6.262 Discrete Stochastic Processes, Spring 2009
Problem Set 5 — Solutions
due: Wednesday, March 11, 2009

Exercise 3.9

We don't need Wald's equality to compute the number of Bernoulli trials up to and including the k^{th} success. (How would we do this in one line using only 6.041 material? Think of the linearity of the expectation.) However, the point is to get you to understand Wald's equality on a simple example before we start applying it to more complex situations.

Let $\mathrm{P}(X_i = 1) = p$ and $\mathrm{P}(X_i = 0) = 1-p$, where the event $\{X_i = 1\}$ corresponds to a "success" on the i^{th} trial. Let $J = \min\{n \mid S_n \geq k\}$ and notice that J is a stopping rule, as the event

$$\{J \geq n\} = \{\text{at most } k-1 \text{ sucesses in first } n-1 \text{ trials}\}$$

only depends on $X_1, \ldots, X_{n-1}$ and therefore not on $X_n, X_{n+1}, \ldots$.[1] By Wald's equality, we have that $\mathrm{E}(S_J) = \mathrm{E}(J)\mathrm{E}(X_1)$. Note that $E(S_J) = k$, since S_n changes in increments of 1. Also, $E(X_1) = p$. It follows that

$$\mathrm{E}(J) = \frac{\mathrm{E}(S_J)}{E(X_1)} = \frac{k}{p}.$$

Exercise 3.10

a Let $S_n = X_1 + \ldots + X_n$ and let $J = \min\{n \mid S_n \leq -d\}$. We'd like to show that J is finite with probability 1. As shown in Problem Set 3 (Exercise 3.1-b), it suffices to prove that $\lim_{n\to\infty} P(J \geq n) = 0$ or $\lim_{n\to\infty} P(J > n) = 0$ (you may wish to revisit the solution of Exercise 3.1 and convince yourself that either condition will do). Now note that

$$P(J > n) = \mathrm{P}(S_1 > -d, S_2 > -d, \ldots, S_n > -d) \leq \mathrm{P}(S_n > -d)$$

and that

$$\mathrm{P}(S_n > -d) = \mathrm{P}(S_n - \mathrm{E}(X) > -d - \mathrm{E}(X)) = \mathrm{P}(S_n/n - \mathrm{E}(X) > -d/n - \mathrm{E}(X)).$$

Since $\mathrm{E}(X) < 0$, we have that $-d/n - \mathrm{E}(X) \geq -\mathrm{E}(X)/2 > 0$ for n large enough. Thus, for n large enough,

$$\begin{aligned}\mathrm{P}(S_n/n - \mathrm{E}(X) > -d/n - \mathrm{E}(X)) &\leq \mathrm{P}(S_n/n - \mathrm{E}(X) > -\mathrm{E}(X)/2)\\ &\leq \mathrm{P}(S_n/n - \mathrm{E}(X) \geq -\mathrm{E}(X)/2)\\ &\leq \mathrm{P}(|S_n/n - \mathrm{E}(X)| \geq -\mathrm{E}(X)/2).\end{aligned}$$

[1] To show that J is a stopping rule, we also need to show that J is finite with probability 1. However, in this case, the easiest way to show that is by computing $\mathrm{E}(J)$ directly, without Wald's equality, and showing it's finite. (Why would that imply that J is finite with probability 1?) Again, keep in mind that the point of this exercise is to illustrate the mechanics of Wald's equality, not to convince you of its power.

But, since X_i are IID with finite mean, the Weak Law tells us that

$$\lim_{n\to\infty} \mathrm{P}(|S_n/n - \mathrm{E}(X)| \geq \epsilon) = 0$$

for all $\epsilon > 0$. Letting $\epsilon = -\mathrm{E}(X)/2$, it follows that

$$\lim_{n\to\infty} P(J > n) \leq \lim_{n\to\infty} \mathrm{P}(|S_n/n - \mathrm{E}(X)| \geq \epsilon) = 0,$$

which is what we wanted to show.

