + m \star F) \\ &= a + a \star m \end{aligned}$$

where we have used the easy observation that $m = F + m \star F$. Thus the proposed solution satisfies the renewal equation.

[3] In general, an equation in which the unknown function appears inside an integral is called an *integral equation*, and, in particular, the renewal equation is called a *Volterra integral equation of the second kind.*

(ii) To verify that the proposed solution is bounded on finite intervals consider, for a finite $T > 0$,

$$\begin{aligned} \sup_{0\le t\le T} |a(t) + (a \star m)(t)| &\le \sup_{0\le t\le T} |a(t)| + \sup_{0\le t\le T} \left|\int_0^t a(t-u)dm(u)\right| \\ &\le \sup_{0\le t\le T} |a(t)| + \int_0^T \sup_{0\le y\le T} |a(y)|dm(u) \\ &= \left(\sup_{0\le t\le T} |a(t)|\right)(1 + m(T)) < \infty \end{aligned}$$

where the finiteness follows since $m(t)$ is bounded (by $\int_0^\infty (1 - F(x))dx > 0$ and Lemma 3.1), and $a(t)$ is given to be bounded.
(iii) To establish uniqueness, assume to the contrary that H_1 and H_2 are two solutions bounded on finite intervals and let

$$D = H_1 - H_2$$

Hence D is also bounded on finite intervals, then

$$\begin{aligned} D \star F &= H_1 \star F - H_2 \star F \\ &= (H_1 - a) - (H_2 - a) \\ &= D \end{aligned}$$

where the second equality is obtained since H_1 and H_2 both are solutions to $H = a + H \star F$. It follows, by recursion, that, for all $n \ge 1$,

$$D = D \star F^{(n)}$$

Hence for all $n \ge 1$

$$\begin{aligned} |D(t)| &= \left|\int_0^t D(t-x)dF^{(n)}(x)\right| \\ &\le \int_0^t |D(t-x)|dF^{(n)}(x) \\ &\le \left(\sup_{u\in[0,t]} |D(u)|\right) F^{(n)}(t) \end{aligned}$$

Now the first term in the last expression is finite, since D is bounded on finite intervals. Also, for each fixed t, $F^{(n)}(t) \to 0$ as $n \to \infty$ since $\int_0^\infty (1 - F(x))dx > 0$ (see the proof of Lemma 3.1). It follows that, for all t,

$$|D(t)| = 0$$

We conclude that $H_1 = H_2$, i.e., the proposed solution is unique. ■

Remarks 3.13.

(a.) Let us now apply Theorem 3.14 to solve the renewal Equation 3.5 for $K(t) = P(Y(t) > y)$. This yields

$$K(t) = (1 - F(t+y)) + \int_0^t (1 - F(t+y-x))\, dm(x)$$

(b.) There is an intuitive way to think about the solution of the renewal equation provided we appropriately interpret $dm(x)$. To this end, consider the following formal "calculation."

$$\begin{aligned} P(\text{a renewal occurs in the interval } (x, x+dx)) &= \sum_{k=1}^{\infty} P(Z_k \in (x, x+dx)) \\ &= \sum_{k=1}^{\infty} dF^{(k)}(x) \\ &= d(\sum_{k=1}^{\infty} F^{(k)}(x)) \\ &= dm(x) \end{aligned}$$

Thus we can interpret $dm(x)$ as the probability that there is a renewal in the interval $(x, x+dx)$. For this reason $\frac{dm(x)}{dx}$ is called the *renewal density*. As an example, consider the Poisson process, which we now know is a renewal process with exponentially distributed life-times. We also know that for a Poisson process of rate λ the renewal function $m(t) = \lambda t$. Hence the renewal density is λ. Now Lemma 3.5 confirms our interpretation of $dm(x)$ in the case of the Poisson process. With this interpretation of $dm(x)$ let us now view the second term in the solution of the renewal equation as conditioning on the *last* renewal in $[0, t]$. This will yield

$$\begin{aligned} K(t) &= (1 - F(t+y)) \\ &\quad + \int_0^t P(\text{a renewal occurs in } (x, x+dx), \\ &\qquad\qquad \text{the next renewal occurs after } t, \\ &\qquad\qquad \text{and the residual life-time at } t \text{ is greater than } y) \\ &= (1 - F(t+y)) + \int_0^t (1 - F((t-x)+y))\, dm(x) \end{aligned}$$

■

Let us now consider a delayed renewal process.

Example 3.7.

