Results from Renewal Theory and Its Applications

SooJean Han[1*]

February 20, 2022

Abstract

In this manuscript, we compile extensive notes on results from renewal theory and martingale theory. We also provide numerous applications of renewal theory related to numerous subfields in engineering, including simple queueing processes and the control of discrete-event systems.

1 The Basic Renewal Process

Renewal processes generalize the Poisson process by allowing the sequence of the interarrival times be any i.i.d. random sequence instead of being distributed exponentially. There are numerous applications of renewal processes in engineering, and we demonstrate some in the sections which follow.

In [1], incremental stability of nonlinear stochastic systems perturbed by Poisson shot noise is developed to address the lack of discussion surrounding the analysis of stochastic systems perturbed by non-Gaussian noise. By invoking renewal processes, we can expand the class of systems further by considering systems that are perturbed by impulsive disturbances which are more general than exponentially-distributed interarrival times.

Much of the material in this manuscript has been adapted from [2], Chapter 7 of [3], Chapter 3 of [4], and [5], with more in-depth treatment of the examples, especially towards extensions which are more relevant to applications in engineering.

1.1 Definitions

Definition 1 (Renewal Process)**.** Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. random variables distributed according to a cumulative distribution function (cdf) F. Define an integer-valued stochastic process $\{N(t), t \geq 0\}$ as $N(t) = \max\{n > 0, T_n \leq t\}$, where $T_0 := 0$, $T_n := \sum_{i=1}^{n} X_i$ for $n \geq 1$ are the arrival times of the process. Then $N(t)$ is called a *renewal process*. □

Note that

$$\mathbb{P}(N(t) = n) = \mathbb{P}(N(t) \geq n) - \mathbb{P}(N(t) \geq n+1) = \mathbb{P}(T_n \leq t) - \mathbb{P}(T_{n+1} \leq t) \tag{1}$$

*1 SooJean Han is with the Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125, USA. Email: `soojean@caltech.edu`

Theorem 1 (Wald's Equation)**.** Let $\{X_i\}_{i=1}^{\infty}$ be i.i.d. sequence such that $\mathbb{E}[X_i] := \mathbb{E}[X] < \infty$ for all i, and let N be a stopping time such that $\mathbb{E}[N] < \infty$. Then the following equality holds:

$$\mathbb{E}\left[\sum_{n=1}^{N} X_n\right] = \mathbb{E}[N]\mathbb{E}[X] \tag{2}$$

Proof. Define the indicator random variable

$$I_n := \begin{cases} 1 & \text{if } n \leq N \\ 0 & \text{else} \end{cases}$$

We can rewrite the left side of (2) as follows

$$\mathbb{E}\left[\sum_{n=1}^{N} X_n\right] = \mathbb{E}\left[\sum_{n=1}^{\infty} X_n I_n\right] = \sum_{n=1}^{\infty} \mathbb{E}[X_n I_n] = \sum_{n=1}^{\infty} \mathbb{E}[X_n]\mathbb{E}[I_n] \tag{3}$$

where the last equality follows from independence between X_n and I_n. Continuing:

$$(3) = \mathbb{E}[X]\sum_{n=1}^{\infty} \mathbb{E}[I_n] = \mathbb{E}[X]\mathbb{E}\left[\sum_{n=1}^{\infty} I_n\right] = \mathbb{E}[X]\mathbb{E}[N]$$

which proves the desired relationship (2). ■

2 Application to Pattern Occurrence

2.1 Motivational Example: Tossing a Coin

Suppose that we are tossing a coin which lands a success with probability p, and fails with probability $q = 1 - p$. Suppose we are interested in computing the expected number of tosses T of this coin it takes to observe a specific pattern, e.g. a succession of $k \in \mathbb{N}$ heads in a row.

One simple method is to condition the umber of tosses on the outcome of the first toss. Use the regular definition of expectation:

$$\mathbb{E}[T] = \sum_{i=1}^{k} i\mathbb{P}(T = i)$$

Condition on the time of the first tail and rewrite the above equation:

$$\mathbb{E}[T] = \sum_{i=1}^{k}(1-p)p^{i-1}(i + \mathbb{E}[T]) + kp^k$$

Solving for $\mathbb{E}[T]$, we get

$$\begin{aligned}
\mathbb{E}[T] &= (1-p)\sum_{i=1}^{k} ip^{i-1} + \mathbb{E}[T](1-p)\sum_{i=1}^{k} p^{i-1} + kp^k \\
&= (1-p)\sum_{i=1}^{k} ip^{i-1} + \mathbb{E}[T](1-p)\frac{1-p^k}{1-p} + kp^k
\end{aligned}$$

$$p^k\mathbb{E}[T] = (1-p)\sum_{i=1}^{k} ip^{i-1} + kp^k$$

$$\mathbb{E}[T] = \frac{1-p}{p^k}\sum_{i=1}^{k} ip^{i-1} + k = \frac{1}{p^k}\sum_{i=0}^{k-1} p^i = \frac{(1-p^k)}{p^k(1-p)}$$

We can consider a similar calculation for more general patterns. Suppose now, we wanted to compute the expected number of coin tosses it takes to get the sequence HHT. Let A be the event where we get the desired sequence. Condition on whether we get a heads or a tails on the first toss.

$$\mathbb{E}[A] = \mathbb{E}[A|H]\mathbb{P}(H) + \mathbb{E}[A|T]\mathbb{P}(T) = \mathbb{E}[A|H]p + \mathbb{E}[A|T]q$$

Note that $\mathbb{E}[A|T] = 1 + \mathbb{E}[A]$ since we want a heads in the first toss in order to begin counting for our sequence, but we've already tossed the coin once, so we add a one.

Now $\mathbb{E}[A|H]$ can be further conditioned on whether we get a heads or a tails on the second toss.

$$\mathbb{E}[A|H] = \mathbb{E}[A|HH]\mathbb{P}(H|H) + \mathbb{E}[A|HT]\mathbb{P}(T|H) = \mathbb{E}[A|HH]p + \mathbb{E}[A|HT]q$$

By similar reasoning as above, we have that $\mathbb{E}[A|HT] = 2 + \mathbb{E}[A]$. Also, $\mathbb{E}[A|HH] = 2 + \frac{1}{q}$, since now that we have the first two heads, all we have to wait is for the first tail to land, and the number of tosses required for such is a geometric random variable with parameter q.

Hence

$$\mathbb{E}[A|H] = (2 + \frac{1}{q})p + (2 + \mathbb{E}[A])q = 2p + \frac{p}{q} + 2q + q\mathbb{E}[A] = 2 + \frac{p}{q} + q\mathbb{E}[A]$$

Substituting this into our original expression for $\mathbb{E}[A]$, we get:

$$\mathbb{E}[A] = \frac{2p^2q + p^2 + 2q^2p + q^2}{p^2q}$$

Note that the numerator of this expression can be simplified as follows:

$$2p^2q + p^2 + 2q^2p + q^2 = 2p^2(1-p) + p^2 + 2(1-p)^2p + (1-p)^2 = 1$$

So our final answer is $\frac{1}{p^2q}$. We can visualize this with the figure Figure 1. If we were to obtain the sequence immediately since we start tossing the coin, the endpoint of the trajectory we desire would be as circled.

Another problem of interest for this setup is the probability of observing one pattern before another, e.g., n consecutive successes before m consecutive failures. The extended version of the problem, which is a problem of interest in more general settings of renewal processes, looks into computing the probability of observing a specific pattern first among a predetermined set of patterns we wish to observe. We investigate the more general problem after addressing this simpler coin tossing problem.

We again use conditional expectations to compute the probability, and solve a system of equations. Let A be the event that we get n consecutive successes before m consecutive failures. Let S denote the event that a success occurred in the first toss, and F denote the event that a failure occurred instead. By conditional probabilities, we get:

$$\mathbb{P}(A) = \mathbb{P}(A|S) * \mathbb{P}(S) + \mathbb{P}(A|F) * \mathbb{P}(F) = \mathbb{P}(A|S) * p + \mathbb{P}(A|F) * q$$

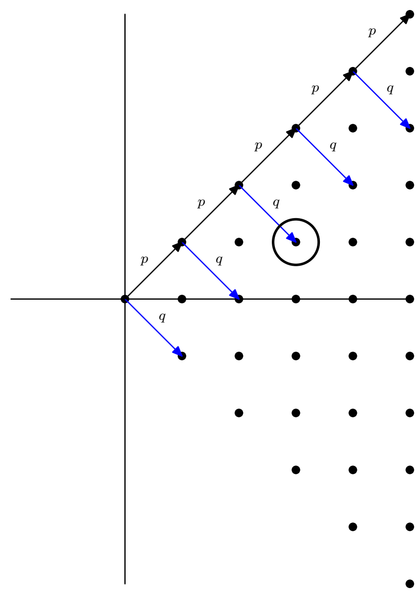

Figure 1: Random walk-like representation of coin toss trials, where the horizontal axis marks the number of tosses made, and the walk travels up one unit if the coin lands heads, down one unit otherwise.

Let us compute an expression for $\mathbb{P}(A|S)$. From the first success, we need $n-1$ more success in order to get n consecutive successes. Then the probability of event A occurring is 1. On the other hand, if at least one failure occurred within the $n-1$ subsequent tosses, then we break the desired sequence, and must restart the whole process, starting with the first failure in there. Hence:

$$\mathbb{P}(A|S) = p^{n-1} * 1 + (1 - p^{n-1}) * \mathbb{P}(A|F) = p^{n-1} + (1 - p^{n-1})\mathbb{P}(A|F)$$

We can similarly compute an expression for $\mathbb{P}(A|F)$. From the first failure, we need $m-1$ more failures in order to get m consecutive failures. Then the probability of event A occurring is 0. On the other hand, if at least one success occurred within the $m-1$ subsequent tosses, then we break the string of failures, and must restart the whole process, starting with the first success in there. Hence:

$$\mathbb{P}(A|F) = q^{m-1} * 0 + (1 - q^{m-1}) * \mathbb{P}(A|S) = (1 - q^{m-1})\mathbb{P}(A|S)$$

We have two equations and two unknowns, $\mathbb{P}(A|S)$ and $\mathbb{P}(A|F)$. Solving for them yields:

$$\mathbb{P}(A|S) = \frac{p^{n-1}}{p^{n-1} + q^{m-1} - p^{n-1}q^{m-1}}, \quad \mathbb{P}(A|F) = \frac{(1 - q^{m-1})p^{n-1}}{p^{n-1} + q^{m-1} - p^{n-1}q^{m-1}}$$

Substituting into the expression for our original desired probability, we get:

$$\mathbb{P}(A) = \mathbb{P}(A|S) * p + \mathbb{P}(A|F) * q = \frac{q^{m-1}(1 - p^n)}{p^{n-1} + q^{m-1} - p^{n-1}q^{m-1}}$$

2.2 Independent, Identically-Distributed Case

Motivated by the above example, we can consider the general problem of **pattern occurrence**, set up in the following way. Let $X_1, X_2, \cdots$ denote a sequence of i.i.d. random variables taking on values from a discrete, finite set $\mathcal{X}$. Denote $p_k := \mathbb{P}(X_i = k)$ for all i, and let T denote the next time the pattern $(x_1, \cdots, x_m)$ occurs after its first occurrence. In this problem, we are interested in the expected value of T. Within the specific context of the pattern occurrence problem, we refer to a "renewal" as the event when the pattern string $(x_1, \cdots, x_m)$ repeats itself. Hence, the interarrival times of such a renewal process are distributed in the same manner as T.

There are two cases that need to be considered, depending on the amount of overlap there is in the pattern we are interested in.

Definition 2 (Pattern Overlap)**.** For a renewal process $X_1, X_2, \cdots$ taking values $x_1, x_2, \cdots$ from a certain probability distribution, we say that a pattern $(x_{(1)}, x_{(2)}, \cdots, x_{(m)})$ has an overlap of size $k < m$ if

$$k \triangleq \max\{\ell < m \mid (x_1, \cdots, x_\ell) = (x_{m-\ell+1}, \cdots, x_m)\}$$

That is, k is the largest value such that the first k elements are identical to the last k elements. □

- ***Case 1: there is no pattern overlap.*** In this case, it is intuitive that the occurrence time T can be treated as a geometric random variable. Hence,

$$\mathbb{E}[T] = \frac{1}{\prod_{i=1}^{m} p_{x_i}} \tag{4}$$

since $\prod_{i=1}^{m} p_{x_i}$ is the probability of observing exactly the sequence $(x_1, \cdots, x_m)$. For a more comprehensive derivation of this formula, note that $T > n + m$ iff the pattern does not occur for the first n values, and the next m values are $(x_1, \cdots, x_m)$. That is:

$$\{T > n + m\} \iff \{T > n \text{ and } (X_{n+1}, \cdots, X_{n+m}) = (x_1, \cdots, x_m)\}$$

In terms of probabilities:

$$\mathbb{P}(T > n + m) = \mathbb{P}(T > n)\mathbb{P}((X_{n+1}, \cdots, X_{n+m}) = (x_1, \cdots, x_m)) = \mathbb{P}(T > n) \prod_{i=1}^{m} p_i$$

Note that by the definition of expected value:

$$1 = \sum_{n=0}^{\infty} \mathbb{P}(T > n + m) = \prod_{i=1}^{m} p_i \sum_{n=0}^{\infty} \mathbb{P}(T > n) = \prod_{i=1}^{m} p_i \mathbb{E}[T]$$

Indeed, dividing through by $\prod_{i=1}^{m} p_i$ yields (4).

- ***Case 2: there is a pattern overlap of size k, which does not have an overlap itself.*** Suppose the pattern $(x_1, \cdots, x_m)$ has an overlapping subsequence $(x_1, \cdots, x_k)$ which itself does not have an overlap. Note that if $(x_1, \cdots, x_k)$ has an overlap, then we can repeat Case 2 by induction.

Define T_y to be the next time after the first occurrence of the augmented pattern $(x_1, \cdots, x_m, y)$ it takes to observe the augmented pattern again. Then it is easy to see that $T_y > T$, that is:

$$T_y = T + \Delta T \tag{5}$$

where ΔT is the time after the next occurrence of the original pattern $(x_1, \cdots, x_m)$ it takes to observe the next occurrence of the augmented pattern. Note that the computation of $\mathbb{E}[T_y]$ is a simple instance of Case 1, since the addition of y to the original sequence removes the overlap. Hence:

$$\mathbb{E}[T_y] = \frac{1}{p_y \prod_{i=1}^{m} p_i}$$

We can compute $\mathbb{E}[\Delta T]$ by conditioning on the next value X, assuming the sequence $(x_1, \cdots, x_m)$ has already been observed.

$$\mathbb{E}[\Delta T | X = z] = \begin{cases} 1 + \mathbb{E}[T_y | x_1, \cdots, x_{k+1}] & \text{if } z = x_{k+1} \\ 1 + \mathbb{E}[T_y | x_1] & \text{if } z = x_1 \\ 1 & \text{if } z = y \\ 1 + \mathbb{E}[T_y] & \text{if } z \notin \{x_1, x_{k+1}, y\} \end{cases}$$

from which we can construct equations and compute $\mathbb{E}[T]$. Namely:

$$\mathbb{E}[\Delta T] = 1 + p_{k+1}\mathbb{E}[T_y | x_1, \cdots, x_{k+1}] + p_1 \mathbb{E}[T_y | x_1] + (1 - p_{k+1} - p_1 - p_y)\mathbb{E}[T_y]$$

where p_y denotes the probability of observing y in a trial. Note that we can directly write, by definition,

$$\begin{aligned} \mathbb{E}[T_y] &= \mathbb{E}[T(x_1, \cdots, x_{k+1})] + \mathbb{E}[T_y | x_1, \cdots, x_{k+1}] \\ \mathbb{E}[T_y] &= \mathbb{E}[T(x_1)] + \mathbb{E}[T_y | x_1] \end{aligned}$$

where $T(x_1, \cdots, x_j)$ represents the time it takes to observe the next occurrence of the pattern $(x_1, \cdots, x_j)$ after its first occurrence. Since neither of the sequences $(x_1, \cdots, x_{k+1})$ nor (x_1) have overlaps, we can use the first case to compute:

$$\mathbb{E}[T(x_1, \cdots, x_{k+1})] = \frac{1}{\prod_{i=1}^{k+1} p_i}, \quad \mathbb{E}[T(x_1)] = \frac{1}{p_1}$$

In combination with (5), we get:

$$\begin{aligned} \mathbb{E}[T_y] = \mathbb{E}[T] + 1 + p_{k+1}\left(\mathbb{E}[T_y] - \mathbb{E}[T(x_1, \cdots, x_{k+1})]\right) + p_1\left(\mathbb{E}[T_y] - \mathbb{E}[T(x_1)]\right) \\ + (1 - p_{k+1} - p_1 - p_y)\mathbb{E}[T_y] \\ \implies p_y \mathbb{E}[T_y] = \mathbb{E}[T] + \cancel{1} - p_{k+1}\mathbb{E}[T(x_1, \cdots, x_{k+1})] - \cancel{p_1 \mathbb{E}[T(x_1)]} \\ \implies \mathbb{E}[T] = \cancel{p_y} \cdot \frac{1}{\cancel{p_y} \prod_{i=1}^{m} p_i} + p_{k+1} \cdot \frac{1}{\prod_{i=1}^{k+1} p_i} \end{aligned}$$