This exercise shows that even though our goal is to prove that J is finite *with probability 1*, we can approach the problem by showing that another statement holds *in probability*, which is frequently easier.

b To find $E(J)$, we use Wald's identity. Part a) shows that J is finite w.p.1. Furthermore,

$$P(J \geq n) = \mathrm{P}(S_1 > -d, S_2 > -d, \dots, S_{n-1} > -d)$$

and so J does not depend on $X_n, X_{n+1}, \dots$. It follows that J is a stopping rule. Since $S_J = -d$ and $\mathrm{E}(X) < 0$, Wald's equality yields

$$E(J) = \frac{-d}{\mathrm{E}(X)}.$$

Exercise 3.12

a Since $J = \min\{n \mid S_n \leq B \text{ or } S_n \geq A\}$, the event

$$\{J \geq n\} = \min\{n \mid S_n \leq B \text{ or } S_n \geq A\} = \{S_1 \in (B, A), \dots, S_{n-1} \in (B, A)\}$$

does not depend on $X_n, X_{n+1}, \dots$. To show that J is a stopping rule, it remains to show that J is not defective. As shown in Problem Set 3 (Exercise 3.1-b), it suffices to prove that $\lim_{n\to\infty} P(J \geq n) = 0$. In other words, we'd like to show that the probability that S_n remains within (A, B) converges to zero with increasing n. One way to do this is through reasoning similar to that of Problem 3 on Problem Set 1. Instead, the following proof yields a useful inequality and gives us some practice with manipulating J.

Let $m = A - B - 1$. Since $X_i \in \{-1, 0, 1\}$, we have that $\mathrm{P}(S_{n+m} \geq A \mid S_n \in (B, A)) \geq (\mathrm{P}(X = 1))^m$. (Why? It may be helpful to notice that the equality holds only if $S_n = B + 1$.) Letting $\epsilon = (\mathrm{P}(X = 1))^m > 0$, it follows that for any $n \geq 0$,

$$\begin{aligned} \mathrm{P}(S_{n+m-1} \in (B, A) \mid S_n - 1 \in (B, A)) &= 1 - \mathrm{P}(S_{n+m-1} \leq B \text{ or } S_{n+m-1} \geq A \mid S_{n-1} \in (B, A)) \\ &\leq 1 - \mathrm{P}(S_{n+m-1} \geq A \mid S_{n-1} \in (B, A)) \leq 1 - \epsilon. \end{aligned}$$

But notice that

$$\mathrm{P}(S_{n+m-1} \in (B, A) \mid S_{n-1} \in (B, A)) = \mathrm{P}(S_{n+m-1} \in (B, A) \mid S_{n-1}, \dots, S_2 \in (B, A), S_1 \in (B, A))$$

since additionally knowing the values of S_i for $i < n-1$ provides no added information on S_{n+m-1}. (This type of processes will be studied in Chapter 7 in far greater detail. In fact, the type of machinery that we will develop there will render the current proof trivial.) Also notice that $\{S_{n-1}, \ldots, S_1 \in (B, A)\} = \{J \geq n\}$ (why?).

It follows that $\mathrm{P}(S_{n+m-1} \in (B, A) \mid J \geq n) \leq 1 - \epsilon$. But $\{J \geq n+m \mid J \geq n\} \subset \{S_{n+m-1} \in (B, A) \mid J \geq n\}$, so $\mathrm{P}(J \geq n+m \mid J \geq n) \leq \mathrm{P}(S_{n+m-1} \in (B, A) \mid J \geq n)$. Therefore,

$$1 - \epsilon \geq \mathrm{P}(J \geq n+m \mid J \geq n) = \frac{\mathrm{P}(J \geq n+m \mid J \geq n)}{\mathrm{P}(J \geq n)} = \frac{\mathrm{P}(J \geq n+m)}{\mathrm{P}(J \geq n)},$$

from which it follows that

$$\mathrm{P}(J \geq n+m) \leq (1-\epsilon)\mathrm{P}(J \geq n) \qquad (\spadesuit).$$