Let us redo Example 3.6 for a delayed renewal process with the c.d.f. of X_1 being $A(\cdot)$ and that of $X_j, j \geq 2$, being $F(\cdot)$. Again, for fixed $y \geq 0$, define $K(t) = P(Y(t) > y)$, where $Y(t)$ is the residual life process. In an identical fashion as in Example 3.6, we can write

$$K(t) \quad = \quad (1 - P(X_1 \leq t + y)) + \int_0^t K_o(t - x)\, dA(x)$$

where $K_o(t) = P(Y(t) > y)$ for the ordinary renewal process with life-time distribution $F(\cdot)$. Thus we have

$$K(t) = A^c(t + y) + (K_0 \star A)(t)$$

where, in turn, $K_o(t)$ satisfies

$$K_o(t) = F^c(t + y) + (K_0 \star F)(t)$$

■

Thus, in situations involving a delayed renewal process, we obtain a pair of equations of the form

$$H = a + H_0 \star A$$

in conjunction with

$$H_o = a_o + H_o \star F$$

Of course, we know that the equation for $H_o(\cdot)$ has the solution

$$H_o = a_o + a_o \star m_0$$

where

$$m_o = \sum_{k=1}^{\infty} F^{(k)}$$

The solution of the delayed renewal equation is given in the following result, and can be intuitively appreciated based on the interpretation of $\frac{dm(t)}{dt}$ as the renewal density for the delayed renewal process.

Theorem 3.15. *The solution to the general renewal equation is*

$$H = a + a_0 \star m$$

where

$$m = \sum_{k=0}^{\infty} A \star F^{(k)}$$

Proof: It is easy to check that the proposed solution indeed satisfies the delayed renewal equation. We omit the proof of uniqueness. ■

3.7 Stationary Renewal Process

Obviously, a renewal counting process $M(t)$ cannot be stationary since it is increasing with time. What then is meant by a stationary renewal process? Consider a general renewal process with X_j distributed as $F(\cdot)$ for $j \geq 2$, and X_1 distributed as $F_e(\cdot)$ where $F_e(t) = \frac{1}{\mathsf{E}(X_2)} \int_0^t (1 - F(x))dx$. Recall that $F_e(\cdot)$ is the time average equilibrium distribution for a renewal process with i.i.d. life-times with distribution $F(\cdot)$. Denote the renewal function for this delayed renewal process by $m_e(t)$.

Let us apply the result in Theorem 3.15 to calculate $P(Y(t) > y)$ for this general renewal process (for some fixed $y > 0$). We obtain

$$P(Y(t) > y) = F_e^c(t + y) + \int_0^t F^c(t + y - x)\, dm_e(x)$$

Hence we need to determine $m_e(t)$. This is easily done by writing a renewal equation for $m_e(\cdot)$. We observe that the following holds

$$m_e(t) \quad = \quad \mathsf{E}(M_e(t)) = \int_0^t (1 + m_o(t - x))dA(x)$$

This can be understood as follows. If the first arrival occurs in (t, ∞) then the mean number of arrivals in $[0, t]$ is 0. Hence we condition on the first renewal in $[0, t]$. We count 1 for this renewal, and then from here on we have an ordinary renewal process. If the first renewal occurs at x, the remaining mean number of renewals in $[x, t]$ is $m_o(t - x)$. Thus we can write, compactly,

$$m_e(t) = A(t) + (m_o \star A)(t)$$

Taking Laplace Stieltjes Transforms (LSTs) across this equation (see Section 1.2.3), we obtain,

$$\begin{aligned} \tilde{m}_e(s) &= \tilde{A}(s) + \tilde{m}_o(s) \cdot \tilde{A}(s) \\ &= \tilde{A}(s) + \frac{\tilde{F}(s)}{1 - \tilde{F}(s)} \cdot \tilde{A}(s) \\ &= \frac{\tilde{A}(s)}{1 - \tilde{F}(s)} \end{aligned}$$

where the second term in the second equality is obtained by applying the first equality to $m_o(\cdot)$. Further, it can be seen by taking the LST that

$$\tilde{A}(s) = \frac{1 - \tilde{F}(s)}{s\mathsf{E}(X_2)}$$

Substituting, we obtain

$$\tilde{m}_e(s) = \frac{1}{s\mathsf{E}(X_2)}$$

which, on inversion, yields

$$m_e(t) = \frac{t}{\mathsf{E}(X_2)}$$

Remark: This result should be compared with what we know for the Poisson process of rate λ. The Poisson process is a renewal process with exponentially distributed life-times with c.d.f. $F(x) = 1 - e^{-\lambda x}$. It can be verified that $F_e(x) = 1 - e^{-\lambda x}$. Thus the Poisson process automatically satisfies the assumption in the previous derivation. Hence, applying the conclusion of the previous derivation, the renewal function will be $\frac{t}{\mathsf{E}(X_2)} = \lambda t$, which we already know to be true.