In conclusion, where a specific pattern $(x_1, \cdots, x_m)$ has an overlap of size k

$$\mathbb{E}[T] = \frac{1}{\prod_{i=1}^{m} p_i} + \frac{1}{\prod_{i=1}^{k} p_i}$$

Example 1 (String of Same Values)**.** Suppose we are interested in computing T for a sequence of m consecutive values, e.g. $(5, 5, \cdots, 5)$, for when $\mathcal{X} := \{0, 1, \cdots, 9\}$. Then following the analysis of Case 2 above, we can define T_y for the augmented sequence $(5, 5, \cdots, 5, y)$, where $y \neq 5$. Then we can compute $\mathbb{E}[T]$ as follows:

$$\mathbb{E}[T] = \mathbb{E}[T_y] - \mathbb{E}[\Delta T]$$

where

$$\mathbb{E}[T_y] = \frac{1}{p_5^m p_y}, \quad \mathbb{E}[\Delta T | X = z] = \begin{cases} 1 + \mathbb{E}[\Delta T] & \text{if } z = 5 \\ 1 + \mathbb{E}[T_y] & \text{if } z \notin \{5, y\} \\ 1 & \text{if } z = y \end{cases}$$

so

$$\begin{aligned}
&\mathbb{E}[\Delta T] = 1 + p_5\mathbb{E}[\Delta T] + (1 - p_5 - p_y)\mathbb{E}[T_y] \\
\Longrightarrow &\mathbb{E}[\Delta T] = \frac{1}{1-p_5} + \frac{1-p_5-p_y}{1-p_5} \cdot \mathbb{E}[T_y] \\
\Longrightarrow &\mathbb{E}[T] = \mathbb{E}[T_y] - \mathbb{E}[\Delta T] = -\frac{1}{1-p_5} + \frac{p_y}{1-p_5} \cdot \mathbb{E}[T_y] = \frac{\frac{1}{p_5^m} - 1}{1-p_5}
\end{aligned}$$

which, one can verify, is the formula obtained when substituting it into Case 2. □

Example 2 (String of Same Values: Coin Tosses)**.** Now we will consider a specific instance of Example 1 for when $\mathcal{X} = \{H, T\}$ and $p := \mathbb{P}(X_n = H), q := 1 - p = \mathbb{P}(X_n = T)$. It is well-known that the formula of the time until one observes m consecutive H's is given by

$$\mathbb{E}[T] = \sum_{k=1}^{m} \frac{1}{p^k}$$

We will derive this formula using the procedure used to prove Case 2 above. First, denote T_T to be the time until we observe the sequence $(H, \cdots, H, T)$. Then $\mathbb{E}[T_y] = 1/(p^m q)$

$$\mathbb{E}[\Delta T] = 1 + p\mathbb{E}[\Delta T] + q \cdot 0 \implies \mathbb{E}[\Delta T] = \frac{1}{q}$$

and

$$\mathbb{E}[T] = \frac{1}{p^m q} - \frac{1}{q} = \frac{1-p^m}{q} = \frac{(1-p)(1+p+\cdots+p^{m-1})}{1-p} = \sum_{k=1}^{m} \frac{1}{p^k}$$

which indeed verifies the formula. □

Example 3 (Coin Tosses: Specific Strings)**.** Suppose we are given a coin with probability of heads p, and we are interested in two specific sequence patterns, $A := HHTTHH$ and $B := HTHTT$. Let $T(A)$ and $T(B)$ denote the number of flips until we observe A and B, respectively. Then we get:

$$\mathbb{E}[T(A)] = \frac{1}{p^4 q^2} + \mathbb{E}[T(HH)] = \frac{1}{p^4 q^2} + \frac{1}{p^2} + \frac{1}{p}$$

from induction and the Case 2 formula above, while:

$$\mathbb{E}[T(B)] = \frac{1}{p^2 q^3}$$

from the Case 1 formula above.

We are interested in the probability that A will occur before B as well as the expected time it takes to observe either one of the patterns A or B. We further denote $T(B|A)$ to be the additional number of flips needed in order to observe B in the sequence, given we've already observed A. Likewise, define $T(A|B)$.

$$\begin{aligned}
\mathbb{E}[T(B|A)] &= \mathbb{E}[T(B)] - \mathbb{E}[T(H)] = \frac{1}{p^2 q^3} - \frac{1}{p} \\
\mathbb{E}[T(A|B)] &= \mathbb{E}[T(A)] = \frac{1}{p^4 q^2} + \frac{1}{p^2} + \frac{1}{p}
\end{aligned}$$

where the first equation follows since the H at the end of sequence A can be reused, while the second equation follows because there is no subsequence at the end of sequence B which can be reused for A.

Now denote P_A to be the probability that A will occur before B. Further denote $T_{\min} := \min\{T(A), T(B)\}$. We can derive a system of equations with these variables as follows:

$$\mathbb{E}[T(A)] = \mathbb{E}[T_{\min}] + \mathbb{E}[T(A) - T_{\min}] = \mathbb{E}[T_{\min}] + \mathbb{E}[T(A|B)](1 - P_A) + 0 \cdot P_A$$
$$\mathbb{E}[T(B)] = \mathbb{E}[T_{\min}] + \mathbb{E}[T(B) - T_{\min}] = \mathbb{E}[T_{\min}] + \mathbb{E}[T(B|A)]P_A + 0 \cdot (1 - P_A)$$

and solving for the variables yields

$$P_A = \frac{\mathbb{E}[T(B)] + \mathbb{E}[T(A|B)] - \mathbb{E}[T(A)]}{\mathbb{E}[T(B|A)] + \mathbb{E}[T(A|B)]}$$
$$\mathbb{E}[T_{\min}] = \mathbb{E}[T(B)] - \mathbb{E}[T(B|A)]P_A$$

Substituting in the values corresponding to each specific pattern yields

$$P_A = \frac{p^2}{p^2q^3 + q + p^2} = \frac{8}{25}$$
$$\mathbb{E}[T_{\min}] = \frac{1}{p^2q^3} - \left(\frac{1}{p^2q^3} - \frac{1}{p}\right) P_A = \frac{112}{5}$$

where the numerical values are evaluated for $p = \frac{1}{2}$.

Now, consider Bernoulli trials with H probability p and T probability $q := 1 - p$. We can use the above generic formula to compute the probability P_A that a sequence of m heads, denoted by sequence A, is observed before a sequence of n tails, denoted by sequence B. We first have

$$T_A = \sum_{i=1}^{m} \frac{1}{p^i}, \quad T_B = \sum_{j=1}^{n} \frac{1}{q^j}$$

Note that since neither of the sequences overlap with each other, $T_{A|B} = T_A$ and $T_{B|A} = T_B$. Hence:

$$T_A = T_{\min} + (T_A - T_{\min}) = T_{\min} + (1 - P_A)T_{A|B}$$
$$T_B = T_{\min} + (T_B - T_{\min}) = T_{\min} + P_A T_{B|A}$$

and thus,

$$P_A = \frac{\sum_{j=1}^{n} \frac{1}{q^j}}{\sum_{i=1}^{m} \frac{1}{p^i} + \sum_{j=1}^{n} \frac{1}{q^j}} = \frac{\frac{1 - \frac{1}{q^n}}{q\left(1 - \frac{1}{q}\right)}}{\frac{1 - \frac{1}{p^m}}{p\left(1 - \frac{1}{p}\right)} + \frac{1 - \frac{1}{q^n}}{q\left(1 - \frac{1}{q}\right)}} = \frac{(1 - q^n)p^{m-1}}{q^{n-1} + p^{m-1} - p^{m-1}q^{n-1}}$$

$\square$

Example 4 (Second Largest vs Second Smallest)**.** Consider the sequence of continuous i.i.d. random variables $X_1, X_2, \cdots \sim F$ with pdf $f = F'$.

Let $X_{(i),n}$ denote the ith-smallest value in the sequence observed up until time n, $X_1, \cdots, X_n$. Further denote

$$N := \min\{n \geq 2 | X_n = X_{(n-1),n}\}, \quad M := \min\{n \geq 2 | X_n = X_{(2),n}\}$$

That is, N is the first time where X_n is the second-largest value in the sequence $(X_1, \cdots, X_n)$ and M is the first time where X_n is the second-smallest. We are interested in determining which value, X_N or X_M tends to be the larger.

Define E_n to be the event that X_n is not the second-largest of $(X_1, \cdots, X_n)$. Then

$$\mathbb{P}(N = n) = \mathbb{P}(E_2 \cap E_3 \cap \cdots \cap E_{n-1} \cap \overline{E}_n)$$

Since the X_i are i.i.d., knowing the rank ordering of $X_1, \cdots, X_n$ yields no information about their actual values. That means that the events E_i are independent of each other over all i, and X_i is equally likely to be the largest, or the second-largest, or $\cdots$, or the smallest. Thus,

$$\mathbb{P}(E_i) = \frac{i-1}{i} \implies \mathbb{P}(N = n) = \frac{1}{2} \cdot \frac{2}{3} \cdot \cdots \cdot \frac{n-2}{n-1} \cdot \frac{1}{n} = \frac{1}{n(n-1)}$$

We can then condition the value of X_N on $N = n$:

$$\mathbb{P}(X_N = x) = \sum_{n=2}^{\infty} \mathbb{P}(X_N = x | N = n)\mathbb{P}(N = n) = \sum_{n=2}^{\infty} \frac{1}{n(n-1)} \cdot \mathbb{P}(X_n = x) \tag{6}$$

Note that

$$\mathbb{P}(X_{(i),n} \leq x) = \sum_{k=i}^{n} \binom{n}{k} F(x)^k (1 - F(x))^{n-k}$$

since by definition of $X_{(i),n}$, $X_{(i),n}$ will be less than or equal to x iff at least i of the n values are $\leq x$. By taking derivatives:

$$\begin{aligned}
\mathbb{P}(X_{(i),n} = x) &= f(x) \sum_{k=i}^{n} \binom{n}{k} k F(x)^{k-1} (1 - F(x))^{n-k} - f(x) \sum_{k=i}^{n} \binom{n}{k} F(x)^k (n-k)(1 - F(x))^{n-k-1} \\
&= f(x) \left[\sum_{k=i}^{n} \frac{n!}{(n-k)!(k-1)!} F(x)^{k-1}(1 - F(x))^{n-k} - \sum_{k=i}^{n} \frac{n!}{(n-k-1)!k!} F(x)^k (1 - F(x))^{n-k-1} \right] \\
&= f(x) \left[\sum_{k=i}^{n} \frac{n!}{(n-k)!(k-1)!} F(x)^{k-1}(1 - F(x))^{n-k} - \sum_{j=i+1}^{n} \frac{n!}{(n-j)!(j-1)!} F(x)^k (1 - F(x))^{n-j} \right] \\
&= f(x) \cdot \frac{n!}{(n-i)!(i-1)!} \cdot F(x)^{i-1}(1 - F(x))^{n-i}
\end{aligned}$$

With $i = n - 1$ and substituting the above expression into (6) yields

$$\begin{aligned}
\mathbb{P}(X_N = x) &= \sum_{n=2}^{\infty} \frac{1}{n(n-1)} \cdot \frac{n!}{(n-2)!} f(x) F(x)^{n-2} (1 - F(x)) \\
&= f(x)(1 - F(x)) \sum_{n=2}^{\infty} F(x)^{n-2} \\
&= f(x)(1 - F(x)) \cdot \frac{1}{1 - F(x)} \quad \text{by geometric series} \\
&= f(x)
\end{aligned}$$

This implies that $X_N \sim X_1$. To obtain the distribution of X_M, just invert the sign and repeat the previous calculation since the second-smallest of $X_1, \cdots, X_n$ and be interpreted as the second-largest of $-X_1, \cdots, -X_n$. Thus, $-X_M \sim -X_1$ and so $X_M \sim X_1$.

Therefore, whether $X_M \leq X_N$ or $X_M > X_N$ is determined completely randomly according to the distribution F. One might note the analogy to the well-known basic probability of computing $\mathbb{P}(U_1 \leq U_2)$ for two i.i.d. random variables U_1 and U_2. $\square$

2.3 Markov Chain Case

We are especially interested in patterns which arise from sequences of vector values, $X_t = (X_{1,t}, \cdots, X_{n,t})$. We further assume that for each timestep $t \in \mathbb{N}$ the components $\{X_{i,t}\}_{i=1}^n$ are independent of each other, and that each sequence of component values over time $\{X_{i,t}\}_{t=1}^\infty$ comes either from an i.i.d. sequence or a memoryless conditional distribution such as a Markov chain. Then note that the vector sequence $\{X_t\}_{t=1}^\infty$ can be represented as a Markov chain itself. In this case, the expected time between consecutive occurrences of a pattern of interest hinges on the following lemma.

Lemma 1 (Mean Return Time to a State)**.** Let $\{X_n\}$ be an irreducible Markov Chain with state-space $\mathcal{X}$, transition probability matrix $P \triangleq \{P_{ij}\}$, and stationary distribution $\{\pi_k\}$. Denote τ_i to be the minimum return time to state i given the chain is started at $X_0 = i$. That is,

$$\tau_i \triangleq \min\{n \geq 1 \mid X_n = X_0 = i\} \tag{7}$$

If X_n is a positive-recurrent Markov chain, i.e. $\mathbb{E}[\tau_i] < \infty$ for all $i \in \mathcal{X}$, then a unique stationary distribution $\{\pi_i\}$ exists and

$$\mathbb{E}[\tau_i] = \frac{1}{\pi_i} \quad \forall i \in \mathcal{X} \tag{8}$$

The following theorem presents the case where $n = 2$, with the first component coming from an i.i.d. sequence and the second coming from a Markov chain.

Theorem 2 (Pattern Occurrence for Random Vector Sequences)**.** Let $\{X_t\}_{t=1}^\infty \subset \mathbb{R}^2$ be such that $X_t := (Y_t, Z_t)$ where $\{Y_t\}_{t=1}^\infty$ is an i.i.d. sequence taking values from a discrete, finite set $\mathcal{Y} \subset \mathbb{R}$, and $\{Z_t\}_{t=1}^\infty \subset \mathbb{R}$ abides by an irreducible Markov chain with discrete state-space $\mathcal{Z} \subset \mathbb{R}$. Denote $q(y) := \mathbb{P}(Y_t = y)$ for any $y \in \mathcal{Y}$ and $t \in \mathbb{Z}_{\geq 0}$, and denote the transition probability matrix for $\{Z_t\}$ to be $P \triangleq \{P_{z,z'}\}$ with stationary distribution $\{\pi_z\}$ for any $z \in \mathcal{Z}$. Furthermore, for any two states $z, z' \in \mathcal{Z}$, denote $\mu(z, z')$ to be the mean time it takes to reach state z' starting from state z. Choose a fixed $m \in \mathbb{N}$. Let T_0 denote the number of transitions until the first occurrence of the pattern $(x_1, \cdots, x_m)$ where $x_i := (y_i, z_i)$ for each of $i = 1, \cdots, m$, given the first state of the Markov chain is Z_0. Let T denote the number of transitions between the first and second occurrences of $(x_1, \cdots, x_m)$. Then:

- When the pattern $(x_1, \cdots, x_m)$ does not contain any overlaps, the expected time between the first and second occurrences is given by

$$\mathbb{E}[T] = \frac{1}{q_{y_1} \pi_{z_1} \prod_{i=1}^{m-1} q_{y_{i+1}} P_{z_i, z_{i+1}}} \tag{9}$$

  and the expected time until the first occurrence is

$$\mathbb{E}[T_0 | Z_0 = z_0] = \mu(z_0, z_1) - \mu(z_m, z_1) + \mathbb{E}[T] \tag{10}$$

  where $z_0 \in \mathcal{Z}$ and $\mathbb{E}[T]$ is the expression from (9).

- When the pattern $(x_1, \cdots, x_m)$ has an overlap $(x_1, \cdots, x_k)$ of size k, and $(x_1, \cdots, x_k)$ itself does not contain any overlaps, the expected time between the first and second occurrences is given as in (9) and the expected time until the first occurrence is

$$\mathbb{E}[T_0 | Z_0 = z_0] = \mu(z_0, z_1) - \mu(z_k, z_1) + \frac{1}{q_{y_1} \pi_{z_1} \prod_{i=1}^{k-1} q_{y_{i+1}} P_{z_i, z_{i+1}}} + \mathbb{E}[T] \tag{11}$$

  where $z_0 \in \mathcal{Z}$.

Before proving Theorem 2, we consider the following simple example, which illustrates a simple relationship between the stationary probabilities of the Markov chain component and the Markov chain constructed from the full vector state, including the i.i.d. component.