Since $\mathrm{P}(J \geq 0) = 1$, setting $n = 0$ yields $\mathrm{P}(J \geq m) \leq (1 - \epsilon)$. But, for any $k \geq 1$, iterating ($\spadesuit$) yields $\mathrm{P}(J \geq km) \leq (1-\epsilon)^k$. Recalling that $\epsilon = (\mathrm{P}(X = 1))^m$ and letting $m = 1$, we have

$$\mathrm{P}(J \geq k) \leq (1 - \epsilon_0)^k,$$

where $\epsilon_0 = \mathrm{P}(X = 1)$. It then follows that

$$\lim_{k \to \infty} \mathrm{P}(J \geq k) = 0,$$

and we're done.

b Since A and B are integer-valued, since S_n changes in increments of ± 1, and since we stop the moment we reach either boundary, we have that $S_J \in \{A, B\}$.

c By the Total Probability Theorem, $\mathrm{E}(S_J) = \mathrm{P}(S_J \geq A)\mathrm{E}(S_J \mid S_J \geq A) + \mathrm{P}(S_J \leq B)\mathrm{E}(S_J \mid S_J \leq B)$. By b), $\mathrm{E}(S_J \mid S_J \geq A) = A$ and $\mathrm{E}(S_J \mid S_J \leq B) = B$. Letting $p = \mathrm{P}(S_J \geq A)$, we have that $\mathrm{P}(S_J \geq A) = 1 - p$ since $S_J \in \{A, B\}$. It follows that

$$\mathrm{E}(S_J) = pA + (1-p)B.$$

d Since J is a stopping rule, we can apply Wald's equality to yield $\mathrm{E}(S_J) = \mathrm{E}(J)\mathrm{E}(X_1)$, where $\mathrm{E}(X_1) = 0$. From part a), $\mathrm{P}(J \geq k) \leq (1 - \epsilon_0)^k$, and thus

$$E(J) = \sum_{k=1}^{\infty} \mathrm{P}(J > k) \leq \sum_{k=1}^{\infty} \mathrm{P}(J \geq k) \leq \sum_{k=1}^{\infty} (1-\epsilon_0)^k < \infty$$

since $0 < 1 - \epsilon_0 < 1$. Then, by Wald's equality, we have $\mathrm{E}(S_J) = \mathrm{E}(X_1)\mathrm{E}(J) = 0\mathrm{E}(J) = 0$. By c), it now follows that $p = \mathrm{P}(S_J \geq A) = -B/(A - B)$.

e Letting $A = 1$, we stop when S_j reaches 1 or falls below some $B < 0$. By d), we have $p = \mathrm{P}(S_J \geq 1) = -B/(1 - B)$ and

$$\mathrm{E}(S_J) = p + (1-p)B = \frac{-B}{1-B}1 + \left(1 + \frac{B}{1-B}\right)B = \frac{-B}{1-B}1 + \frac{1}{1-B}B = 0.$$

It follows that

$$\lim_{B\to-\infty} \mathrm{P}(S_J \geq 1) = 1 \quad \text{and} \quad \lim_{B\to-\infty} \mathrm{E}(S_J) = 0.$$

The above calculations shed a deeper insight into the "bizarre" coin-tossing example of Lecture 8. Since for any finite interval (A, B), the probability that S_j is contained within that interval converges to zero as $j \to \infty$. Expanding that interval by letting $B \to \infty$, we make S_j less and less likely to fall below B, and we obtain that the probability that S_j exceeds A must therefore be equal to 1 in the limit. However, as the probability of S_j eventually falling below B goes as $1/|B|$, the term still has a significant effect on the mean.

Furthermore, note that making A be any positive integer (say 10^{15}) and letting $B \to -\infty$, it follows that the probability of eventually reaching wealth A converges to 1. In class, we've shown that the scheme is not practical as the number of games one is expected to play is infinite. However, now we can also see that in order to achieve gain A, we must be able to deal with the possibility of seeing our wealth become arbitrarily low in the process. If we let our computers do the gambling, we can squeeze many millions of games within the space of a second, so the first issue may not be as problematic. (In other words, unless $J = \infty$ with probability 1, there is some positive probability that we may eventually earn the 10^{15}\$.) But the second issue is what makes the scheme a truly bad idea from the gambler's perspective.