Let us now return to obtaining $P(Y(t) > y)$ for the renewal process that we had constructed above. Substituting $m_e(\cdot)$ into Equation 3.6, we obtain (via self evident manipulations)

$$\begin{aligned}
P(Y(t) > y) &= F_e^c(t+y) + \int_0^t F^c(t+y-x)\frac{dx}{\mathsf{E}(X_2)} \\
&= F_e^c(t+y) + \frac{1}{\mathsf{E}(X_2)}\int_y^{t+y} F^c(u)du \\
&= F_e^c(t+y) + \frac{1}{\mathsf{E}(X_2)}\left(\int_0^{t+y} F^c(u)du - \int_0^{y} F^c(u)du\right) \\
&= F_e^c(t+y) + F_e(t+y) - F_e(y) \\
&= 1 - F_e(y) = F_e^c(y)
\end{aligned}$$

Thus for the renewal process that we constructed, for all $t \geq 0$,

$$P(Y(t) > y) = F_e^c(y)$$

i.e., the marginal distribution of the residual life process is invariant with time. It is easy to see that $Y(t)$ is a Markov process. It follows that, with the marginal distribution being stationary, the $Y(t)$ process is itself a stationary process. It is for this reason that the process we have constructed is called a *stationary* renewal process.

Now, denoting the counting process for this renewal process by $M_e(\cdot)$, let us consider, for $x > 0$, and $t > 0$, $M_e(x)$ and $M_e(t+x) - M_e(t)$. Considering t as the starting time, since the distribution of $Y(t)$ is the same as that of $Y(0)$, we have another renewal process statistically identical to the one that started at time 0. Hence we conclude that $M_e(x)$ and $M_e(t+x) - M_e(t)$ have the same distributions, and thus $M_e(t)$ has *stationary increments*.

Remark: It is important to note, however, that $M_e(t)$ need not have independent increments. As an example, suppose that $X_j, j \geq 2$, are deterministic random variables taking the value T with probability 1. Then $F_e(t)$ is the uniform distribution over $[0, T]$, and $M_e(t)$ is a stationary increment process. But we notice that in $[0, T]$ there is exactly one renewal. Hence $M(T/2)$ and $M(T) - M(T/2)$ are not independent, even though they have the same distribution.

Definition 3.5. *A point process with stationary and independent increments is called a* compound *or* batch *Poisson process.*

Remark: Basically a batch Poisson arrival process comprises a Poisson process, at the points of which i.i.d. batches of arrivals occur. Each batch has at least one arrival. If the batch size is exactly one then we are back to a Poisson process. Thus, in summary, a general stationary renewal process has stationary increments. If the property of independent increments is added, we obtain a batch Poisson point process. If, in addition, each batch is exactly of size 1, we have a Poisson point process.

3.8 From Time Averages to Limits

Until this point in our discussions, we have focused on obtaining long run time averages of processes and certain time varying quantities associated with these processes. Results such as the elementary renewal theorem, or the renewal reward theorem were the tools we used in establishing the existence of and the forms of such limits. In this section we turn to the important question of studying the limits of the quantities themselves. The approach is by obtaining a renewal equation for the desired quantity, solving this renewal equation and then taking the limit as $t \to \infty$ in the solution. The Key Renewal Theorem is an important tool in this approach.

The following example shows a typical situation in which the time average may exist, with the limit of the associated quantity failing to exist.

Example 3.8.

Consider an ordinary renewal process with life-times $X_i, i \geq 1$, such that $X_i = T$ for all $i \geq 1$, where $T > 0$ is given. It is elementary to see that, given $y \geq 0$,

$$P(Y(t) \leq y) = \begin{cases} 0 & (Tk \leq t < T(k+1) - y \;\; k \geq 0 \\ 1 & \text{otherwise} \end{cases}$$

Hence $P(Y(t) \leq y)$ is an "oscillatory" function of time and does not converge with t. However, as we already know, for $0 \leq y \leq T$,

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t P(Y(t) \leq y) dy = \frac{y}{T}$$

the uniform distribution over $[0, T]$. ■

The life-time in this example has a property that is captured in the following definition.

Definition 3.6. *A random variable X, such that $P(|X| < \infty) = 1$, is called* lattice *if there exists $d > 0$ such that*

$$\sum_{n=-\infty}^{\infty} P(X = nd) = 1$$

otherwise X is called nonlattice. The largest d with this property is called the span *of X. We also say that the distribution of X is lattice.* ■

Notice that in the example above the renewal life-time was lattice with $P(X_1 = T) = 1$.

We state the following result without proof. This result helps in the understanding of the Key Renewal Theorem which we will state and discuss next.