Example 5. Using the notation of Theorem 2, let $\mathcal{Y} := \{1, 2, 3\}$ and $\mathcal{Z} := \{A, B, C\}$ with, respectively, probability distribution $\{q_1, q_2, q_3\}$ and transition probability matrix

$$P := \begin{bmatrix} P_{AA} & P_{AB} & 0 \\ 0 & P_{BB} & P_{BC} \\ P_{CA} & 0 & P_{CC} \end{bmatrix}$$

For the Markov chain $\{Z_t\}$, the stationary probability equation $\pi = \pi P$ expands as follows:

$$\begin{aligned} \pi_A &= \pi_A P_{AA} + \pi_C P_{CA} \\ \pi_B &= \pi_B P_{BB} + \pi_A P_{AB} \\ \pi_C &= \pi_C P_{CC} + \pi_B P_{BC} \end{aligned}$$

The transition probability matrix for the Markov chain corresponding to the original sequence $\{X_t\}$ structurally resembles the transition probability matrix for $\{Z_t\}$. Note that all states $(1, z), (2, z), (3, z)$ are able to transition to each other with positive probability for any $z \in \{A, B, C\}$ due to the i.i.d. nature of the first component. For $\{X_t\}$, the stationary probability equation writes as

$$\begin{aligned} \pi(y, A) &= \sum_{j=1}^{3} \pi(j, A) q_y P_{AA} + \sum_{j=1}^{3} \pi(j, C) q_y P_{CA} \\ \pi(y, B) &= \sum_{j=1}^{3} \pi(j, B) q_y P_{BB} + \sum_{j=1}^{3} \pi(j, A) q_y P_{AB} \\ \pi(y, C) &= \sum_{j=1}^{3} \pi(j, C) q_y P_{CC} + \sum_{j=1}^{3} \pi(j, B) q_y P_{BC} \end{aligned}$$

where $y \in \{1, 2, 3\}$, yielding a total of nine equations. However, each group of three equations can be summed over y, and with the fact that $\sum_{y=1}^{3} q_y = 1$, designating a new variable $\psi_z := \sum_{y=1}^{3} \pi(y, z)$ for each of $z \in \{A, B, C\}$ yields exactly the original stationary equations for the Markov chain $\{Z_t\}$, with ψ_z in place of π_z. Hence, we obtain the relationship $\pi(y, z) = q_y \pi_z$ for all y and z. It is easy to extend this result to general n, where any number $m \leq n$ of the components come from i.i.d. sequences. □

We further define a few notations. For any $x'_1, \cdots, x'_\ell \in \mathcal{X}$ and $\ell \in \mathbb{N}$, let $\Delta T(x'_1, \cdots, x'_\ell)$ be the additional number of transitions needed to observe $(x_1, \cdots, x_m)$ assuming $(x'_1, \cdots, x'_\ell)$ has already been observed. Further define associated Markov chains $\{\zeta_t\}$ and $\{\chi_t\}$ with states given by the past m sequences of the Markov chain $\{Z_t\}$ and the original sequence $\{X_t\}$, respectively. That is, $\zeta_t \triangleq (Z_{t-m+1}, \cdots, Z_t)$ for $t \in \mathbb{Z}_{\geq 0}$ and $\chi_t := (X_{t-m+1}, \cdots, X_t)$. It is easy to verify that $\{\zeta_t\}$ has the memorylessness property, and because the appended component Y_t comes from an i.i.d. sequence, $\{\chi_t\}$ is also memoryless. Denote the stationary probabilities of ζ_t to be $\rho(z'_1, \cdots, z'_m)$ for any sequence of states $(z'_1, \cdots, z'_m) \in \mathcal{Z}^m$. Similarly, the stationary probabilities of χ_t to be $\phi(x'_1, \cdots, x'_m)$ for any sequence of states $(x'_1, \cdots, x'_m) \in \mathcal{Y}^m \times \mathcal{Z}^m$, where $x_i := (y_i, z_i)$. Then we have:

$$\rho(z'_1, \cdots, z'_m) := \frac{1}{\pi_{z'_1} \prod_{i=1}^{m-1} P_{z'_i, z'_{i+1}}} \tag{12a}$$

$$\phi(x'_1, \cdots, x'_m) := \frac{1}{\prod_{i=1}^{m} y_i} \cdot \rho(z'_1, \cdots, z'_m) \tag{12b}$$

Proof of (9) and (10). Let $\tau(z_1, \cdots, z_m)$ denote the number of transitions of the Markov chain ζ_t between consecutive visits to the state $(z_1, \cdots, z_m), z_k \in \mathcal{Z}$ for all $k = 1, \cdots, m$. By the expression for the mean return time for Markov chains, we have:

$$\mathbb{E}[\tau(z_1, \cdots, z_m)] = \frac{1}{\rho(z_1, \cdots, z_m)} \tag{13}$$

Moreover, the probability for the corresponding sequence of Y_t is given by

$$\mathbb{P}((Y_{t-m+1}, \cdots, Y_t) = (y_1, \cdots, y_m)) = \prod_{i=1}^{n} q_{y_i} \tag{14}$$

By virtue of Example 5, the expected time is simply the inverse of the product of the two probabilities, yielding (9).

To compute (10), note that

$$\mathbb{E}[T] = \mathbb{E}[\Delta T(x_m)] = \mu(z_m, z_1) + \mathbb{E}[\Delta T(x_1)] \tag{15}$$

That is, the expected number of transitions between two consecutive occurrences of $(z_1, \cdots, z_m)$ is equivalent to the expected additional amount of time it takes to observe $(z_1, \cdots, z_m)$ given z_m (from x_m) has already been observed. This is then equivalent to the expected number of transitions it takes to observe z_1 (from x_1) given z_m, plus the expected number of transitions it takes to observe $(z_1, \cdots, z_m)$ given z_1. Because the component Y_t appended to Z_t is i.i.d., there is always nonzero probability of observing y_{t+1} given y_t. Thus, we obtain the expression (10). ■

Proof of (11). Suppose the pattern $(x_1, \cdots, x_m)$ has an overlapping subsequence $(x_1, \cdots, x_k)$, $k \leq m$, which itself does not have an overlap. Hence, by (13):

$$\mathbb{E}[T] = \mathbb{E}[\Delta T(x_{m-k+1}, \cdots, x_m)] = \mathbb{E}[\Delta T(x_1, \cdots, x_k)] = \mathbb{E}[\tau(x_1, \cdots, x_m)] = \frac{1}{\phi(x_1, \cdots, x_m)} \tag{16}$$

Note that due to the Markov chain structure that becomes imposed on the original sequence $\{X_t\}$ via the second component $\{Z_t\}$ (see Example 5), the expression of $\mathbb{E}[T]$ is the same as the case with no overlaps.

Denote $T_0(x_1, \cdots, x_k)$ to be the time it takes to observe the sequence $(x_1, \cdots, x_k)$ for the first time. Then note that

$$\mathbb{E}[T_0|Z_0 = z_0] = \mathbb{E}[T_0(x_1, \cdots, x_k)|Z_0 = z_0] + \mathbb{E}[\Delta T(x_1, \cdots, x_k)] \tag{17}$$

That is, the expected number of transitions until the first occurrence of $(x_1, \cdots, x_m)$ is equivalent to the expected number of transitions it takes to observe $(x_1, \cdots, x_k)$, plus the expected additional time it takes to observe $(x_1, \cdots, x_m)$ after $(x_1, \cdots, x_k)$ has been observed. Since $(x_1, \cdots, x_k)$ has no overlaps, we can use the formula (10) to obtain

$$\mathbb{E}[T_0(z_1, \cdots, z_k)|Z_0 = z_0] = \mu(z_0, z_1) - \mu(z_k, z_1) + \frac{1}{q_{y_1} \pi_{z_1} \prod_{i=1}^{k-1} q_{y_{i+1}} P_{z_i, z_{i+1}}} \tag{18}$$

Substituting (18) and (16) into (17) yields the desired result (11). ■

2.4 Interpretation via Martingale Theory

We discuss the pattern-occurrence problem from an alternative perspective using martingales, which we adapt from [6]. Recall that a sequence of random variables $\{X_n\}_{n=1}^{\infty}$ on probability space $(\Omega, \mathcal{F}, P)$ with filtration $\{\mathcal{F}\}_{n=1}^{\infty}$ defined by $\mathcal{F}_n := \sigma(X_0, X_1, \cdots, X_n)$ is a *martingale* if for all $n \in \mathbb{N}$, $\mathbb{E}[|X_n|] < \infty$ and $\mathbb{E}[X_n|\mathcal{F}_m] = X_m$ for all $m, n \in \mathbb{N}$ such that $m \leq n$. Associated with a martingale, we often have a *stopping time* $\tau \in \mathbb{R}^+$, which is a random variable for the time at which to stop the process $\{X_n\}_{n=1}^{\infty}$ based on its current and past values. For instance, a random walk $\{S_n\}_{n=1}^{\infty}$, where $S_n = \sum_{n=0}^{\infty} X_n$ where X_n are i.i.d. with $\mathbb{E}[X_n] = 0$ is a martingale, and $\tau := \min\{n \in \mathbb{N}^+ | S_n = a\}$ for some $a \in \mathbb{R}$ is an example of a stopping time.

Recall that the pattern-occurrence problem was posed as follows: given the sequence of i.i.d. random variables $\{X_n\}_{n=1}^{\infty}$, suppose we are interested in the expected time it takes to observe the pattern $B := (b_1, \cdots, b_k)$ for the first time. That is, if $\tau_B \in \mathbb{R}^+$ be the random variable defined as

$$\tau_B := \min\{n \in \mathbb{N} \mid (X_n, \cdots, X_{n+k-1}) = B\} \tag{19}$$

then we are interested in computing $\mathbb{E}[\tau_B]$.

We first make the following notations. For two patterns $A := (a_1, \cdots, a_m)$ and $B := (b_1, \cdots, b_k)$:

$$\delta_{i,j}(A, B) := \begin{cases} \frac{1}{\mathbb{P}(X=b_j)} & \text{if } a_i = b_j \\ 0 & \text{else} \end{cases} \tag{20}$$

for all $i \in \{1, \cdots, m\}$, $j \in \{1, \cdots, k\}$.

The martingale interpretation for the pattern-occurrence problem goes as follows. Suppose, at each time instant $n \in \mathbb{N}$, a gambler arrives and bets money on the event $\{(X_n, X_{n+1}, \cdots, X_{n+k-1}) = B\}$, using betting strategy $\{Y_{n,j}\}_{j=0}^{k-1}$, starting from $Y_{n,0} = 1$. Each gambler's strategy is defined such that the game has fair odds, i.e.

$$Y_{n,j+1} := \begin{cases} \left(\frac{1}{\mathbb{P}(X=b_{j+1})}\right) Y_{n,j} - Y_{n,j} & \text{if } X_{n+k-j} = b_{j+1} \\ -Y_{n,j} & \text{if } X_{n+k-j} \neq b_{j+1} \end{cases} \tag{21}$$

for each $j \in \{0, \cdots, k-1\}$. When a gambler first loses everything he has, he leaves the game and never returns. Hence, gambler n's total earnings by some time $\ell \in \mathbb{N}$ is given by $R_{n,\ell} = \sum_{j=0}^{k-1} Y_{n,j} \mathbb{1}\{n + j \leq \ell\}$, given by

$$R_{n,\ell} = \begin{cases} \left(\prod_{\substack{j \in \{0, \cdots, k-1\} \\ n+j \leq \ell}} \mathbb{P}(X = b_j) \right)^{-1} - 1 & \text{if } \forall j \in \{0, \cdots, k-1\} \text{ s.t. } n+j \leq \ell,\ \delta_{n+j,j+1}(X^n, B) > 0 \\ -1 & \text{if } \exists j \in \{0, \cdots, m-1\} \text{ s.t. } n+j \leq \ell,\ \delta_{n+j,j+1}(X^n, B) = 0 \end{cases} \tag{22}$$

where we denote $X^n := (X_1, \cdots, X_n)$. We set up the betting strategy of each gambler in this way so that the accumulated net reward over all gamblers $\{\sum_{n=1}^{\ell} R_{n,\ell}\}_{\ell=1}^{\infty}$ of the game is a martingale. Then by the optional stopping theorem, we can have that $\mathbb{E}[\sum_{n=1}^{\ell \wedge \tau_B} R_{n,\ell \wedge \tau_B}] = 0$. This makes it easy to compute $\mathbb{E}[\tau_B]$, as we demonstrate concretely using the example below.

Example 6 (12-Sided Die)**.** Suppose we roll a 12-sided die with the sequence of possible outcomes $X_1, X_2, \cdots$ distributed as

$$\mathbb{P}(X_n = k) = \begin{cases} p_1 := \frac{1}{4} & \text{if } k = 1 \\ p_2 := \frac{1}{3} & \text{if } k = 2 \\ p_3 := \frac{1}{4} & \text{if } k = 3 \\ p_4 := \frac{1}{6} & \text{if } k = 4 \end{cases}$$

for all $n \in \mathbb{N}$. Suppose we are interested in observing the pattern $B := (b_1, \cdots, b_k) := (2, 3, 4, 2, 3)$ of length $k = 5$. Using the formula from Section 2.2 and the fact that B has an overlap of size 2, we get

$$\mathbb{E}[\tau_B] = \left(\frac{1}{3 \cdot 4 \cdot 6 \cdot 3 \cdot 4}\right)^{-1} + \left(\frac{1}{3 \cdot 4}\right)^{-1} = 876$$

We can obtain the same number via the martingale interpretation. Suppose the current sequence of die outcomes until time $\ell = 11$ is given by $X^\ell = (1, 2, 4, 2, 3, 1, 2, 3, 4, 2, 3)$, then only two gamblers remain by time 11: gambler 7 and gambler 10. The other gamblers, $1, 2, 3, 4, 5, 6, 8, 9, 11$ leave the game at timesteps $1, 3, 3, 5, 5, 6, 8, 9, 11$, respectively. Using (22), we have total earnings

$$\sum_{n=1}^{\tau_B} R_{n,\tau_B} = W_B(\tau_B) - \tau_B \tag{23}$$

where $W_B(\tau_B)$ denotes the total winnings accumulated over all the gamblers, who bet on observing pattern B, until time τ_B. Using the numerical values and the specific X^n defined above, the expected winnings of gambler 7 is given by $1/(p_2p_3p_4p_2p_3) = 864$ at time 11, while the expected winnings of gambler 10 is given by $1/(p_2p_3) = 12$; the other gamblers have left the game with winnings 0. Hence:

$$0 = \mathbb{E}\left[\sum_{n=1}^{\tau_B} R_{n,\tau_B}\right] = \mathbb{E}[W_B(\tau_B)] - \mathbb{E}[\tau_B] = (864 + 12) - \mathbb{E}[\tau_B] \tag{24}$$

where the left-hand side of the above is equal to zero because it is a martingale by the optional stopping theorem. Therefore, algebraic manipulation yields $\mathbb{E}[\tau_B] = 864 + 12 = 876$, which matches with the result obtained using the method of Section 2.2. □

A quantity of interest related to $\mathbb{E}[\tau_B]$, for $B := (b_1, \cdots, b_k)$, is $\mathbb{E}[\tau_{B|A}]$ for $A := (a_1, \cdots, a_m)$, where

$$\tau_{B|A} := \min\{n \in \mathbb{N} \mid (X_{n+\tau_A}, \cdots, X_{n+\tau_A+k-1}) = B\} \tag{25}$$

denotes the additional time it takes to observe B given the occurrence of A. Computing $\mathbb{E}[\tau_{B|A}]$ is a simple extension of computing $\mathbb{E}[\tau_B]$. Similar to (23):

$$\sum_{n=\tau_A+1}^{\tau_B} R_{n,\tau_B} = \sum_{n=1}^{\tau_B} R_{n,\tau_B} - \sum_{n=1}^{\tau_A} R_{n,\tau_B} = W_B(\tau_B) - \tau_B - W_B(\tau_A) + \tau_A = W_B(\tau_B) - W_B(\tau_A) + \tau_{B|A} \tag{26}$$

since $\tau_{B|A} = \tau_B - \tau_A$, and $W_B(\tau_A)$ denotes the total winnings obtained by all the gamblers who gamble to view outcome B until time τ_A. Below, we elaborate on the 12-sided die example to discuss the computation of $W_B(\tau_A)$.