Problem 4

Recall that a function $f : [0, \infty) \to \mathbb{R}$ that is piecewise continuous and of exponential order (i.e. such that there exists some $\alpha \in \mathbb{R}$ and $M \geq 0$ so that $|f(t)| \leq Me^{\alpha t}$ for all t greater or equal to some $t_0 \geq 0$) will have a Laplace transform L_f given by

$$L_f(r) = \int_0^\infty e^{-rt} f(t)\mathrm{d}t,$$

where any $r \in \mathbb{C}$ for which the above integral converges is said to belong to the *region of convergence* A associated with the Laplace transform of f.

a We're given that both $f(t)$ and $f'(t)$ have Laplace transforms that converge in some common region A. For $r \in A$, by definition of the Laplace transform, we have

$$L_{f'}(r) = \int_0^\infty f'(t)e^{-rt}\mathrm{d}t = \left[f(t)e^{-rt}\right]_0^\infty - \int_0^\infty (-r)e^{-rt} f(t)\mathrm{d}t$$

where the second equality follows from integration by parts. Now note that

$$-\int_0^\infty (-r)e^{-rt} f(t)dt = r\int_0^\infty e^{-rt} f(t)dt = rL_f(r),$$

where the last equality follows from the definition of the Laplace transform, since $r \in A$ and $f(t)$ has a Laplace transform that converges in A. Furthermore, $f(t)e^{-rt}\Big|_{t=0} = f(0)$ for any complex number r. It remains to evaluate $f(t)e^{-rt}\Big|_{t=\infty}$, which is shorthand for

$\lim_{t\to\infty} f(t)e^{-rt} = 0$ (remember the definition of the extended integral from calculus). Since f is of exponential order (recall the above definition of the Laplace transform), we have that for t large enough, $|f(t)| \leq Me^{\alpha t}$ for some $\alpha \in \mathbb{R}$. It follows that

$$\lim_{t\to\infty} |f(t)e^{-rt}| \leq \lim_{t\to\infty} Me^{\alpha t}e^{-rt} = 0$$

for all $r > \alpha$. But, $-|f(t)e^{-rt}| \leq f(t)e^{-rt} \leq |f(t)e^{-rt}|$. Taking limits, we obtain that $\lim_{t\to\infty} f(t)e^{-rt} = 0$. It follows that for $r \in A$,

$$L_{f'}(r) = -f(0) + rL_f(r),$$

which is what we wanted to show.

b If we want to compute $\lim_{r\to 0} rL_g(r)$, we first need to ensure that there exists some sequence $r_1, r_2, \ldots \in A$ such that $r \to 0$. (Can you see why the expression $\lim_{r\to 0} rL_g(r)$ does not make sense otherwise?) Here, we essentially allowed you to assume that $A \supset \{r \in \mathbb{C} \mid \mathrm{Re}(r) > 0\}$, that is, to assume that A contains the right half-plane, in which case any real sequence converging to 0 from the right will do.

(**Note 1:** It is not too hard to show that the above assumption actually holds true. We have that g is differentiable, which implies that g is continuous. The continuity of g in turn implies that g is bounded on any interval $[0, t_0]$, that is, that there exists some $M \in \mathbb{R}$ (which depends on the choice of t_0) such that $g(t) \leq M$ for $t \in [0, t_0]$. (This is another basic fact from calculus.) So far, we know that whatever $t_0 \in [0, \infty)$ we pick $\int_0^{t_0} g(t)e^{-rt}\mathrm{d}t \leq \int_0^{t_0} Me^{-rt}\mathrm{d}t < \infty$ for any $r \in \mathbb{C}$. So far, so good. Now consider the $\int_{t_0}^{\infty} g(t)e^{-rt}\mathrm{d}t$ term, and recall that $\lim_{t\to\infty} g(t) = g(\infty) \in \mathbb{R}$. By a similar argument as in the previous note, we let $\epsilon = |g(\infty)|/2$ and let t_0 be such that $g(t) \in (g(\infty) - \epsilon, g(\infty) + \epsilon)$ for all $t \geq t_0$. Supposing that $g(\infty) > 0$ (a similar argument holds if $g(\infty) < 0$), we have that $\int_{t_0}^{\infty} g(t)e^{-rt}\mathrm{d}t \leq \int_{t_0}^{\infty} (3g(\infty)/2)e^{-rt}\mathrm{d}t$, which is finite for any $r \in \mathbb{C}$ such that $\mathrm{Re}(r) > 0$. And we're done!)