Theorem 3.16 (Blackwell's Renewal Theorem (BRT))**.** *(i) For a general (or delayed) renewal process with arbitrary A and nonlattice F, and any $h > 0$.*

$$\lim_{t \to \infty} [m(t + h) - m(t)] = h\mu$$

where $\frac{1}{\mu} = \int_0^\infty (1 - F(x))dx$.

(ii) For an ordinary renewal process, i.e., $A = F$, with lattice F with span d, for $h = d, 2d, 3d, \cdots$,

$$\lim_{t \to \infty} [m(t + h) - m(t)] = \mu h$$

Remarks 3.14.

(a.) We first note that part (ii) of the theorem deals with the case of a lattice F, and in this case the result is restricted to the situation in which $A = F$. This is because without such a restriction on the "initial" life-time, different conclusions can be obtained depending on the choice of A. For example, if $A = F_e$ then (recall from Section 3.7) we have a stationary renewal process and $m(t) = \mu t$ for all t, thus yielding $m(t + h) - m(t) = \mu h$ for every t and $h \geq 0$.

(b.) We also note that, in the same sense as the convergence of a sequence of numbers is stronger than the convergence of the averages of the sequence, BRT is stronger than the expectation version of ERT. This can be seen by the following simple argument that shows that BRT implies ERT. Consider the nonlattice case of BRT (part (i), above). We observe that, for every n, with $m(0) = 0$,

$$m(n) = \sum_{k=0}^{n-1} (m(k + 1) - m(k))$$

Now, by BRT, $m(k+1) - m(k) \to \mu$, as $k \to \infty$. Hence, the average of this sequence also converges, i.e.,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (m(k+1) - m(k)) = \mu$$

i.e.,

$$\lim_{n\to\infty} \frac{m(n)}{n} = \mu$$

which establishes ERT for the case when we pass to the limit via integer values of time. But we also see that

$$\frac{\lfloor t \rfloor}{t} \frac{m(\lfloor t \rfloor)}{\lfloor t \rfloor} \leq \frac{m(t)}{t} \leq \frac{m(\lfloor t \rfloor + 1)}{\lfloor t \rfloor + 1} \frac{\lfloor t \rfloor + 1}{t}$$

Since the upper and lower bounds both converge to μ, we conclude that $\lim_{t\to\infty} \frac{m(t)}{t} = \mu$, which is the expectation version of ERT.

■

We have seen that the solution of the renewal equation

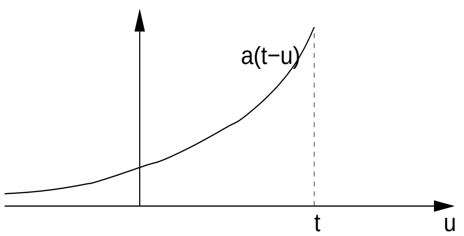
$$H = a + H \star F$$

is

$$H = a + a \star m$$

where $m(t)$ is the renewal function of the ordinary renewal process with life-time distribution F. Our aim next is to understand how to use this solution to obtain $\lim_{t\to\infty} H(t)$. It is usually straightforward to obtain $\lim_{t\to\infty} a(t)$; recall that, in the examples we have seen, $a(t)$ is defined by the tail of a distribution function and hence goes to 0 as $t \to \infty$. The problem remains of obtaining $\lim_{t\to\infty}(a \star m)(t)$. So let us consider

$$(a \star m)(t) = \int_0^t a(t-u)\, dm(u)$$



The diagram on the left shows the way a convolution is performed between a function $a(\cdot)$, which is 0 for negative arguments, and the renewal function $m(t)$. For fixed $t > 0$, the diagram shows $a(t-u)$ as a function of u. Multiplying $a(t-u)$ with $dm(u)$ for $u \geq 0$ and integrating up to t yields the value of the convolution at t.

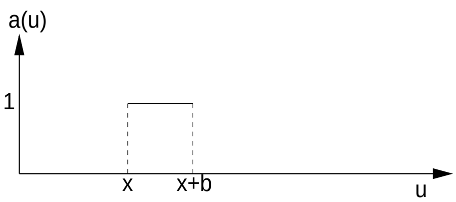
Now suppose that for large u, $dm(u) \approx \mu du$, as would be suggested by BRT. Suppose, for a moment that $a(\cdot)$ is 0 for large arguments. Then, for large enough t, over the range of integration, the approximation $dm(u) \approx \mu du$ will be better and better as t increases. To see this, look at the picture and note that if $a(\cdot)$ is 0 for large arguments, for large enough t all of the nonzero part of it will be "pushed" into the positive quadrant, and larger values of t will push this nonzero part further to the right, thus multiplying it with $dm(u)$ for larger and larger u. Thus, we can write, for large t,

$$\begin{aligned}(a \star m)(t) &\approx \int_0^t a(t-u)\, \mu du \\ &= \mu \int_0^t a(x)dx\end{aligned}$$

This suggests the following result

$$\lim_{t\to\infty} (a \star m)(t) = \mu \int_0^\infty a(x)dx$$

In general, we would like to permit $a(\cdot)$ to be nonzero over all $\mathbb{R}^+$. The Key Renewal Theorem provides an important technical condition that constrains the behaviour of $a(t)$ for large t, and permits the above limiting result to hold. We develop an intuition into this condition before stating the theorem.