Example 7 (12-Sided Die Continued)**.** In continuation of the setup of Example 6, suppose that we have already observed a die outcome of $A := (a_1, \cdots, a_m) := (2, 4, 2, 3)$, with length $m = 4$. Thus, a total of 4 gamblers play a game throughout A, and obtain various rewards depending on (partial) observance of B. Based on (22) for this particular sequence, the only gambler who has positive expected winnings the third gambler, who wins $1/(p_2 p_3) = 12$; all the other gamblers have 0 expected winnings. Thus, $\mathbb{E}[W_B(\tau_A)] = 12$, and so

$$0 = \mathbb{E}\left[\sum_{n=\tau_A+1}^{\tau_B} R_{n,\tau_B}\right] = \mathbb{E}[W_B(\tau_B)] - \mathbb{E}[W_B(\tau_A)] + \mathbb{E}[\tau_{B|A}] \implies \mathbb{E}[\tau_{B|A}] = 876 - 12 = 864$$

$\square$

For the sake of easier notation, we further denote

$$A\Diamond B = \prod_{i=1}^{\min(m,k)} \delta_{i,i} + \prod_{i=1}^{\min(m-1,k)} \delta_{i+1,i} + \prod_{i=1}^{\min(m-2,k)} \delta_{i+2,i} + \cdots + \delta_{\min(m,k),1} \tag{27}$$

Essentially, (27) conveys the total winnings obtained by a gambler who is betting to observe (partial) occurrences of B in the sequence A. As a result, we can rewrite the expressions for $\mathbb{E}[\tau_B]$ and $\mathbb{E}[\tau_{B|A}]$ defined above using (27), described in the following theorem.

Theorem 3 (Expected Waiting Time Until a Sequence)**.** The expected waiting time for a sequence $B := (b_1, \cdots, b_n)$ is given by

$$\mathbb{E}[\tau_B] := B\Diamond B \tag{28}$$

Given a starting sequence $A := (a_1, \cdots, a_m)$, the expected waiting time for a sequence $B := (b_1, \cdots, b_n)$ is given by

$$\mathbb{E}[\tau_{B|A}] := B\Diamond B - A\Diamond B \tag{29}$$

The proof of Theorem 3 basically hinges on the fact that the stopped process $\{\sum_{n=1}^{\ell\wedge\tau} R_{n,\ell\wedge\tau}\}_{\ell=1}^{\infty}$, for some stopping time τ, is a martingale. Formally, we require the following important theorem for martingales.

Lemma 2 (Doob's Martingale Convergence)**.** Let $\{X_n\}_{n=1}^{\infty}$ be a martingale on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, P)$, i.e. $X_m = \mathbb{E}[X_n|\mathcal{F}_m]$ for $m \leq n$. Let τ be a stopping time for the martingale. If $\mathbb{E}[X_\tau] < \infty$ and X_n is uniformly-integrable, i.e.

$$\liminf_{n\to\infty} \int_{\{\omega\in\Omega:|\tau(\omega)|>n\}} |X_n(\omega)| dP(\omega) = 0$$

then $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$.

Proof of Theorem 3. By the notation introduced in (21) and (22):

$$R_{n,\ell\wedge\tau_{B|A}} = \sum_{j=0}^{k-1} Y_{n,j}\mathbb{1}\{n+j \leq \ell\wedge\tau_{B|A}\}, \qquad \sum_{n=m-1}^{r} R_{n,\ell\wedge\tau_{B|A}} = AX^r\Diamond B - (r+m),\ r \geq \tau_{B|A} \tag{30}$$

where AX^r denotes the concatenated sequence $(a_1, \cdots, a_m, X_1, \cdots, X_r)$. We have that $\{Y_{n\wedge\tau_{B|A}}\}$ is a martingale as well.

Recall that $\sum_{n=1}^{\tau_{B|A}} R_{n,\tau_{B|A}} = B\Diamond B - \tau_{B|A}$. Because $\mathbb{E}[\tau_{B|A}] < \infty$, we have that $\mathbb{E}[|\sum_{n=1}^{\tau_{B|A}} R_{n,\tau_{B|A}}|] \leq B\Diamond B - \tau_{B|A} < \infty$, and on the event set $\{\tau_{B|A} > r\}$ for some $r \in \mathbb{N}$, $|\sum_{n=1}^{\tau_{B|A}} R_{n,\tau_{B|A}}| \leq |AX^r \Diamond B| + r \leq B\Diamond B + \tau_{B|A}$. We can apply Doob's martingale convergence lemma to get

$$\lim_{r\to\infty} \int_{\{\omega\in\Omega:\tau_{B|A}(\omega)>r\}} |\sum_{n=1}^{r} R_{n,r}(\omega)| dP(\omega) \leq \lim_{r\to\infty} \int_{\{\omega\in\Omega:\tau_{B|A}(\omega)>r\}} (B\Diamond B + \tau_{B|A}) dP(\omega) = 0 \tag{31}$$

Because $\{\sum_{n=1}^{\tau_{B|A}} R_{n,\tau_{B|A}}\}$ is a martingale, $\mathbb{E}[\sum_{n=1}^{\tau_{B|A}} R_{n,\tau_{B|A}}] = \mathbb{E}[R_{1,1}] = A\Diamond B$ since $X^0 = ()$ and the above formula shows that $\mathbb{E}[\sum_{n=1}^{\tau_{B|A}} R_{n,\tau_{B|A}}] = B\Diamond B - \mathbb{E}[\tau_{B|A}]$ and so $\mathbb{E}[\tau_{B|A}] = B\Diamond B - A\Diamond B$. This yields our desired results. ■

Based on Theorem 3, we have two additional important results.

Theorem 4. Suppose $\{X_n\}_{n=1}^{\infty}$ is a sequence of discrete i.i.d. random variables, and let $\mathscr{A} = \{A_1, \cdots, A_K\}$, for $K \in \mathbb{N}$, be a set of patterns of interest. Denote m_i to be the length of pattern A_i. Given some starting sequence of random variables $B := (b_1, \cdot, b_k)$, denote

$$\tau_i := \{n \in \mathbb{N} \mid (X_{n+k}, \cdots, X_{n+k+m_i-1}) = A_i\} \tag{32}$$

to be the time of the first occurrence of A_i after observing B, and denote $\tau := \min\{\tau_1, \cdots, \tau_K\}$ to be the time of the first occurrence of any one of the patterns in $\mathscr{A}$. We have the following results.

- For all $i \in \{1, \cdots, K\}$, let $p_i := \mathbb{P}(\tau = \tau_i)$. Then

$$\sum_{j=1}^{K} p_j (A_j \Diamond A_i) = \mathbb{E}[\tau] + B\Diamond A_i \tag{33}$$

- If A_i and A_j are not subsequences of each other, then the odds that A_j precedes A_i in $X_1, X_2, \cdots$ are

$$(A_i\Diamond A_i - A_i\Diamond A_j) : (A_j\Diamond A_j - A_j\Diamond A_i) \tag{34}$$

Proof. For the first result, note

$$\mathbb{E}[\tau_i] = \mathbb{E}[\tau] + \mathbb{E}[\tau_i - \tau] = \mathbb{E}[\tau] + \sum_{j=1}^{K} p_j \mathbb{E}[\tau_i - \tau_j | \tau = \tau_j] \tag{35}$$

Note that by Theorem 3, $\mathbb{E}[\tau_i] = A_i\Diamond A_i - B\Diamond A_i$ and $\mathbb{E}[\tau_{i|j}] = A_i\Diamond A_i - A_j\Diamond A_i$. Substituting both into (35), we get

$$A_i\Diamond A_i - A\Diamond A_i = \mathbb{E}[\tau] + \sum_{j=1}^{K} p_j (A_i\Diamond A_i - A_j\Diamond A_i) \implies \mathbb{E}[\tau] = A\Diamond A_i - \sum_{j=1}^{K} p_j (A_j\Diamond A_i) \tag{36}$$

which is exactly the desired equation.

The second result follows directly from Theorem 3. ■

Theorem 5. From observing a discrete-valued i.i.d. sequence $\{X_n\}_{n=1}^{\infty}$ of random variables, let $A := (a_1, a_2, \cdots, a_m)$ and $B := (a_2, \cdots, a_m)$, for $m \in \mathbb{N}$.

$$\mathbb{P}(\tau_A = \tau_B) = \frac{B \lozenge B - B \lozenge A}{A \lozenge A - A \lozenge B - B \lozenge A + B \lozenge B} \tag{37}$$

Proof. Note that $\tau_A \geq \tau_B$ always because B is an ending subsequence of A. Denote $X^{\tau_B} := (X_1, \cdots, X_\ell, a_2, \cdots, a_m)$. Theorem 3 shows that the process $\{\sum_{n=1}^{\ell} R_{n,\ell \wedge \tau_{B|A}}\}_{\ell=1}^{\infty}$ is a martingale. Thus

$$\mathbb{E}[X^{\tau_B} \lozenge A - \tau_B] = 0, \quad \mathbb{E}[X^{\tau_B} \lozenge B - \tau_B] = 0$$

The second equality is obvious because there is always exactly one occurrence of B in the sequence X^{τ_B}. For the first equality, $X^{\tau_B} \lozenge A - \tau_B$ can be thought of as the total rewards of all the gamblers who are betting for sequence A until time τ_B, which is clearly a martingale process.

Combining the two together, note:

$$X^{\tau_B} \lozenge A - X^{\tau_B} \lozenge B = \begin{cases} A \lozenge A - A \lozenge B & \text{if } \tau_A = \tau_B \\ B \lozenge A - B \lozenge B & \text{if } \tau_A > \tau_B \end{cases}$$

because $\tau_A = \tau_B$ if X_ℓ in the sequence X^{τ_B} is equal to a_1, and $\tau_A > \tau_B$ otherwise. Using conditional expectations, we get:

$$0 = \mathbb{E}\left[X^{\tau_B} \lozenge A - X^{\tau_B} \lozenge B\right] = (A \lozenge A - A \lozenge B)\mathbb{P}(\tau_A = \tau_B) - (B \lozenge A - B \lozenge B)\mathbb{P}(\tau_A > \tau_B)$$

Rearranging the terms yields the desired expression. ∎

2.5 Extension of the Martingale Theoretic Interpretation

Consider the same problem as above. Suppose that we are given $\{X_n\}_{n=1}^{\infty} \subseteq \mathcal{X}$, a sequence of scalar-valued i.i.d. trials, with probability distribution $p(x) := \mathbb{P}(X_n = x)$ for any $n \in \mathbb{N}$ and $x \in \mathcal{X}$. We are interested in 1) the distribution of the time it takes to observe the pattern $B := (b_1, \cdots, b_r)$ for the first time, and 2) the distribution of the time it takes to observe the pattern B given the initial sequence $\{(X_0, X_1, \cdots, X_{m-1}) = A\}$ for $A := (a_1, \cdots, a_m)$, where $m, r \in \mathbb{N}$. As before, we denote $\tau_{0,B}$ to be the time of interest in the first question, and $\tau_{B|A}$ to be the time of interest in the second question. While the approach described in Section 2.4 derives formulas for the expected values of our interested quantities, we are also interested in the probability distributions, namely, their moment-generating functions.

Towards this end, it is useful to associate the following Markov chain to our problem. Let $\{Z_n\}_{n=1}^{\infty}$ be a Markov chain with state-space $\mathcal{Z} := \{0, 1, \cdots, r\}$, such that the state at time $n \in \mathbb{N}$ is

$$Z_n := \max\{j \in \mathbb{N} : X_{n-j+1} = b_1, \cdots, X_n = b_j, (X_0, X_1, \cdots, X_{m-1}) = A\}$$

Note that for all intermediate states of the chain $j \in \{1, \cdots, r-1\}$, transitions only occur to either $j+1$ with probability $p(b_{j+1})$ or 0 with probability $1 - p(b_{j+1})$. Moreover, we can write

$$\tau_{B|A} := \min\{n \in \mathbb{N} : Z_n = r\} \tag{38}$$

A corresponding Markov chain for rewriting $\tau_{0,B}$ can be derived similarly.

To compute the generating function of $\tau_{B|A}$, redefine the notations

$$\delta_{i,j}(z, A, B) := \begin{cases} \frac{1}{z\mathbb{P}(X=b_j)} & \text{if } a_i = b_j \\ 0 & \text{else} \end{cases} \tag{39}$$

and $A\Diamond B(z)$ equal to (27) with all instances of $\delta_{i,j}(A, B)$ replaced by $\delta_{i,j}(z, A, B)$. Then the cumulative reward obtained by time $n \in \mathbb{N}$ for the ℓth gambler is given by

A version of Theorem 3 for the generating function can then be described as follows.

Theorem 6 (Generating Function of the Waiting Time Until a Sequence)**.** Under the setup described above,

$$\mathbb{E}[z^{\tau_{B|A}}] = \frac{1 + (1-z)A\Diamond B(z)}{1 + (1-z)B\Diamond B(z)} \tag{40}$$

Proof. ■

Example 8. Consider an i.i.d. sequence of coin tosses with state-space $\mathcal{X} := \{H, T\}$.

□

3 Common Types of Renewal Processes

3.1 Important Theorems

Definition 3 (Renewal Reward Process)**.** Consider a basic renewal process $\{N(t), t \geq 0\}$ of the form defined in Definition 1. Suppose that at the time of the nth renewal from the process $\{N(t), t \geq 0\}$, a reward of R_n is gained, and $\{R_n\}_{n=1}^{\infty}$ forms an i.i.d. sequence. Such types of renewal processes are referred to as *renewal reward processes.* □

Definition 4 (Delayed Renewal Process)**.** Suppose that the distribution of the first interarrival time $X_1 \sim G$ of a renewal process $N(t)$ is different from the distribution of each successive interarrival time $X_2, X_3, \cdots \sim F$, where both G and F have finite mean. Then $N(t)$ is referred to as a *delayed renewal process.* We henceforth denote the renewal process $\{N(t), t \geq 0\}$ as $\{N_D(t), t \geq 0\}$ if it is delayed. □

Definition 5 (Regenerative Renewal Process)**.** A renewal process $\{N(t), t \geq 0\}$ with state space $\mathbb{Z}^+$ is said to be a *regenerative renewal process* if there exist times $\{T_i^{(0)}\}_{i=1}^{\infty}$ at which the process restarts itself with probability 1. We henceforth denote $C_j^{(0)}$ to be the time duration of the jth *cycle*, i.e. time between consecutive visits of the repeated state. □

Definition 6 (Alternating Renewal Process)**.** An *alternating renewal process* is a renewal process which switches among $M \in \mathbb{N}$ states. The duration of time spent in each possible state $k \in \{1, \cdots, M\}$ is distributed as F_k, and are independent of the time spent in the other states. The most common alternating renewal process considers only $M = 2$ states. □

Before we describe some important limiting theorems pertaining to the basic renewal process Definition 1 and its extensions defined above, we make a few additional definitions.

Definition 7 (Renewal Function)**.** The *renewal function* $m(t) := \mathbb{E}[N(t)]$ denotes the mean number of renewals of the renewal process $N(t)$ by time t. $\square$

Definition 8 (Lattice Distribution)**.** A nonnegative random variable S with cdf F is *lattice* if there exists a $c \geq 0$ such that S only takes on values which are integer multiples of c.

$$\sum_{n=0}^{\infty} \mathbb{P}(S = nc) = 1$$

The largest such c in which this property holds is referred to as the *period* of S. $\square$

We overload the terminology of "lattice" to describe both the random variable S and its distribution function F.

Definition 9 (Ladder Variables)**.** An *ascending variable of ladder height* S_n occurs at time n if $S_n > \max\{0, S_1, \cdots, S_{n-1}\}$.
Similarly, define *descending variable of ladder height* S_n at time n if $S_n < \max\{0, S_1, \cdots, S_{n-1}\}$. $\square$

Suppose there is an ascending ladder variable of height S_n at time n. Then the next ladder variable occurs at time $n + m$ if m is the first value such that

$$S_{n+m} > S_n \iff X_{n+1} + \cdots + X_{n+m} > 0$$

But note that since X_i are i.i.d., the probability of this event happening is the same regardless of which timestep we start on.

$$\mathbb{P}(X_{n+1} + \cdots + X_{n+m} > 0) = \mathbb{P}(X_1 + \cdots + X_m > 0) \tag{41}$$

Theorem 7 (Basic Theorems for Renewal Processes)**.** We provide multiple results on the limiting behavior of certain types of renewal processes.

1. **Elementary Renewal Theorem**: Let $\{X_i\}_{i=1}^{\infty}$ be i.i.d. sequence of interarrival times such that $\mathbb{E}[X_i] := \mathbb{E}[X] = \mu < \infty$ for all i, and $N(t)$ be the corresponding renewal process. With this setup, the following limiting relationship holds:

$$\frac{m(t)}{t} \to \frac{1}{\mu} \text{ as } t \to \infty \tag{42}$$

   where $m(t)$ is the renewal function.

2. **Blackwell's Theorem**: Let $\{X_i\}_{i=1}^{\infty}$ be i.i.d. sequence such that $\mathbb{E}[X_i] := \mathbb{E}[X] = \mu < \infty$ for all i, and let F be the distribution function of X_i. Let $N(t)$ be the corresponding renewal process, with $m(t)$ its renewal function.

   (a) If F is not lattice, then

$$m(t + a) - m(t) \to \frac{a}{\mu} \quad \text{as} \quad t \to \infty \tag{43}$$

   for all $a \geq 0$.