Now, we can apply the differentiation theorem in part (a) to yield $rL_g(r) = L_{g'}(r) + g(0)$. Taking the limits, we have

$$\begin{aligned}
\lim_{r\to 0} rL_g(r) &= \lim_{r\to 0} L_{g'}(r) + g(0) = \lim_{r\to 0} \int_0^{\infty} g'(t)e^{-rt}dt + g(0) \\
&= \int_0^{\infty} \lim_{r\to 0} g'(t)e^{-rt}dt + g(0) \qquad \text{(see note below on this interchange)} \\
&= \int_0^{\infty} g'(t)dt + g(0) \\
&= \lim_{t\to\infty} g(t)
\end{aligned}$$

(**Note 2:** That we can move the limit inside of the integral in this case is a consequence of the Monotone Convergence Theorem, as $g'(t)e^{-rt}dt$ monotonically approaches a limit in decreasing r. However, this is not a course in integration, and as engineers dealing with Laplace transform, we're used to performing this swap (we never even question in in 6.003). So, don't worry about it for this problem, but be aware that this step is non trivial and is, in fact, frequently wrong.)

c Here, we are told that $m(t)$ and $m'(t)$ exist and are continuous. Combining parts a) and b) (let $f = m$ and $g = f' = m'$, with A containing at least the right-half plane), we have,

$$\lim_{t\to\infty} m'(t) = \lim_{r\to 0} rL_{m'}(r)$$

and

$$L_{m'}(r) = rL_m(r) - m(0) = rL_m(r)$$

where we used the fact that $m(0) = 0$ (why?). It follows that

$$\lim_{t\to\infty} m'(t) = \lim_{r\to 0} r^2 L_m(r)$$

Using the expansion given as Eqn. (3.7) in the notes, we have $L_m(r) = \frac{L_X(r)}{r(1-L_X(r))}$, from which it follows that

$$\lim_{t\to\infty} m'(t) \quad = \quad \lim_{r\to 0} r^2 \frac{L_X(r)}{r(1-L_X(r))} = \lim_{r\to 0} \frac{rL_X(r)}{(1-L_X(r))}$$

Since $\lim_{r\to 0} rL_X(r) = \lim_{r\to 0}(1 - L_X(r)) = 0$ (why?), we can evaluate this limit using the de L'Hôpital's rule. We then have

$$\begin{aligned}
\lim_{t\to\infty} m'(t) &= \lim_{r\to 0} \frac{\frac{d}{dr} rL_X(r)}{\frac{d}{dr}(1-L_X(r))} \\
&= \lim_{r\to 0} \frac{\frac{d}{dr} r\int_0^\infty e^{-rt} f_X(t)dt}{\frac{d}{dr}(1-\int_0^\infty e^{-rt} f_X(t)dt)} \\
&= \lim_{r\to 0} \frac{-r\int_0^\infty te^{-rt} f_X(t)dt + \int_0^\infty e^{-rt} f_X(t)dt}{\int_0^\infty te^{-rt} f_X(t)dt},
\end{aligned}$$

where in the last equality, we used the fact from calculus (the Liebniz integral rule) that for a real-valued function h, we can differentiate $\int_{\mathbb{R}} h(x,t)\mathrm{d}t$ with respect to x by integrating $\partial h(x,t)/\partial x$, as long as both h and $\partial h(x,t)/\partial x$ are continuous over the region of interest. (It's easy to check that it is the case here.) Taking the limit of the numerator yields

$$\lim_{r\to 0} \left(-r\int_0^\infty te^{-rt} f_X(t)dt + \int_0^\infty e^{-rt} f_X(t)dt \right) = \int_0^\infty f_X(t)dt = 1,$$

and that of the denominator yields

$$\lim_{r\to 0} \int_0^\infty te^{-rt} f_X(t)dt = \int_0^\infty tf_X(t)dt = \mathrm{E}(X).$$