Consider the function $a(u)$ shown on the left, i.e.,

$$a(u) = \begin{cases} 1 & \text{if} \quad x \le u \le x+b \\ 0 & \text{otherwise} \end{cases}$$

Consider $\int_0^t a(t-u)dm(u)$. As discussed earlier, this converges to $\mu \int_0^\infty a(u)du = \mu b$. Note that this is also a consequence of BRT (Theorem 3.16).

Next we consider a bounded function, $a(t)$, defined for $t \ge 0$. Define the following function, for all $i \ge 1$,

$$\underline{a}_i(u) = \begin{cases} \inf_{x\in[(i-1)b,ib)} a(x) & \text{for } u \in [(i-1)b, ib) \\ 0 & \text{otherwise} \end{cases}$$

and consider

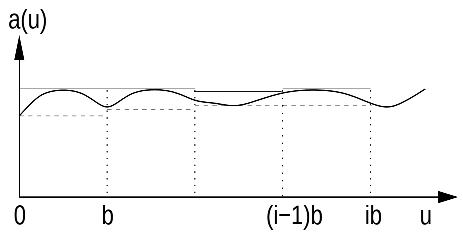
$$\underline{a}(u) = \sum_{i=1}^{\infty} \underline{a}_i(u)$$

Similarly, define, for all $i \geq 1$,

$$\overline{a}_i(u) = \begin{cases} \sup_{x\in[(i-1)b,ib)} a(x) & \text{for } u \in [(i-1)b, ib) \\ 0 & \text{otherwise} \end{cases}$$

and consider

$$\overline{a}(u) = \sum_{i=1}^{\infty} \overline{a}_i(u)$$



We show on the left a general function $a(t)$. The positive x-axis is partitioned into equal intervals of width b. The piece-wise flat function drawn with thin solid lines is $\overline{a}(\cdot)$ and the piece-wise flat function drawn with thin dashed lines is $\underline{a}(\cdot)$.

By the definitions of $\underline{a}(u)$ and $\overline{a}(u)$, it is clear that, for all t,

$$\int_0^t \underline{a}(t-u)dm(u) \leq \int_0^t a(t-u)dm(u) \leq \int_0^t \overline{a}(t-u)dm(u)$$

Thus, if it can be shown that the upper and lower bounds in this expression converge to the same quantity, as $t \to \infty$, then that will be $\lim_{t\to\infty} \int_0^t a(t-u)dm(u)$.

Now we can expect the following to hold

$$\lim_{t\to\infty} \int_0^t \underline{a}(t-u)dm(u) = \lim_{t\to\infty} \int_0^t \sum_{i=1}^{\infty} \underline{a}_i(t-u)dm(u) = \mu b \sum_{i=1}^{\infty} \inf_{x\in[(i-b),\, ib)} a(x)$$

and, similarly,

$$\lim_{t\to\infty} \int_0^t \overline{a}(t-u)dm(u) = \lim_{t\to\infty} \int_0^t \sum_{i=1}^{\infty} \overline{a}_i(t-u)dm(u) = \mu b \sum_{i=1}^{\infty} \sup_{x\in[(i-1)b,\, ib)} a(x)$$

Notice that the expressions $b\ \sum_{i=1}^{\infty}\ \inf_{x\in[(i-b),\, ib)} a(x)$ and $b\ \sum_{i=1}^{\infty} \sup_{x\in[(i-1)b,\, ib)} a(x)$ are like lower and upper Riemann partial sums, except that they are taken over all $[0,\infty)$. Now letting $b \to 0$, if the upper and lower partial sums over $[0,\infty)$ converge to the same number then that must be equal to $\int_0^{\infty}\ a(u)du$ the Riemann integral, and a is said to be Directly Riemann Integrable (DRI). Thus if $a(\cdot)$ is DRI, the above argument suggests that $\lim_{t\to\infty} \int_0^t a(t-u)dm(u) = \mu \int_0^{\infty} a(u)du$.

Remarks 3.15.