   (b) If F is lattice with period $c \geq 0$, then

$$\mathbb{E}[N(nc)] \to \frac{c}{\mu} \quad \text{as} \quad n \to \infty \tag{44}$$

   Note that $N(nc)$ denotes the number of renewals which have occurred by time nc.

3. **Blackwell's Theorem for Renewal Reward Processes**: Suppose the renewal process has i.i.d. rewards $\{R_i\}$. Assuming the distribution of the interarrival times $\{X_i\}_{i=1}^{\infty}$ of the renewal reward process is not lattice with $\mathbb{E}[X_i] := \mathbb{E}[X] := \mu$, then

$$\mathbb{E}[R(t, t+a)] \to \frac{a\mathbb{E}[R]}{\mu} \qquad \text{as } t \to \infty \tag{45}$$

4. **Blackwell's Theorem for Random Walks on the Line**: Let $S_n := \sum_{i=1}^{n} X_i$ be a random walk with $\mu := \mathbb{E}[X] > 0$ and $S_0 = 0$. Denote

$$U(t) := \sum_{n=1}^{\infty} I_n \quad \text{where} \quad I_n = \begin{cases} 1 & \text{if } S_n \leq t \\ 0 & \text{else} \end{cases}$$

If $\mu > 0$ and X_i are nonlattice, then

$$\frac{u(t+a) - u(t)}{t} \to \frac{a}{\mu} \quad \text{as} \quad t \to \infty \; \forall a > 0 \tag{46}$$

Example 9 (Theorem 7 to the Poisson Process)**.** Before we proceed onwards with the proof, the Elementary Renewal Theorem has a very intuitive interpretation which is easy to see in the well-known case of the Poisson process, in which the interarrival times X_n are distributed according to exponential random variables. Suppose vehicles arrive at an intersection at an average rate of $\lambda = 3$ per hour. This implies that the average time between two consecutive vehicle arrivals is $1/\lambda = 1/3$ hour, or 20 min. Hence, with time period t denoting one hour, we have that $m(t)/t = 3$, and $1/\mu = 1/(1/3) = 3$, which verifies (42).

Blackwell's theorem also admits an intuitive interpretation. Suppose that instead of one hour, the time period of consideration is 5 hours. With a rate of $\lambda = 3$, it is easy to expect that the average number of vehicles observed is $3 \cdot 5 = 15$. This corresponds to the statement of the theorem with $a = 5$.

Blackwell's Theorem for renewal reward processes also admits an intuitive interpretation. Suppose that vehicles enter a parking lot which charges different costs to park depending on the vehicle's characteristics. Let the values of the costs range across three discrete values $R_n \in \{10, 20, 30\}$. Suppose that $\lambda = 3$ vehicles arrive within a time range of one hour, with cost random variables R_1, R_2, R_3. Then the average cost among the three vehicles is computed as $\mathbb{E}[R_1 + R_2 + R_3]/3 = 20$. In addition, for a time interval of 5 hours, we multiply this amount by 5: the average reward accumulated over a timespan of 5 hours is then given by 100. $\square$

Proof. 1. **Elementary Renewal Theorem**: We carry out this proof through two parts, we show that $\lim_{t\to\infty} \frac{m(t)}{t} \geq \frac{1}{\mu}$, then show that $\lim_{t\to\infty} \frac{m(t)}{t} \leq \frac{1}{\mu}$.

Note that $T_{N(t)+1}$ denotes the time of the first renewal after time t, which can alternatively be represented as

$$T_{N(t)+1} = t + \Delta T_{N(t)} \tag{47}$$

where we refer to $\Delta T_{N(t)}$ as the "excess" time from t until the next renewal.

By Wald's equation, note that

$$\mathbb{E}[T_{N(t)+1}] = \mathbb{E}\left[\sum_{n=1}^{N(t)+1} X_n\right] = \mathbb{E}[X]\mathbb{E}[N(t)+1] = \mu(m(t)+1)$$

which, in combination with (47), yields

$$\mu(m(t)+1) = t + \mathbb{E}[\Delta T_{N(t)}] \implies \frac{m(t)}{t} + \frac{1}{t} = \frac{1}{\mu} + \frac{\mathbb{E}[\Delta T_{N(t)}]}{t\mu} \tag{48}$$

where the second equality follows from dividing through by $t\mu$. Since $\Delta T_{N(t)} \geq 0$, it follows that

$$\frac{m(t)}{t} \geq \frac{1}{\mu} - \frac{1}{t} \implies \lim_{t\to\infty} \frac{m(t)}{t} \geq \frac{1}{\mu} \tag{49}$$

To prove the other half, suppose there exists a value $C < \infty$ such that $\mathbb{P}(X_i < C) = 1$. This implies that $\Delta T_{N(t)} < C$, and so (48) implies

$$\frac{m(t)}{t} \leq \frac{1}{\mu} + \frac{C}{t\mu} - \frac{1}{t} \implies \lim_{t\to\infty} \frac{m(t)}{t} \leq \frac{1}{\mu} \tag{50}$$

Thus, when the interarrival times are bounded, the Elementary Renewal Theorem holds. In the case where they are unbounded, again fix $C > 0$, and define $\{N_C(t), t \geq 0\}$ to be the renewal process with interarrival times $\min(X_n, C), n \geq 1$. Since $\min(X_n, C) \leq X_n$ for all $n \geq 0$, $N_C(t) \geq N(t)$ for all $t \geq 0$ since the interarrival times are shorter. Consequently:

$$\lim_{t\to\infty} \frac{\mathbb{E}[N(t)]}{t} \leq \lim_{t\to\infty} \frac{\mathbb{E}[N_C(t)]}{t} = \frac{1}{\mathbb{E}[\min(X_n, C)]} \tag{51}$$

where the second equality follows from the first case, where the interarrival times are bounded. Since $\lim_{C\to\infty} \mathbb{E}[\min(X_n, C)] = \mathbb{E}[X_n] = \mu$, (51) becomes

$$\lim_{t\to\infty} \frac{m(t)}{t} \leq \frac{1}{\mu}$$

The two cases together, in combination with (49) yields the desired result (42).

2. **Blackwell's Theorem**: We will only prove the nonlattice case of the theorem statement, since the lattice case follows from a similar argument. For easier notation, let us denote $g(a) := \lim_{t\to\infty}(m(t+a) - m(t))$. Then note that

$$g(a+b) = \lim_{t\to\infty}(m(t+a+b) - m(t+a)) + \lim_{t\to\infty}(m(t+a) - m(t)) = g(b) + g(a)$$

The form of g which satisfies this equation is given by $g(a) = c \cdot a$, for some constant c. Now we want to show that $c = 1/\mu$. Consider the following successive increments of $a = 1$:

$$\begin{aligned} \Delta m_1 &:= m(1) - m(0) \\ \Delta m_2 &:= m(2) - m(1) \\ &\vdots \\ \Delta m_n &:= m(n) - m(n-1) \end{aligned}$$

Note that $\lim_{n\to\infty} \Delta m(n) = c$.

On one hand, by the law of large numbers:

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \Delta m(k)}{n} = c \tag{52}$$

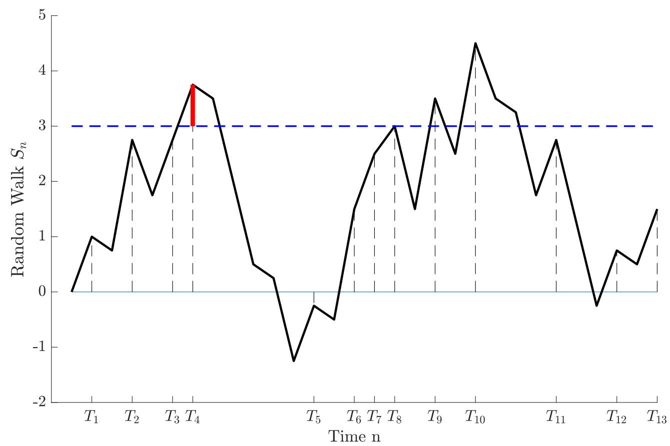

Figure 2: The ascending ladder height renewal process with the excess at height $c = 3$. The excess $Y(t)$ is the length of the thick red line.

On the other hand, by the Elementary Renewal theorem:

$$\frac{\sum_{k=1}^{n} \Delta m(k)}{n} = \frac{m(n)}{n} \to \frac{1}{\mu} \text{ as } n \to \infty \tag{53}$$

Combining (52) together with (53), we have that $c = 1/\mu$. This concludes the proof of the first part of the theorem.

3. **Blackwell's Theorem for Renewal Reward Processes**: Note that

$$\begin{aligned}
\mathbb{E}[R(t, t+a)] &:= \mathbb{E}\left[\sum_{n=0}^{N(t+a)} R_n - \sum_{n=0}^{N(t)} R_n\right] \\
&= \left(\mathbb{E}[N(t+a)] - \mathbb{E}[N(t)]\right)\mathbb{E}[R] \quad \text{by Wald's equation} \\
&= (m(t+a) - m(t))\mathbb{E}[R] \quad \text{by } m(t) \text{ definition} \\
&\to \frac{a}{\mu} \cdot \mathbb{E}[R] \quad \text{as } t \to \infty
\end{aligned}$$

where the limit follows from Blackwell's theorem (item 2 above).

4. **Blackwell's Theorem for Random Walks on the Line**: Define a renewal process using successive ascending ladder heights as the renewals, and denote $Y(t)$ to be the excess height of the renewal process past height $t > 0$. Then $S_n = Y(t) + t$ is the first value of the random walk that exceeds t. See Figure 2 for visualization.

   Hence, the primary difference between Blackwell's Theorem for Random Walks and the original Blackwell's Theorem or the Elementary Renewal Theorem is that the renewal process has renewals which are dependent upon the previous renewals.

   The proof of this theorem adheres very closely to the technique of the proof for the original Blackwell's theorem (43). Define

$$h(a) := \lim_{t \to \infty} (u(t+a) - u(a))$$

   Then we can show that $h(a+b) = h(a) + h(b)$ using the same logic as in the proof of (43). The solution to such an equation is then given by $h(a) = \theta a$ for some constant θ to be determined.

Let $\tau(t) := \min\{n|S_n > t\}$. We look at two cases of X_i, just as in the proof of (43).

First, if there exists $M > 0$ such that $X_i \leq M$ for all i, then

$$\begin{aligned} t &< \sum_{i=1}^{\tau(t)} X_i < t + M \\ \implies t &< \mathbb{E}[\tau(t)]\mathbb{E}[X_i] < t + M \quad \text{by Wald} \\ \implies \frac{1}{\mu} &< \frac{\mathbb{E}[\tau(t)]}{t} < \frac{1 + \frac{M}{t}}{\mu} \end{aligned}$$

which implies that

$$\lim_{t\to\infty} \frac{\mathbb{E}[\tau(t)]}{t} = \frac{1}{\mu} \tag{54}$$

On the other hand, if X_i are unbounded, denote

$$N^*(t) := |\{n > \tau(t)|S_n \in (-\infty, t]\}|$$

to be the number of times after exceeding c once S_n lands in $(-\infty, t]$.

Then note that

$$U(t) = |\{n|S_n \leq t\}| = (\tau(t) - 1) + N^*(t) \tag{55}$$

where the -1 comes from excluding the time S_n exceeded t.

Note that given the value of $Y(t) = y$, the distribution of $U(t + a) - U(t)$ is independent of t. If the first point of the random walk past c occurs at $y + t$, then the number of points in $(t, t + a)$ has the same distribution as the number of points in $(0, a)$ given the first positive value of the random walk occurs at y.

Thus,

$$\mathbb{E}[N^*(t)] \leq \mathbb{E}\left[|\{n > \tau(0)|S_n < 0\}|\right]$$

Since $\mu > 0$, $\mathbb{E}[\tau(0)] < \infty$. At time $\tau(0)$, there is probability $1 - p_- > 0$ such that $S_n > S_{\tau(0)}$ for all $n > \tau(0)$. Otherwise, if there is such an n where $S_n < S_{\tau(0)}$, then the expected additional time $m > 0$ such that $S_{n+m} > 0$ is finite since $\mu > 0$. From time $n + m$, there is again a probability of $1 - p_-$ $S_{n+k} > S_{\tau(0)}$ for all $k > m$. Thus:

$$(1 - p_-)\mathbb{E}\left[|\{n > \tau(0)|S_n < 0\}|\right] \leq \mathbb{E}[\tau(0)|X_1 < 0] \implies \mathbb{E}\left[|\{n > \tau(0)|S_n < 0\}|\right] \leq \frac{\mathbb{E}[\tau(0)|X_1 < 0]}{1 - p_-} < \infty$$

This shows that $\mathbb{E}[N^*(t)] < \infty$, and combined with (55), yields

$$\lim_{t\to\infty} \frac{\mathbb{E}[\tau(t)]}{t} = \lim_{t\to\infty} \frac{u(t)}{t} \tag{56}$$

Finally, to conclude the result from both (54) and (56), note that $u(1 + a) - u(a) \to \theta$ as $a \to \infty$. This implies

$$\frac{u(n + 1) - u(1)}{n} = \frac{\sum_{a=1}^{n} u(1 + a) - u(a)}{n} \to \theta \quad \text{as} \quad n \to \infty$$

Indeed, $\theta = 1/\mu$ and this concludes the proof.

$\blacksquare$

3.2 The Renewal Equation

Lemma 3 (Renewal Equation)**.** Given a renewal process $N(t)$ with interarrival time cdf $X_i \sim F$ for all $i \in \mathbb{N}$, its renewal function $m(t)$ satisfies the *renewal equation*

$$m(t) = F(t) + \int_0^t m(t-s)dF(s) \tag{57}$$

If F is a continuous cdf, we can write $dF(s) = F'(s)ds = f(s)ds$, where f is the pdf of the interarrival times X_i.

Proof. Condition on the first interarrival time:

$$m(t) := \mathbb{E}[N(t)] = \int_0^\infty \mathbb{E}[N(t)|X_1 = x]dF(x) \tag{58}$$

Note that we have two cases depending on whether $X_1 \leq t$ and when $X_1 > t$. If $X_1 > t$, this means there have been no arrivals until time t, i.e. $N(t) = 0$. Moreover, we can use the memoryless and stationarity property of the renewal process to conclude the recursive relationship $\mathbb{E}[N(t)|X_1 = x] = 1 + \mathbb{E}[N(t-x)]$. Thus:

$$(58) = \int_0^t \mathbb{E}[N(t)|X_1 = x]dF(x) = \int_0^t (1 + \mathbb{E}[N(t-x)])\, dF(x) = F(t) + \int_0^t \mathbb{E}[N(t-x)]dF(x) \tag{59}$$

and recognizing that $\mathbb{E}[N(t-x)] = m(t-x)$ yields the desired result. ■

Definition 10 (Convolution)**.** For two functions $h, g : [0,\infty) \to \mathbb{R}$, define the *convolution* $h * g$ by the function

$$(h * g)(t) = \int_0^t h(t-x)dg(x) \tag{60}$$

In particular, (57) is rewritten as $m(t) = F(t) + (m * F)(t)$. □

We can use the *Laplace-Stieltjes transform* to derive a closed-form expression of the renewal function $m(t)$ from (57). Recall the definition of the transform: for any function $H : [0,\infty) \to \mathbb{R}$ with derivative $h : [0,\infty) \to \mathbb{R}$, its Laplace-Stieltjes transform is given by

$$\hat{H}(z) = \int_0^\infty e^{-zt}dH(t) = \int_0^\infty e^{-zt}h(t)dt \tag{61}$$

Under the Laplace-Stieltjes transform, convolutions of functions turn into the multiplication of their transforms. Hence, (57) becomes

$$\hat{M}(z) = \hat{F}(z) + \hat{M}(z)\hat{F}(z) \implies \hat{M}(z) = \frac{\hat{F}(z)}{1 - \hat{F}(z)} \tag{62}$$

from which we can obtain $m(t)$ by performing the inverse Laplace-transform on $\hat{M}(z)$.

Definition 11 (Riemann-Integrable)**.** Let $f : \mathbb{R}^+ \to \mathbb{R}$ such that

$$\underline{f}_n(a) \leq f(t) \leq \overline{f}_n(a) \quad \text{for} \quad t \in [(n-1)a, na]$$

We say that f is *Riemann-integrable* if

- $\sum_{n=1}^{\infty} \underline{f}_n(a)$ and $\sum_{n=1}^{\infty} \overline{f}_n(a)$ are finite for all $a > 0$
- $\lim_{a \to 0} a \sum_{n=1}^{\infty} \underline{f}_n(a) = \lim_{a \to 0} a \sum_{n=1}^{\infty} \overline{f}_n(a)$

Furthermore, a sufficient condition for f to be Riemann integrable is that 1) $f(t) \geq 0$ for all $t \geq 0$, 2) $f(t)$ is nonincreasing, 3) $\int_0^\infty f(t)dt < \infty$. □

Theorem 8 (Key Renewal Theorem)**.** Let $\{N(t), t \geq 0\}$ be a renewal process with interarrival times $\{X_i\}_{i=1}^\infty$ being an i.i.d. sequence with distribution function F (i.e., $F(t) := \mathbb{P}(X_i = t)$) such that $\mathbb{E}[X_i] := \mathbb{E}[X] = \mu < \infty$ for all i. In addition, let $f : \mathbb{R}^+ \to \mathbb{R}$ be a Riemann-integrable function. Then the following equality holds:

$$\lim_{t\to\infty} \int_0^t f(t-s)dm(s) = \frac{1}{\mu} \int_0^\infty f(s)ds$$

where $m(t)$ is the renewal function, which can be written as

$$m(t) := \mathbb{E}[N(t)] = \sum_{n=1}^{\infty} F_n(t), \quad \mu := \int_0^\infty \overline{F}(t)dt$$

and F_n is the distribution of arrival time T_n. Note that T_n is the n-fold convolution of interarrival distribution F, and $\overline{F}(t) := \mathbb{P}(X_i > t)$ for all i.

Example 10 (Renewal Function for Poisson Arrivals)**.** Consider arrivals of a Poisson process with intensity $\lambda > 0$. This means the pdf of each interarrival time is $f(t) := \lambda e^{-\lambda t}$ and the cdf is $F(t) := 1 - e^{-\lambda t}$. The Laplace transform of $F(t)$ is given by

$$\hat{F}(z) = \int_0^\infty \lambda e^{-\lambda t} \cdot e^{-zt} dt = \frac{\lambda}{z+\lambda} \tag{63}$$

Thus, (62) becomes

$$\hat{M}(z) = \frac{\frac{\lambda}{z+\lambda}}{1 - \frac{\lambda}{z+\lambda}} = \frac{\lambda}{z}$$

for which the inverse Laplace transform is given by $m(t) = \lambda t$. □

The renewal function for delayed renewal processes can be obtained by using the same means as (62). The only difference is that the initial interarrival time X_1 is distributed differently (cdf G) than the other interarrival times $X_2, X_3, \cdots$ (cdf F). Hence, the renewal equation (57) is modified as

$$m(t) = G(t) + \int_0^t m(t-s)dF(s) \tag{64}$$

and using the Laplace transform, (64) becomes

$$\hat{M}(z) = \hat{G}(z) + \hat{M}(z)\hat{F}(z) \implies \hat{M}(z) = \frac{\hat{G}(z)}{1 - \hat{F}(z)} \tag{65}$$

The importance of the Key Renewal theorem arises in computing the limiting value of some probability or expectation-like function $g(t)$. This yields an equation of the form

$$g(t) = h(t) + \int_0^t h(t-s)dm(s)$$

for some Riemann-integrable function h. One specific function is the distribution of $T_{N(t)}$, the arrival time of the last renewal before time t.

$$
\begin{aligned}
\mathbb{P}(T_{N(t)} \leq s) &= \sum_{n=0}^{\infty} \mathbb{P}(T_n \leq s, N(t) \leq n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(T_n \leq s, T_{n+1} > t) \quad \text{since } \{N(t) \leq n\} \iff \{T_{n+1} > t\} \\
&= \mathbb{P}(T_0 \leq s, T_1 > t) + \sum_{n=1}^{\infty} \mathbb{P}(T_n \leq s, T_{n+1} > t) \\
&= \mathbb{P}(T_1 > t) + \sum_{n=1}^{\infty} \mathbb{P}(T_n \leq s, T_{n+1} > t) \quad \text{since } T_0 := 0 \leq s \text{ always} \\
&= \overline{F}(t) + \sum_{n=1}^{\infty} \mathbb{P}(T_n \leq s, T_{n+1} > t) \\
&= \overline{F}(t) + \sum_{n=1}^{\infty} \int_0^s \mathbb{P}(T_n = r, T_{n+1} > t) dF_n(r) \\
&= \overline{F}(t) + \sum_{n=1}^{\infty} \int_0^s \mathbb{P}(T_{n+1} - T_n > t - r) dF_n(r) \\
&= \overline{F}(t) + \int_0^s \overline{F}(t-r) d\left(\sum_{n=1}^{\infty} F_n(r)\right) \\
&= \overline{F}(t) + \int_0^s \overline{F}(t-r) dm(r)
\end{aligned}
$$

Hence:

$$
\mathbb{P}(T_{N(t)} \leq s) = \overline{F}(t) + \int_0^s \overline{F}(t-r) dm(r) \tag{66}
$$

3.3 Other Results

For renewal processes, we can derive expressions for quantities of interest other than the limiting rate of renewals.

Lemma 4 (Limiting Expected Value of the Next Reward)**.** For a renewal reward process $N(t)$ with interarrival times $X_i \sim A$ with $\mathbb{E}[X] =: \mu$, the following limit holds:

$$
\mathbb{E}[R_{N(t)+1}] \to \frac{\mathbb{E}[RX]}{\mu} \quad \text{as } t \to \infty \tag{67}
$$

Proof. We can condition on the value of $T_{N(t)}$:

$$
\mathbb{E}[R_{N(t)+1}] = \mathbb{E}[R_{N(t)+1} | T_{N(t)} = 0] \mathbb{P}(T_{N(t)} = 0) + \int_0^t \mathbb{E}[R_{N(t)+1} | T_{N(t)} = s] \mathbb{P}(T_{N(t)} = s) ds \tag{68}
$$

Note that we can use (66) to simplify the probabilities:

$$
\mathbb{P}(T_{N(t)} = 0) = \mathbb{P}(T_{N(t)} \leq 0) = \overline{F}(t)
$$

$$\mathbb{P}(T_{N(t)} = s) = d\mathbb{P}(T_{N(t)} \leq s) = \overline{F}(t-s)dm(s) \ \text{ by FTC}$$

Substituting:

$$\begin{aligned}
(68) &= \mathbb{E}[R_{N(t)+1} | T_{N(t)} = 0]\overline{F}(t) + \int_0^t \mathbb{E}[R_{N(t)+1} | T_{N(t)} = s]\overline{F}(t-s)dm(s) \\
&= \mathbb{E}[R_1 | X_1 > t]\overline{F}(t) + \int_0^t \mathbb{E}[R_n | X_n > t-s]\overline{F}(t-s)dm(s), \ n \geq 1 \\
&\to 0 + \frac{1}{\mu}\int_0^t \mathbb{E}[R_n | X_n > s]\overline{F}(s)ds \ \text{ by Key Renewal Theorem} \\
&= \frac{1}{\mu}\int_0^\infty \int_s^\infty \mathbb{E}[R_n | X_n = u]dF(u)ds
\end{aligned} \tag{69}$$

Note that we can perform a change of variables such that the integral from s to ∞ is equivalent to taking the integral from 0 to u.

$$(69) = \frac{1}{\mu}\int_0^\infty \left(\int_0^u ds\right)\mathbb{E}[R_n | X_n = u]dF(u) = \frac{1}{\mu}\int_0^\infty u\mathbb{E}[R_1 | X_1 = u]dF(u) = \frac{\mathbb{E}[R_1 X_1]}{\mu}$$

which indeed proves (67). ∎

Lemma 5 (Renewal Reward)**.** Let $R(t)$ represent the total reward earned by time t, i.e., $R(t) := \sum_{n=1}^{N(t)} R_n$. Denote $\nu := \mathbb{E}[R_n] < \infty, \mu := \mathbb{E}[X_n] < \infty$ for all $n \geq 1$. Then the following hold:

$$\lim_{t\to\infty} \frac{R(t)}{t} = \frac{\nu}{\mu} \quad \text{w.p. 1} \tag{70a}$$

$$\lim_{t\to\infty} \frac{\mathbb{E}[R(t)]}{t} = \frac{\nu}{\mu} \tag{70b}$$

Proof of Lemma 5. Since the proof of (70b) follows similarly to the proof of (70a), we will only prove (70a). Write:

$$\frac{R(t)}{t} := \frac{1}{t}\sum_{n=1}^{N(t)} R_n = \left(\frac{1}{N(t)}\sum_{n=1}^{N(t)} R_n\right)\left(\frac{N(t)}{t}\right)$$

By the strong law of large numbers, $(\sum_{n=1}^{N(t)} R_n)/N(t) \to \nu$ as $t \to \infty$, while $N(t)/t \to 1/\mu$ as $t \to \infty$ follows from (42). The combination yields (70a). ∎

Theorem 9 (Properties of a Delayed Renewal Process)**.** Consider a delayed renewal process $\{N_D(t), t \geq 0\}$, with first interarrival time distribution $X_1 \sim G$, where G has finite mean, and successive interarrival time distribution $X_2, X_3, \cdots \sim F$, where F is nonlattice with $\int s^2 dF(s) < \infty$. Then the following properties hold:

- Denoting $m_D(t) := \mathbb{E}[N_D(t)]$ to be its renewal function:

$$m_D(t) = G(t) + \int_0^t \sum_{n=1}^\infty F_n(t-s)dG(s) \tag{71}$$

- Denoting $H_D(t) := t - T_{N_D(t)}$ to be the age of the process by time t, we have $t\overline{G}(t) \to 0$ as $t \to \infty$ and

$$\mathbb{E}[H_D(t)] \to \frac{\int_0^\infty s^2 dF(s)}{2\int_0^\infty s dF(s)} \quad \text{as } t \to \infty \tag{72}$$

Proof. First, by utilizing the equivalence in (1) and conditioning on the first interarrival time:

$$\begin{aligned}
m_D(t) := \mathbb{E}[N_D(t)] &= \sum_{n=0}^{\infty} \mathbb{P}(N_D(t) > n) \\
&= \mathbb{P}(N_D(t) > 0) + \sum_{n=1}^{\infty} \mathbb{P}(N_D(t) > n) \\
&= \mathbb{P}(T_1 < t) + \sum_{n=1}^{\infty} \mathbb{P}(N_D(t) > n) \\
&= G(t) + \sum_{n=1}^{\infty} \int_0^t \mathbb{P}(N_D(t) > n | T_1 \leq s) dG(s) \\
&= G(t) + \sum_{n=1}^{\infty} \int_0^t \mathbb{P}(T_{n+1} \leq t | T_1 \leq s) dG(s) \\
&= G(t) + \sum_{n=1}^{\infty} \int_0^t \mathbb{P}(T_n \leq t - s) dG(s) \\
&= G(t) + \int_0^t \sum_{n=1}^{\infty} F_n(t-s) dG(s)
\end{aligned}$$

This proves (71).

Second, we have by the dominated convergence theorem

$$\lim_{t\to\infty} t\overline{G}(t) = \lim_{t\to\infty} t \int_t^{\infty} G(s)ds \leq \lim_{t\to\infty} \int_t^{\infty} sG(s)ds \tag{73}$$

Note that because the mean of G is assumed to be finite, the integrand of (73) is finite. Hence, when taking $t \to \infty$, the value of the overall integral tends to 0 since the upper and lower limits converge to the same value.

Now we can consider the age of the delayed renewal process. By conditioning on the value of $T_{N_D(t)}$ and substituting in (66) (and taking care to ensure that the distribution of the first interarrival time is given by G, not F), we get

$$\begin{aligned}
\mathbb{E}[H_D(t)] &= \mathbb{E}[H_D(t)|T_{N_D(t)} = 0]\mathbb{P}(T_{N_D(t)} = 0) + \int_0^t \mathbb{E}[H_D(t)|T_{N_D(t)} = s]\mathbb{P}(T_{N_D(t)} = s)ds \\
&= \mathbb{E}[H_D(t)|T_{N_D(t)} = 0]\overline{G}(t) + \int_0^t \mathbb{E}[H_D(t)|T_{N_D(t)} = s]\overline{F}(t-s)dm_D(s) \\
&= \mathbb{E}[H_D(t)|X_1 > t]\overline{G}(t) + \int_0^t \mathbb{E}[H_D(t)|T_{N_D(t)} = s]\overline{F}(t-s)dm_D(s), \quad \text{for } n \geq 2 \\
&= t\overline{G}(t) + \int_0^t (t-s)\overline{F}(t-s)dm_D(s)
\end{aligned} \tag{74}$$

since note that the age is simply the time elapsed since the last renewal. Now, note that under the assumption that G is finite, we showed above that $t\overline{G}(t) \to 0$ as $t \to \infty$, which implies that the first term of (74) tends to 0. We will thus ignore the first term in our analysis and focus primarily on the second term. Using the Key Renewal Theorem, we get:

$$\int_0^t (t-s)\overline{F}(t-s)dm_D(s) \to \frac{1}{\mu}\int_0^{\infty} s\overline{F}(s)ds \tag{75}$$

Note the relationship

$$\overline{F}(s) = \mathbb{P}(Y > s) = \int_0^\infty \mathbb{P}(Y = u)du$$

Substituting, we get:

$$\begin{aligned}
(75) &= \frac{1}{\mu}\int_0^\infty s\left(\int_s^\infty dF(u)\right)ds \\
&= \frac{1}{\mu}\int_0^\infty \left(\int_0^u sds\right)dF(u) \quad \text{via change of variables} \\
&= \frac{1}{2\mu}\int_0^\infty u^2 dF(u) = \frac{\int_0^\infty s^2 dF(s)}{2\int_0^\infty sdF(s)}
\end{aligned}$$

This proves (72). ■

Using (71), we can compute the distribution of the arrival time of the last renewal before time t for delayed renewal processes $N_D(t)$ with first interarrival distribution G and successive interarrival distribution F. Denote $H_n := G * F * \cdots * F$, where F is convolved $n-1$ times, to be the distribution of the nth arrival time of $N_D(t)$. We make a similar computation as in (66).

$$\begin{aligned}
\mathbb{P}(T_{N_D(t)} \leq s) &= \mathbb{P}(T_1 > t) + \sum_{n=1}^\infty \mathbb{P}(T_n \leq s, T_{n+1} > t) \\
&= \overline{G}(t) + \sum_{n=1}^\infty \int_0^\infty \mathbb{P}(T_n \leq s, T_{n+1} > t | T_n \leq r)dH_n(r) \\
&= \overline{G}(t) + \sum_{n=1}^\infty \int_0^\infty \mathbb{P}(T_{n+1} - T_n > t - r)dH_n(r) \\
&= \overline{G}(t) + \int_0^\infty \overline{F}(t-r)d\underbrace{\left(\sum_{n=1}^\infty H_n(r)\right)}_{= m_D(r) \text{ from } (71)}
\end{aligned}$$

Hence,

$$\mathbb{P}(T_{N_D(t)} \leq s) = \overline{G}(t) + \int_0^s \overline{F}(t-r)dm_D(r) \tag{76}$$

Theorem 10 (Time Spent in a State of a Regenerative Process)**.** Denote $C_j^{(0)}$ to be the time length of the jth cycle for regenerative process $\{N(t), t \geq 0\}$, and denote $C^{(0)}$ to be the time length of a generic cycle such that $C_j^{(0)}$ is distributed the same as $C^{(0)}$ for all $j \in \mathbb{N}$. Denote $C_{ji}^{(0)}$ to be the amount of time the process spends in state i during cycle j, and let $C_{\cdot i}^{(0)}$ denote the amount of time spent in state i during any generic cycle such that $C_{ji}^{(0)}$ is distributed the same as $C_{\cdot,i}^{(0)}$ for all $i, j \in \mathbb{N}$. Further suppose the distribution F of a cycle has a density over some interval of time and $\mathbb{E}[C_i^{(0)}] < \infty$ for all i. Then the following property holds:

$$P_i := \lim_{t\to\infty} P_i(t) := \lim_{t\to\infty} \mathbb{P}(N(t) = i) = \frac{\mathbb{E}[C_{\cdot,i}^{(0)}]}{\mathbb{E}[C^{(0)}]} \tag{77}$$

Proof. Conditioning on the time of the last cycle before time t, and using (66) yields:

$$P_i(t) = \mathbb{P}(N(t) = i | T^{(0)}_{N(t)} = 0)\overline{F}(t) + \int_0^t \mathbb{P}(N(t) = i | T^{(0)}_{N(t)} = s)\overline{F}(t-s)dm(s) \tag{78}$$

where $m(t) := \mathbb{E}[N(t)]$. We've seen before that

$$\mathbb{P}(N(t) = i | T^{(0)}_{N(t)} = 0) = \mathbb{P}(N(t) = i | C^{(0)}_1 > t)$$
$$\mathbb{P}(N(t) = i | T^{(0)}_{N(t)} = s) = \mathbb{P}(N(t) = i | C^{(0)}_1 > t - s)$$

and hence:

$$\begin{aligned}(78) &= \mathbb{P}(N(t) = i | C^{(0)}_1 > t)\overline{F}(t) + \int_0^t \mathbb{P}(N(t) = i | C^{(0)}_1 > t-s)\overline{F}(t-s)dm(s) \\ &\to \frac{1}{\mathbb{E}[C^{(0)}_1]} \int_0^\infty \mathbb{P}(N(t) = i | C^{(0)}_1 > s)\overline{F}(s)ds \quad \text{by Key Renewal Thm}\end{aligned} \tag{79}$$

and this proves the result since the integral expression above is exactly $C^{(0)}_{\cdot,i}$ for state i. ∎

Theorem 11 (Key Renewal Theorem for Alternating Renewal Processes)**.** Consider a two-state alternating renewal process defined by sequences of tuples $\{(U_i, D_i)\}_{i=1}^\infty$, where U_i is the duration of the ith up period and D_i is the duration of the ith down period. A single "cycle" is considered to be a single up period and a single down period. Define $X(t)$ denote the state of the renewal process, which takes value 1 when it is up and 0 when it is down. Further denote $G(t) := \mathbb{P}(U \le t)$ to be the cdf of the U_i and $F(t) := \mathbb{P}(C \le t)$ to be the cdf of the cycle duration of the renewal process. Then the probability $P_1(t)$ that the process is up at time t is given by

$$P_1(t) := \mathbb{P}(X(t) = 1) = \overline{G}(t) + \int_0^t \overline{G}(t-s)dm(s) \tag{80}$$

where $\overline{G}(t) := 1 - G(t)$, and $m(t) = \sum_{n=1}^\infty F_n(t)$.