(a.) Note that in standard Riemann integration we obtain $\int_0^\infty a(u)du = \lim_{t\to\infty} \int_0^t a(u)du$, where $\int_0^t a(u)du$ is the Riemann integral over $[0, t]$, obtained by taking partial sums over a partition of $[0, t]$ and then letting the partition width (say, b) go to 0. Thus standard Riemann integrability over $[0, \infty)$ requires

$$\lim_{t\to\infty} \lim_{b\to 0} b \sum_{i=1}^{\lceil t/b \rceil} \inf_{x\in[(i-b),\, ib)} a(x) = \lim_{t\to\infty} \lim_{b\to 0} b \sum_{i=1}^{\lceil t/b \rceil} \sup_{x\in[(i-b),\, ib)} a(x)$$

On the other hand in direct Riemann integration over $[0, \infty)$ we set up the partial sums by partitioning all of $[0, \infty)$, and then letting the partition width $b \to 0$. Thus direct Riemann integrability over $[0, \infty)$ requires

$$\lim_{b\to 0} \lim_{t\to\infty} b \sum_{i=1}^{\lceil t/b \rceil} \inf_{x\in[(i-b),\, ib)} a(x) = \lim_{b\to 0} \lim_{t\to\infty} b \sum_{i=1}^{\lceil t/b \rceil} \sup_{x\in[(i-b),\, ib)} a(x)$$

Note that the two requirements differ in the order in which the limits with respect to b and t are taken.

(b.) It can be shown that if $a(u)$ is DRI over $[0, \infty)$ then it is Riemann integrable but a Riemann integrable function over $[0, \infty)$ need not be DRI. Here is an example of a Riemann integrable function that is not DRI. Define $a(u) = 1$ for $u \in [n - \frac{1}{2n^2}, n + \frac{1}{2n^2}]$, for $n = 1, 2, 3, \cdots$, and $a(u) = 0$ for all other $u \in \mathbb{R}^+$. Thus, the graph of $a(u)$ comprises "pulses" of height 1 and width $\frac{1}{n^2}$ centred at the positive integers. The Riemann integral of $a(u)$ exists since $\lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n^2} < \infty$. However, $a(u)$ is not DRI since, for each b, there exists N_b, such that for all $n > N_b$, $\frac{1}{n^2} < b$, and hence the upper Riemann sum over $[0, \infty)$ will have an infinite number of 1s contributed by the pulses centered at $n > N_b$ and is, therefore, ∞ for all b.

■

In general, the DRI property may not be easy to recognise, but the following result is useful in many applications. We state the result without proof.

Lemma 3.6. *If a function $a : \mathbb{R}^+ \to \mathbb{R}$ is (i) nonnegative, (ii) monotone nonincreasing, and (iii) Riemann integrable, then it is directly Riemann integrable.*

■

Finally, we state the following important result without proof.

Theorem 3.17 (Key Renewal Theorem (KRT)). *Let A and F be proper distribution functions with $F(0^+) < 1$ and $\int_0^\infty (1 - F(u))du = \frac{1}{\mu}$ with $\mu = 0$ if the integral is infinite. Suppose $a(\cdot)$ is directly Riemann integrable. Then the following hold.*

(i) *If F is non lattice then*

$$\lim_{t\to\infty} \int_0^t a(t-u)dm(u) = \mu \int_0^\infty a(u)du$$

(ii) *If $A = F$, and if F is lattice with span d, then for all $t > 0$*

$$\lim_{n\to\infty} \int_0^{t+nd} a(t+nd-u)dm(u) = \mu d \sum_{n=0}^{\infty} a(t+nd)$$

■

Example 3.9.

Consider a delayed renewal process with initial life-time distribution $A(\cdot)$ and the subsequent life-time distribution being $F(\cdot)$ with mean $\mathsf{E}(X_2)$. We recall that, for given $y \geq 0$, the residual life distribution is given by

$$P(Y(t) > y) = A^c(t+y) + \int_0^t F^c(t+y-x)dm(x)$$

where $m(\cdot)$ is the renewal function. Suppose F is nonlattice and $A(\cdot)$ is a proper distribution. It follows from KRT that

$$\begin{aligned}\lim_{t\to\infty} P(Y(t) > y) &= 0 + \frac{1}{\mathsf{E}(X_2)} \int_0^\infty F^c(u+y)du \\ &= \frac{1}{\mathsf{E}(X_2)} \int_y^\infty (1-F(x))dx \\ &= F_e^c(y)\end{aligned}$$

which, as would be expected, is the same as the time average result. Notice that this result states that the random process $Y(t)$ converges in distribution to a random variable with distribution F_e. ■

3.9 Limits for Regenerative Processes

What we have learnt above can be applied to obtain a fairly general condition for the convergence in distribution of regenerative processes. Consider a delayed regenerative process $B(t),\ t \geq 0$. We are interested in $\lim_{t\to\infty} P(B(t) \leq b)$. There is a random time X_1 at which the process regenerates. Let the law of the process obtained after the regeneration be denoted by $P_o(\cdot)$, where the subscript "o" relates to the term "ordinary." To appreciate this notation, notice that we can write $A(t) = P(X_1 \leq t)$, and $F(t) = P_o(X_1 \leq t)$; thus for this delayed regenerative process $A(\cdot)$ is the distribution of the first cycle time and $F(\cdot)$ is the distribution of the cycle times after the first regeneration.

It is now easy to see that the following renewal equation is obtained for any $b \in \mathbb{R}$

$$P(B(t) \leq b) \quad = \quad P(B(t) \leq b,\ X_1 > t) + \int_0^t P_o(B(t-x) \leq b)\, dA(x)$$

where $A(\cdot)$ is the distribution of the time until the first regeneration instant. Let us denote $a(t) = P(B(t) \leq b, X_1 > t)$, and $a_o(t) = P_o(B(t) \leq b, X_1 > t)$. Then it can be seen that the renewal equation has the following solution

$$P(B(t) \leq b) = a(t) + \int_0^t a_o(t-x)dm(x)$$

where $m(t)$ is the renewal function of the delayed renewal process induced by the regenerative process. Now, if $a_o(t)$ is directly Riemann integrable, and $F(\cdot)$ is nonlattice then, by Theorem 3.17,

$$\lim_{t\to\infty} \int_0^t a_o(t-x)dm(x) = \mu \int_0^\infty a_o(u)du$$

where $\frac{1}{\mu} = \int_0^\infty (1 - F(x))dx$. Further, if $A(\cdot)$ is proper,

$$\lim_{t\to\infty} a(t) = 0$$

Therefore, in such situation, we can conclude that

$$\lim_{t\to\infty} P(B(t) \leq b) = \mu \int_0^\infty P_o(B(u) \leq b,\ X_1 > u)du$$

Hence, if $\mu > 0$ we find that the regenerative process converges in distribution to a proper distribution (why is it proper?). It remains to obtain a condition that ensures that $a_o(\cdot)$ is DRI. We need the following definition

Definition 3.7. *(a.) A real valued function is said to belong to $D[0, \infty)$ if it is right continuous and has left hand limits.*

(b.) *A stochastic process* $B(t),\ t \geq 0$, *is said to belong to* $\mathcal{D}$ *(written* $B(t) \in \mathcal{D}$*) if* $P\{w : B(t,w) \in D[0,\infty)\} = 1$

■

The following is the main result, which we state without proof. We state only the nonlattice version.

Theorem 3.18. *For a generalised regenerative process* $\{B(t), t \geq 0\}$, *with finite mean cycle length* $\mathsf{E}(X_i) < \infty,\ i \geq 2$, *and with* $A(\infty) = 1$, $B(t), t \geq 0$, *converges in distribution to*

$$\mu \int_0^\infty P_o(B(t) \leq b,\ X_1 > t)dt$$

if $B(t) \in \mathcal{D}$ *and* $F(\cdot)$ *is nonlattice.* ■

Remarks 3.16.

(a.) The above theorem provides a very simple but powerful tool for proving the stability of a stochastic process that can be shown to be regenerative. Under a fairly general condition, from the point of view of applications, it suffices to show that the mean cycle time is finite.

(b.) Let us look at the form of the limiting distribution and observe that

$$\int_0^\infty I_{\{B(t,w)\leq b,\ X_1(w)>t\}} dt = \int_0^{X_1(w)} I_{\{B(t,w)\leq b\}} dt$$

Taking expectation on both sides with respect to the probability law $P_o(\cdot)$, we obtain

$$\int_0^\infty P_o(B(t) \leq b,\ X > t)dt = \mathsf{E}_o \left(\int_0^{X_1} I_{\{B(t)\leq b\}} dt \right)$$

where $\mathsf{E}_o(\cdot)$ denotes expectation with respect to $P_o(\cdot)$. Then multiplying with μ on both sides we find that the limit provided by Theorem 3.18 is consistent with that obtained in Section 3.5.1.

3.10 Some Topics in Markov Chains

We now turn to applying some of the results in this chapter to DTMCs.