Proof. Condition on the first cycle time C_1:

$$\begin{aligned}P_1(t) := \mathbb{P}(X(t) = 1) &= \mathbb{P}(X(t) = 1, C_1 > t) + \mathbb{P}(X(t) = 1, C_1 \le t) \\ &= \mathbb{P}(U_1 > t) + \int_0^t \mathbb{P}(X(t) = 1 | C_1 = s)\mathbb{P}(C_1 = s)ds\end{aligned} \tag{81}$$

Note that conditioning on $C_1 = s$ means that one full cycle of up and down has already passed by time t. By memorylessness of the renewal process, $\mathbb{P}(X(t) = 1 | C_1 = s) = \mathbb{P}(X(t - s) = 1)$. Hence:

$$(81) = \mathbb{P}(U_1 > t) + \int_0^t \mathbb{P}(X(t-s) = 1)\mathbb{P}(C_1 = s)ds = \overline{G}(t) + \int_0^t P_1(t-s)dF(s) \tag{82}$$

Take the Laplace-Stieltjes transform of (82). Denote $\hat{G}(z) := \mathcal{L}\{\overline{G}(t)\}$, $\hat{P}_1(z) := \mathcal{L}\{P_1(t)\}$, and $\hat{F}(z) := \mathcal{L}\{F(t)\}$. Then

$$\begin{aligned}P_1(t) = \overline{G}(t) + P_1(t) * F(t) &\iff \hat{P}_1(z) = \hat{G}(z) + \hat{P}_1(z)\hat{F}(z) \\ &\iff \hat{P}_1(z) = \frac{\hat{G}(z)}{1 - \hat{F}(z)} = \hat{G}(z)\left(1 + \frac{\hat{F}(z)}{1 - \hat{F}(z)}\right)\end{aligned} \tag{83}$$

and we have by (62)

$$(83) \implies \hat{P}_1(z) = \hat{G}(z) + \hat{G}(z)\hat{M}(z) \implies P_1(t) = \overline{G}(t) + \overline{G}(t) * m(t) \tag{84}$$

which is exactly the desired result (80). ∎

Definition 12 (Lifetimes)**.** Define $N(t)$ to be a renewal process with arrival times $\{T_i\}_{i=1}^{\infty}$. We define the following quantities:

1. the *residual lifetime* at time t is the time until the next arrival, given by $Y(t) := T_{N(t)+1} - t$.
2. the *age* at time t is the time since the last arrival, given by $Z(t) := t - T_{N(t)}$.

Summed together, the residual lifetime and the age form the interarrival time $X(t) = Y(t) + Z(t)$. □

Theorem 12 (Residual Lifetime Distribution)**.** Given a renewal process $N(t)$ with i.i.d. interarrival times $\{X_n\}_{n=1}^{\infty}$ distributed with cdf F, the cdf of the residual lifetime is given by

$$\mathbb{P}(Y(t) \leq y) = F(t+y) - \int_0^t (1 - F(t) + y - x)\, dm(x) \tag{85}$$

where $x, y, t > 0$, and $m(t) := \mathbb{E}[Y(t)]$.

Proof. Condition on the first arrival time T_1.

$$\mathbb{P}(Y(t) > y) = \int_0^{\infty} \mathbb{P}(Y(t) > y | X_1 = x) dF(x) \tag{86}$$

Note that

$$\mathbb{P}(Y(t) > y | X_1 = x) = \begin{cases} \mathbb{P}(Y(t-x) > y) & \text{if } x \leq t \\ 0 & \text{if } x \in (t, t+y] \\ 1 & \text{if } x > t+y \end{cases}$$

which yields

$$(86) = \int_0^t \mathbb{P}(Y(t) > y | X_1 = x) dF(x) + \int_{t+y}^{\infty} dF(x) = \int_0^t \mathbb{P}(Y(t-x) > y) dF(x) + (1 - F(t+y)) \tag{87}$$

Define $\phi(t) := \mathbb{P}(Y(t) > y)$ and $\psi(t) := 1 - F(t+y)$. Then (87) becomes

$$\phi(t) = \psi(t) + \int_0^t \phi(t-x) dF(x) \tag{88}$$

which has a closed-form solution that can be obtained by taking the Laplace-Stieltjes transform:

$$\hat{\Phi}(z) = \hat{\Psi}(z) + \hat{\Phi}(z)\hat{F}(z) \quad \implies \quad \hat{\Phi}(z) = \frac{\hat{\Psi}(z)}{1 - \hat{\Phi}(z)} = \hat{\Psi}(z)\left(1 + \frac{\hat{\Phi}(z)}{1 - \hat{\Phi}(z)}\right) =: \hat{\Psi}(z) + \hat{\Phi}(z)\hat{m}(z)$$

where $\hat{m}(z) := \hat{\Phi}(z)/(1 - \hat{\Phi}(z))$. Hence, (88) has solution

$$\phi(t) = \psi(t) + \int_0^t \psi(t-x) dm(x), \qquad m(t) = \mathcal{L}^{-1}\left\{\frac{\hat{\Phi}(z)}{1 - \hat{\Phi}(z)}\right\}$$

Therefore, (87) becomes

$$\begin{aligned}
\mathbb{P}(Y(t) > y) &= 1 - F(t+y) + \int_0^t (1 - F(t-x+y))\, dm(x) \\
\implies \mathbb{P}(Y(t) \leq y) &= F(t+y) - \int_0^t (1 - F(t-x+y))\, dm(x)
\end{aligned}$$

which is the desired result. ■

A corollary result follows by deriving the distribution of the age $Z(t)$, which is given by

$$\mathbb{P}(Z(t) \geq z) = \begin{cases} 0 & \text{if } z > t \\ \mathbb{P}(Y(t-z) > z) & \text{if } z \leq t \end{cases} \tag{89}$$

This is because $Z(t) \leq z$ iff there are no arrivals in the interval of time $(t-z, z]$.

3.4 Examples of Applications

Example 11 (Gambling Based on Two Outcomes)**.** In this example, we illustrate an application of Lemma 5. Suppose a gambler is playing a game in which she wins a bet (denoted with event W) with probability p and loses (denoted with event L) with probability $q := 1 - p$. She stops playing the first time she wins k consecutive bets, denoted by event W^k. Thus, we have the expected duration of the game to be

$$\mathbb{E}[T(W^k)] = \sum_{i=1}^{k} \frac{1}{p^i}$$

We are interested in her expected total winning amount and the expected number of bets she has won. In order to determine either quantity, we only need to define the reward function accordingly. For the gambler's expected total winnings, define:

$$R_i := \begin{cases} +1 & \text{if } i\text{th bet won} \\ -1 & \text{if } i\text{th bet lost} \end{cases}$$

Then by Wald's equation, the gambler's expected total winnings is given by:

$$\mathbb{E}[R] = \mathbb{E}\left[\sum_{i=1}^{T(W^k)} R_i\right] = (p-q)\mathbb{E}[T(W^k)] = (p-q)\sum_{i=1}^{k} \frac{1}{p^i}$$

Now instead, define:

$$R_i := \begin{cases} +1 & \text{if } i\text{th bet won} \\ 0 & \text{if } i\text{th bet lost} \end{cases}$$

Then by Wald's equation, the gambler's expected number of bets won is given by:

$$\mathbb{E}[R] = \mathbb{E}\left[\sum_{i=1}^{T(W^k)} R_i\right] = p\mathbb{E}[T(W^k)] = p\sum_{i=1}^{k} \frac{1}{p^i} = \sum_{i=0}^{k-1} \frac{1}{p^i}$$

□

Example 12 (Shuffling Cards)**.** Consider drawing, with replacement, one card at a time from a standard deck of cards. We are interested in the expected number N of draws until four consecutive cards come from the same suite. Since there are 4 suites, the probability of drawing a card from a specific suite is $p = \frac{1}{4}$. If we define a renewal process with renewals being the event that four consecutive cards come from the same suite, then such a process is delayed since the distribution of the first interarrival time is different from the others.

We will first compute the expected number of draws between consecutive renewals because it is easier to compute than N. Denote $\mathbb{E}[N_k]$ to be the expected number of additional draws needed, assuming $k \in \{1, 2, 3\}$ cards of the same suite have already appeared. Then:

$$\mathbb{E}[N_1] = \begin{cases} 1 + \mathbb{E}[N_1] \text{ w.p. } \frac{3}{4} \\ 1 + \mathbb{E}[N_2] \text{ w.p. } \frac{1}{4} \end{cases}, \quad \mathbb{E}[N_2] = \begin{cases} 1 + \mathbb{E}[N_1] \text{ w.p. } \frac{3}{4} \\ 1 + \mathbb{E}[N_3] \text{ w.p. } \frac{1}{4} \end{cases}, \quad \mathbb{E}[N_3] = \begin{cases} 1 + \mathbb{E}[N_1] \text{ w.p. } \frac{3}{4} \\ 1 \text{ w.p. } \frac{1}{4} \end{cases}$$

where in all three cases, we revert to the N_1 case because any new card which is not of the same suite as the previous card marks the beginning of a new renewal automatically; we do not care which specific suite we need to obtain four consecutive cards of.

Solving this system of equations is straightforward:

$$\mathbb{E}[N_1] = \frac{1}{\left(\frac{1}{4}\right)^3} = 64$$

Now, note that the relationship between N_1 and N is as follows:

$$\mathbb{E}[N_1] = \frac{1}{4}(1) + \frac{3}{4}\left(1 + (\mathbb{E}[N] - 1)\right)$$

where the first term comes from the case where the next card is of the same suite (with probability $\frac{1}{4}$) and the second term arises because the next card of a different suite starts a new renewal, and we subtract the first card to remove double-counting. Substituting, we get:

$$\mathbb{E}[N] = 85$$

and so it takes an average of 85 cards to obtain four consecutive cards of the same suite. □

Example 13 (Trucks Shipping Packages)**.** Suppose packages arrive at a facility according to Poisson(λ) arrivals, and trucks arrive with interarrival times distributed according to a nonlattice F with mean μ_F. Denote $N(t)$ to be the number of packages waiting to be picked up at time t. Note that $\{N(t), t \geq 0\}$ can then be interpreted as a regenerative process because every time a truck arrives, the number of packages waiting at the facility return to 0.

We are interested in computing the long-run proportion of time in a cycle during which a facility has i packages waiting. Denote the length of the nth cycle to be C_n and let $C_{n,i}$ to be the duration of time in the cycle for which the facility has i packages. Note that by Theorem 10:

$$\lim_{t\to\infty} \mathbb{P}(N(t) = i) = \frac{\mathbb{E}[C_{n,i}]}{\mathbb{E}[C_n]} = \frac{\mathbb{E}[C_{1,i}]}{\mu_F} \tag{90}$$

where the expected length of a cycle is μ_F because the process $N(t)$ restarts every time a truck arrives.

The numerator expression is simplified as

$$\mathbb{E}[C_{n,i}] = \mathbb{E}[C_{n,i}|N(T_n^{(T)}) - N(T_{n-1}^{(T)}) \geq i]\mathbb{P}(N(T_n^{(T)}) - N(T_{n-1}^{(T)}) \geq i) \text{ where } T_n^{(T)} \text{ is arrival of } n\text{th truck}$$

$$
\begin{aligned}
&= \mathbb{E}[C_{1,i}|N(T_1^{(T)}) \geq i]\mathbb{P}(N(T_1^{(T)}) \geq i) \text{ by stationarity} \\
&= \int_0^\infty \mathbb{E}[C_{1,i}|N(s) \geq i]\mathbb{P}(N(s) \geq i)\mathbb{P}(T^{(T)} = s)ds \\
&= \int_0^\infty \sum_{k=i}^\infty \mathbb{E}[C_{1,i}|N(s) = k]\mathbb{P}(N(s) = k)\mathbb{P}(T^{(T)} = s)ds \\
&= \int_0^\infty \sum_{k=i}^\infty \mathbb{E}[C_{1,i}|N(s) = k]\frac{(\lambda s)^k}{k!}e^{-\lambda s}dF(s)
\end{aligned}
$$

The interpretation of $\mathbb{E}[C_{1,i}|N(s) = k]$ is the extent of time there are i packages in the facility given a truck arrives in between the k and the $(k+1)$th arrival. Note that the probability of choosing the interval corresponding to the duration of time for which there are i packages at the facility, i.e. the proportion of the duration from the 0 to k during which there are i packages at the facility, is simply $1/(k+1)$, and since the duration itself is given to be s, we have $\mathbb{E}[C_{1,i}|N(s) = k] = s/(k+1)$. Substituting:

$$
\mathbb{E}[C_{1,i}] = \int_0^\infty \sum_{k=i}^\infty \frac{s}{k+1} \cdot \frac{(\lambda s)^k}{k!}e^{-\lambda s}dF(s)
$$

Now suppose instead, there is a bit of a delay in producing the first package immediately after a truck has left the facility, and this time is distributed according to H while all other interarrival times are distributed i.i.d. according to G. We are interested in computing the same quantity for this setting.

We still have (90). The length of the cycle is $\mathbb{E}[C_n] = \mu_F$ for any $n \in \mathbb{N}$. The numerator of (90) simplifies similarly to the non-delayed case:

$$
\mathbb{E}[C_{1,i}] = \int_0^\infty \sum_{k=0}^\infty \mathbb{E}[C_{1,i}|N(s) = k]\mathbb{P}(N(s) = k)dF(s) \tag{91}
$$

Following the uniform argument presented for the non-delayed case, we have

$$
\mathbb{E}[C_{1,i}|N(s) = k] = \begin{cases} \frac{s\mu_H}{\mu_H + k\mu_G)} & \text{if } i = 0 \\ \frac{s\mu_G}{\mu_H + k\mu_G)} & \text{if } i \geq 1 \end{cases}
$$

because the long-run average time it takes to accumulate $k \geq 0$ packages in the facility is $\mu_H + k\mu_G$, and the given duration of time to reach k packages is s. Note that if $k = 0$, $\mathbb{P}(N(s) = 0) = \mathbb{P}(T_0 \leq s) - \mathbb{P}(T_1 \leq s) = 1 - G(s)$. To compute $\mathbb{P}(N(s) = k)$ for $k \geq 1$, we condition on the first arrival time:

$$
\begin{aligned}
\mathbb{P}(N(s) = k) &= \mathbb{P}(T_k \leq s) - \mathbb{P}(T_{k+1} \leq s) \\
&= \mathbb{P}(X_1 + X_2 + \cdots + X_k \leq s) + \mathbb{P}(X_1 + X_2 + \cdots + X_{k+1} \leq s) \\
&= \int_0^s \mathbb{P}(X_2 + \cdots + X_k \leq s - r|X_1 = r)\mathbb{P}(X_1 = r)dr \\
&\qquad - \int_0^s \mathbb{P}(X_2 + \cdots + X_{k+1} \leq s - r|X_1 = r)\mathbb{P}(X_1 = r)dr \\
&= \int_0^s G_{k-1}(s-r)dH(r) - \int_0^s G_k(s-r)dH(r)
\end{aligned}
$$

Putting everything together:

$$
\lim_{t\to\infty} \mathbb{P}(N(t) = i) = \frac{1}{\mu_F}\int_0^\infty \sum_{k=0}^\infty \frac{s\mu_H}{\mu_H + k\mu_G} \cdot \int_0^s (G_{k-1}(s-r) - G_k(s-r))dH(r)dF(s) \tag{92}
$$

$\square$

Regenerative renewal processes also have connections to Markov chain formulations, as illustrated by the following example.

Example 14 (Rainy Days)**.** Suppose an individual owns m umbrellas. Whenever it is raining, he uses an umbrella on his way from home to work and vice versa, provided there is at least one at his current location. The probability of a day being rainy is p, independent of the past days. We are interested in the long-run proportion of time the individual walks home wet.