3.10.1 Relative Rate of Visits

Consider a positive recurrent irreducible Markov Chain $X_k, k \geq 0$, taking values in the discrete set $\mathcal{S}$. Suppose $X_0 = i$, and consider any other state j. Consider first the instants of visits to state j. The Markov chain is now a delayed regenerative process with respect to these instants. By the elementary renewal theorem, with probability 1,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} I_{\{X_k=j\}} = \frac{1}{\nu_j}$$

where $\nu_j = \sum_{n=1}^{\infty} n f_{jj}^{(n)}$, the mean recurrence time of j. Also, since the DTMC is positive recurrent, denoting by $\pi_j, j \in \mathcal{S}$, its stationary probability vector, we know that $\pi_j = \frac{1}{\nu_j}, j \in \mathcal{S}$. Consider now visits to i. These visits constitute an ordinary renewal process. Let the reward in each renewal interval be the number of visits to j. Then, by the renewal reward theorem, with probability 1,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} I_{\{X_k=j\}} = \frac{\mathsf{E}(V_{ij})}{\nu_i}$$

where V_{ij} is the random variable of the number of visits to j between visits to i. We have obtained the same limit in two forms, and these must be equal. Thus we obtain

$$\frac{\mathsf{E}(V_{ij})}{\nu_i} = \frac{1}{\nu_j}$$

or,

$$\mathsf{E}(V_{ij}) = \frac{\pi_j}{\pi_i}$$

Thus we conclude that in a positive recurrent DTMC the mean number of visits to state j between consecutive visits to the state i is the ratio of the stationary probabilities of these two states. This result is, however, not limited to positive recurrent DTMCs but holds more generally under the condition of recurrence alone. We state the following result which we shall have occasion to use in the theory of continuous time Markov chains.

Theorem 3.19. *For a recurrent irreducible Markov chain with transition probability matrix* $\mathbf{P}$*, there exist solutions to* $\mathbf{u} = \mathbf{u}\mathbf{P}$*,* $\mathbf{u} > 0$*. The vector space of such solutions has dimension 1. Thus, for all i and j, the ratios* $\frac{u_j}{u_i}$ *are uniquely determined and*

$$\mathsf{E}(V_{ij}) = \frac{u_j}{u_i}$$

■

Remark: Note that if the recurrent DTMC is positive then the solutions of $\mathbf{u} = \mathbf{u}\mathbf{P}$, $\mathbf{u} > 0$, will also be *summable* thus yielding a probability measure on $\mathcal{S}$. In general, however, for a recurrent DTMC the solutions of $\mathbf{u} = \mathbf{u}\mathbf{P}$, $\mathbf{u} > 0$, will not be summable, but the ratios of their components, and hence the relative frequency of visits to states, are uniquely determined.

3.10.2 Limits of DTMCs

Consider again an irreducible recurrent DTMC $X_n, n \geq 0, X_n \in \mathcal{S}$. Let us view the DTMC as evolving over multiples of a time-step, say 1. In this view point, all times between visits to various states are lattice random variables. For example, letting T_j denote the random variable for the time to return to j, we observe that $P(T_j = n) = f_{jj}^n$, and hence

$$f_{jj} = \sum_{n=1}^{\infty} P(T_j = n) = \sum_{n=1}^{\infty} f_{jj}^{(n)} = 1$$

since we have assumed a recurrent DTMC. Thus T_j is a lattice random variable. Let d_j be the span of T_j. If $d_j > 1$ then we can say that j is periodic, as first returns to j only occur at a number of steps that is a multiple of $d_j > 1$. In general, for $f_{jj} \leq 1$, we have the following definition.

Definition 3.8. *For a state j of a Markov Chain if $d_j := g.c.d.\{n : f_{jj}^{(n)} > 0\} > 1$ then j is called periodic, otherwise j is called aperiodic.* ■

The following lemma is an easy exercise that is left for the reader to prove.

Lemma 3.7.

$$g.c.d.\{n : f_{jj}^{(n)} > 0\} = g.c.d.\{n : p_{jj}^{(n)} > 0\}$$

Proof: Exercise. Hint: note that each element in the set $\{n : p_{jj}^{(n)} > 0\}$ is a sum of one or more elements of the set $\{n : f_{jj}^{(n)} > 0\}$. ■

In the following result we learn that the period of a state is also a class property.

Theorem 3.20. *All states in a communicating class have the same period or are all aperiodic.*

Proof: Consider two states j and k in a communicating class. Then there exist $r > 0$ and $s > 0$ such that $p_{jk}^{(r)} > 0$ and $p_{kj}^{(s)} > 0$. Hence

$$p_{jj}^{(r+s)} > 0 \text{ and } p_{kk}^{(r+s)} > 0$$

Hence, reading $m|n$ as "m divides n," and using Lemma 3.7, we have

$$d_j|(r+s) \text{ and } d_k|(r+s)$$

Further, for all $n \geq 1$,

$$p_{jj}^{(r+n+s)} \geq p_{jk}^{(r)}\, p_{kk}^{(n)}\, p_{kj}^{(s)}$$

Hence, for all $n \geq 1$,

$$p_{kk}^{(n)} > 0 \text{ implies } p_{jj}^{(r+n+s)} > 0$$