We can define the state X_n to be the number of umbrellas at the individual's current location (either his office or his home) at time n. Then the transition probabilities are

$$p_{0,m} = 1,\ p_{i,j} = \begin{cases} q & \text{if } j = m - i \\ p & \text{if } j = m - i + 1 \\ 0 & \text{else} \end{cases},\ \text{for } i \neq 0$$

where $q := 1 - p$, and the first equality is true because if there are no umbrellas at the person's current location at time n, there will always be m umbrellas at the other location (his next location at time $n+1$).

Solving the stationary equation $\pi = \pi P$ using these transition probabilities yields

$$\begin{aligned} \pi_0 &= \pi_r q \\ \pi_1 &= \pi_{m-1} q + \pi_m p \\ &\vdots \\ \pi_j &= \pi_{m-j} q + \pi_{m-j+1} p \\ &\vdots \\ \pi_m &= \pi_0 + \pi_1 p \end{aligned}$$

which has the solution

$$\pi_k = \begin{cases} \frac{1}{m+q} & \text{if } k \geq 1 \\ \frac{q}{r+q} & \text{if } k = 0 \end{cases}$$

However, since there is only probability p of being a rainy day, the proportion the individual spends being wet is given by $p\pi_0$. □

Example 15 (Telemarketer's Calling Strategy)**.** A telemarketer is selling a product over telephone. The probability that she will be able to sell the product to a customer at time t is given by

$$P(t) := \begin{cases} 2t^2 - t^3 & \text{if } t \leq 0.8 \text{ hours} \\ 0.87 & \text{if } t > 0.8 \text{ hours} \end{cases} \tag{93}$$

For each call, if the telemarketer either 1) manages to persuade the customer to buy the product or 2) does not manage to persuade the customer within $\tau \in \mathbb{R}^+$ hours, she hangs up and starts dialing the next customer immediately.

We are interested in determining the optimal value of τ which maximizes the long-term number of sold products. This requires us to define the reward $R_i = 1$ product sold if the telemarketer manages to persuade the ith customer, and 0 otherwise. Further define T_i to be the time spent on the ith customer.

We consider two cases: 1) $\tau \leq 0.8$ and 2) $\tau > 0.8$.

- Case 1: $\tau \leq 0.8$. Then we can compute the expected time of the call with the ith customer as:

$$\mathbb{E}[T_i] = \int_0^\tau \mathbb{P}(T_i > t)dt = \int_0^\tau \left(1 - (2t^2 - t^3)\right) dt = \tau - \frac{2}{3}\tau^3 + \frac{1}{4}\tau^4$$

Furthermore, the expected reward obtained from customer i is

$$\mathbb{E}[R_i] = 1 \cdot \mathbb{P}(R_i = 1) = 2\tau^2 - \tau^3$$

By Lemma 5, we have

$$\lim_{t\to\infty} \frac{R(t)}{t} = \frac{\mathbb{E}[R_i]}{\mathbb{E}[T_i]} = \frac{2\tau^2 - \tau^3}{\tau - \frac{2}{3}\tau^3 + \frac{1}{4}\tau^4} = \frac{24\tau - 12\tau^2}{12 - 8\tau^2 + 3\tau^3}$$

To obtain the maximizing value of τ, take the derivative and set it equal to zero. The numerator of the expression becomes:

$$0 =: (24 - 24\tau)(12 - 8\tau^2 + 3\tau^3) - (-16\tau + 9\tau^2)(24\tau - 12\tau^2) = 24 - 24\tau + 16\tau^2 - 12\tau^3 + 3\tau^4$$

- Case 2: $\tau > 0.8$. Similar to the $\tau \leq 0.8$ case,

$$\mathbb{E}[T_i] = \int_0^\tau tP(t)dt = \int_0^{0.8} t(2t^2 - t^3)dt + \int_{0.8}^\tau t(0.87)dt =$$

□

Example 16 (Computing the Renewal Functions of Different Processes)**.** In this example, we demonstrate computation of the renewal function for various types of renewal processes.

1. Consider the renewal process $N_e(t)$ constructed by taking only the even arrivals of a homogeneous Poisson process with intensity λ.
2. Consider the renewal process with interarrival times X_i distributed as follows:

$$X_i := \begin{cases} 0 & \text{w.p. } p \\ s & \text{w.p. } (1-p)\lambda e^{-\lambda s},\ s > 0 \end{cases}$$

3. Consider the delayed renewal process where $X_1 \sim \mathcal{U}[0, a]$ for $a \in \mathbb{R}^+$, and $X_2, X_3, \cdots \sim \text{Exp}(\lambda)$.

□

Example 17 (A Small Town Bank)**.** In a small town, there is a bank with only one counter, and customers arrive with interarrival times X_i distributed $\text{Exp}(\lambda)$. However, a potential customer arriving to the bank enters the bank only if there are no other customers at the counter; otherwise, they leave and never return. Assume that the service times S_i are distributed $\text{Exp}(\mu)$.

Note that when the bank is opened, there are no customers. Hence, the arrivals form a delayed renewal process. First, we compute the long-term proportion of the customers which enter the bank, among all customers who arrive. This is equivalent to the the long-run probability that there is no person at the counter. Because the service times are exponentially distributed with mean μ (cdf denoted by $G(t)$), its corresponding Laplace transform is given by $\hat{G}(z) := \mu/(z + \mu)$. Therefore, the probability is given by:

$$P := \mathbb{P}(S_i < X_i) = \int_0^\infty \mathbb{P}(S_i < a)\mathbb{P}(X_i = a)da$$

$$= \int_0^\infty \left(1 - e^{-\lambda a}\right) \cdot \mu e^{-\mu a} da = \frac{\lambda}{\lambda + \mu}$$

Note that customers who enter the bank after the first arrival abide by a thinned Poisson process with intensity $\lambda\mu/(\lambda + \mu)$. Denoting the corresponding exponential interarrival distribution as $F(t)$, we get Laplace transform

$$\hat{F}(z) = \frac{\frac{\lambda\mu}{\lambda+\mu}}{z + \frac{\lambda\mu}{\lambda+\mu}}$$

We can compute the corresponding renewal function $m(t)$ using (65).

$$\hat{M}(z) = \frac{\frac{\lambda}{z+\lambda}}{1 - \frac{\frac{\lambda\mu}{\lambda+\mu}}{z+\frac{\lambda\mu}{\lambda+\mu}}} = \frac{\lambda\left((\lambda+\mu)z + \lambda\mu\right)}{(\lambda+\mu)z(z+\lambda)} = \frac{\lambda}{\lambda+\mu}\left(\frac{\mu}{z} + \frac{\lambda}{z+\lambda}\right)$$

$$\implies m(t) = \frac{\lambda}{\lambda+\mu}\left(\mu t + \int_0^t e^{-\lambda s} ds\right) = \frac{\lambda\mu}{\lambda+\mu} t + \frac{1}{\lambda+\mu}\left(1 - e^{-\lambda t}\right)$$

□

Example 18 (Deploying the Police)**.** A policeman is responsible for the surveillance of two regions in a city, A and B. While the policeman is surveying region A, his supervisor arrives to meet him according to a Poisson process with intensity λ_A meetings per month. At each meeting, there is a p_A chance the supervisor will tell the policeman to relocate to region B. Similarly, while the policeman is in region B, his supervisor meets with him according to a Poisson process with intensity λ_B meetings per month, with a p_B chance of being told to relocate to region A. Assume that the decisions on relocation are independent of each other as well as of the supervisor meeting times.

We want to compute the renewal function of the supervisor arrivals. First, note that they form a delayed renewal process. Since the policeman starts in region A, the interarrival time until his first supervisor meeting is exponentially-distributed with mean $1/\lambda_A$. All future interarrival times between supervisor visits are distributed conditionally based on region: with probability p_A, the interarrival time is $\text{Exp}(\lambda_A)$ and with probability p_B, the interarrival time is $\text{Exp}(\lambda_B)$. Hence, the supervisor meetings form a delayed renewal process with $G(t) = 1 - e^{-\lambda_A t}$, i.e. $\hat{G}(z) = \lambda_A/(z + \lambda_A)$, and $F(t) = p_A(1 - e^{-\lambda_A t}) + p_B(1 - \lambda_B t)$, i.e.

$$\hat{F}(z) = p_A \frac{\lambda_A}{z + \lambda_A} + p_B \frac{\lambda_B}{z + \lambda_B}$$

The Laplace-Stieltjes transform of the renewal function is given by (65):

$$\hat{M}(z) = \frac{\frac{\lambda_A}{z+\lambda_A}}{p_A\frac{\lambda_A}{z+\lambda_A} + p_B\frac{\lambda_B}{z+\lambda_B}} = \frac{\lambda_A(z+\lambda_B)}{p_A\lambda_A(z+\lambda_B) + p_B\lambda_B(z+\lambda_A)} = \frac{Az + B}{z + C} = A + \frac{B - CA}{z + C}$$

where we denote for simplicity

$$A := \frac{\lambda_A}{p_A\lambda_A + p_B\lambda_B}, \quad B := \frac{\lambda_A\lambda_B}{p_A\lambda_A + p_B\lambda_B}, \quad C := \frac{\lambda_A\lambda_B(p_A + p_B)}{p_A\lambda_A + p_B\lambda_B}$$

The inverse Laplace transform yields the renewal function

$$m(t) = A + \frac{B - CA}{C} e^{-Ct}$$

□

Example 19 (Computer with Multiple Parts)**.** Suppose that a computer has $M \in \mathbb{N}$ critical parts, all of which need to be functioning simultaneously in order for the computer to work. Each part $k \in \{1, \cdots, M\}$ has a lifetime (in weeks) which is distributed as $\text{Exp}(\lambda_k)$, independently of the lifetimes of the other parts. Each part k also takes $G_k(\mu_k)$ time (in weeks) to repair once it fails, where G_k is some probability distribution with mean μ_k. Finally, each part k costs R_k dollars to repair.

Note that the computer's state forms an alternating renewal process with two states: up, when all the parts are working, and down, when at least one of the parts has failed. Because all parts are exponentially-distributed, the time taken until at least one of the parts fails is distributed exponentially with mean $\sum_{k=1}^{M} \lambda_k$. Hence, the mean time to a computer failure is $\mathbb{E}[W] = 1/(\sum_{k=1}^{M} \lambda_k)$ weeks, where W denotes the duration of a single working period. Moreover, the probability that part $i \in \{1, \cdots, M\}$ is the first one to fail is given by $p_i := \lambda_i/(\sum_{k=1}^{M} \lambda_k)$.

Given these basic results, we are now interested in 1) the long-run proportion P_W of time the computer is working, and 2) the long-run average cost r of replacement parts per week. First, note that a cycle of the renewal process is equivalent to a computer failure period followed by a computer working period. Hence, by Theorem 10, we get that the long-run proportion P_W of time the computer spends working is

$$P_W = \frac{\mathbb{E}[W]}{\mathbb{E}[W] + \mathbb{E}[F]}$$

where F denotes the duration of a single failure period. $\mathbb{E}[W]$ has been determined above and $\mathbb{E}[F]$ can be computed by conditioning on which part failed:

$$\mathbb{E}[F] = \sum_{k=1}^{M} p_k \mathbb{E}[F|\text{part } k \text{ failed}] = \sum_{k=1}^{M} p_k \mu_k$$

and so:

$$P_W = \frac{1/(\sum_{k=1}^{M} \lambda_k)}{1/(\sum_{k=1}^{M} \lambda_k) + \sum_{k=1}^{M} p_k \mu_k} = \frac{1}{1 + (\sum_{k=1}^{M} \lambda_k)(\sum_{k=1}^{M} p_k \mu_k)}$$

The cost of replacement parts per week can be determined in a similar way. We combine the results of Theorem 10 and Lemma 5 to get

$$r := \frac{\mathbb{E}[R]}{\mathbb{E}[W] + \mathbb{E}[F]} = \frac{\mathbb{E}[R]}{1/(\sum_{k=1}^{M} \lambda_k) + \sum_{k=1}^{M} p_k \mu_k}$$

where $\mathbb{E}[R]$ can be computed similarly by taking conditional expectation on the failed part

$$\mathbb{E}[R] = \sum_{k=1}^{M} p_k \mathbb{E}[R|\text{part } k \text{ failed}] = \sum_{k=1}^{M} p_k R_k$$

and so, overall:

$$r = \frac{\sum_{k=1}^{M} p_k R_k}{1/(\sum_{k=1}^{M} \lambda_k) + \sum_{k=1}^{M} p_k \mu_k} = \frac{(\sum_{k=1}^{M} \lambda_k)(\sum_{k=1}^{M} p_k R_k)}{1 + (\sum_{k=1}^{M} \lambda_k)(\sum_{k=1}^{M} p_k \mu_k)}$$

□

Example 20 (Geiger Counter)**.** A Geiger counter is a device which detects the presence of radioactive particles in a system. Each time the counter detects a radioactive particle, it requires a certain interval of time to completely register it in the count. During these "dead" periods, the counter fails to detect arriving particles.

We denote $N(t)$ to represent the renewal process of arriving particles, with i.i.d. interarrival times $\{X_n\}_{n=1}^{\infty}$ distributed with cdf F and corresponding renewal function $m(t)$. Denote L_n to be the length of the dead period corresponding to the nth detected arrival. Assume $\{L_n\}_{n=0}^{\infty}$ are i.i.d. distributed according to cdf G and L_0 begins at time 0, i.e. the counter starts dead. Denote $\tilde{N}(t)$ to be the stochastic process counting the number of detected arrivals over time, with interarrival times $\tilde{X}_{n+1} := L_n + D_n$, where D_n is the excess lifetime of $\tilde{N}$ at the end of the nth dead period. Note that $\tilde{N}$ is not necessarily a renewal process because the interarrival times $\{\tilde{X}_n\}_{n=1}^{\infty}$ are not i.i.d.

We are interested in computing the distribution of $\tilde{X}_1$, i.e. the time elapsed until the first detected particle. By conditioning on the length of the initial dead period:

$$\begin{aligned}
\mathbb{P}(\tilde{X}_1 \leq t) &= \int_0^x \mathbb{P}(\tilde{X}_1 \leq t | L_0 = s) dG(s) \\
&= \int_0^x \mathbb{P}(L_0 + D_0 \leq t | L_0 = s) dG(s) \\
&= \int_0^x \mathbb{P}(D_0 \leq t - s) dG(s) = \int_0^x \mathbb{P}(Y(s) \leq t - s) dG(s) \qquad (94)
\end{aligned}$$

where we interpret $Y(s)$ as the residual lifetime at time $s \in [0, t)$.

Note that by Theorem 12, we have

$$\mathbb{P}(Y(s) \leq t - s) = F(t) - \int_0^s (1 - F(t - r))\, dm(r) \qquad (95)$$

and we can alternatively write (57) as

$$\begin{aligned}
m(t) = F(t) + \int_0^t F(t-s) dm(s) &\implies \int_0^t dm(s) - \int_0^t F(t-s) dm(s) = F(t) \\
&\implies F(t) = \int_0^t (1 - F(t-s))\, dm(s)
\end{aligned}$$

Substituting this back into (95) and substituting the result into (94) yields:

$$\begin{aligned}
\mathbb{P}(\tilde{X}_1 \leq t) &= \int_0^x \left(\int_s^t (1 - F(t-r))\, dm(r) \right) dG(s) \\
&= \int_0^t \int_0^r (1 - F(t-r))\, dG(s) dm(r) \quad \text{by change of variables} \\
&= \int_0^t (1 - F(t-r))\, G(r) dm(r) \qquad (96)
\end{aligned}$$

Given the mean length of dead periods $\mathbb{E}[L]$ and the count $\tilde{N}(t)$ of detected particles, we can create an estimate of the unknown emission rate λ of radioactive particles First, the density of observed particles is approximated empirically by $\gamma_t \approx \tilde{N}(t)/t$. Using the elementary renewal theorem, we have

$$\gamma_t \approx \frac{\tilde{N}(t)}{t} \approx \frac{1}{\mathbb{E}[\tilde{X}_1]} = \frac{1}{\mathbb{E}[L] + \frac{1}{\lambda}}$$

for large t. Hence: $\lambda \approx \gamma_t / (1 - \gamma_t \mathbb{E}[L])$. ☐

References

[1] S. Han and S.-J. Chung, "Incremental nonlinear stability analysis for stochastic systems perturbed by Lévy noise," *International Journal of Robust and Nonlinear Control*, Submitted 2021.

[2] D. R. Cox, *Renewal Theory.* London: John Wiley & Sons Inc., 1962.

[3] S. Ross, *Introduction to Probability Models*, 10th ed. San Diego, CA, USA: Academic Press, 2010.

[4] ——, *Stochastic Processes*, 2nd ed. New York, USA: John Wiley & Sons Inc., 1996.

[5] G. Grimmett and D. Stirzaker, *Probability and random processes*, 3rd ed. Oxford; New York: Oxford University Press, 2001.

[6] S.-Y. R. Li, "A Martingale Approach to the Study of Occurrence of Sequence Patterns in Repeated Experiments," *The Annals of Probability*, vol. 8, no. 6, pp. 1171 – 1176, 1980.