Since both limits are finite, it follows that

$$\begin{aligned}
\lim_{t\to\infty} m'(t) &= \frac{\lim_{t\to\infty}(-r\int_0^\infty te^{-rt} f_X(t)\mathrm{d}t + \int_0^\infty e^{-rt} f_X(t)\mathrm{d}t)}{\lim_{t\to\infty} \int_0^\infty te^{-rt} f_X(t)dt} \\
&= \frac{1}{\mathrm{E}(X)},
\end{aligned}$$

which is what we set out to show.

Exercise 3.14

a Let X_i be IID and take on values in $\{2,4,8\}$ with equal probability. Let $J = \min\{n \mid X_n = 2\}$. Note that X_i denote the travel times on each try and that, at any given try, the travel time of exactly two days implies that the miner has reached safety. To show that J is indeed a stopping time, note that the event $\{J \geq n\} = \{X_1 \neq 2, \ldots, X_{n-1} \neq 2\}$ depends only on $X_1, \ldots, X_{n-1}$, and does therefore not depend on $X_n, X_{n+1}, \ldots$. Furthermore, $J \geq 0$, which implies J cannot be defective if $\mathrm{E}(J) < \infty$ (why?). But note that J is a geometric random variable, with success probability equal to $p = P(X_1 = 2)$. It follows that $E(J) = 1/p = 3 < \infty$ and J is a stopping time.

b Since J is a stopping time, Wald's equality yields $\mathrm{E}(T) = \mathrm{E}(X_1)\mathrm{E}(J) = \frac{2+4+8}{3} 3 = 14$.

c By the linearity of expectation,

$$\mathrm{E}\left(\sum_{i=1}^{J} X_i \;\middle|\; J = n\right) = \sum_{i=1}^{n} \mathrm{E}(X_i \mid J = n)$$

Given that $J = n$, we have that on the first $n-1$ tries, the miner failed to escape, only to escape on the n^{th} trial. It follows that $\mathrm{P}(X_i = 4 \mid J = n) = \mathrm{P}(X_i = 8 \mid J = n) = 1/2$ for $i = 1, \ldots, n = 1$ and $\mathrm{P}(X_n = 2) = 1$. Thus,

$$\mathrm{E}\left(\sum_{i=1}^{J} X_i \;\middle|\; J = n\right) = (n-1)\left(\frac{8}{2} + \frac{4}{2}\right) + 2 = 6(n-1) + 2 = 6n - 4.$$

Now compare this to

$$\mathrm{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{E}(X_i) = n\left(\frac{2}{3} + \frac{4}{3} + \frac{8}{3}\right) = \frac{14}{3}n,$$

the lesson here being that $\{\sum_{i=1}^{J} X_i \mid J = n\} \neq \{\sum_{i=1}^{n} X_i\}$ and that you should be careful before "discarding" the conditional universe.

d By the Total Probability Theorem,

$$\sum_{n=1}^{\infty} \mathrm{E}\left(\sum_{i=1}^{J} X_i \;\middle|\; J = n\right) \mathrm{P}(J = n) = \sum_{n=1}^{\infty} (6n - 4)\mathrm{P}(J = n) = 6\,\mathrm{E}(J) - 4 = 14,$$

as obtained in a). Wald's equality frequently isn't the only way to solve a simple problem, but it frequently is the unique tool that lets us approach more complicated situations.

Exercise 3.16

Let $Y(t) = S_{N(t)+1} - t$ and let the inter-renewal times be distributed as $f_X(x) = \lambda e^{-\lambda x}$ for $x \geq 0$. From Eq. 3.24 in the notes,

$$\lim_{t \to \infty} \frac{\int_{\tau=0}^{\infty} Y(\tau) \mathrm{d}\tau}{t} = \frac{\mathrm{E}(X^2)}{2\mathrm{E}(X)} = \frac{2/\lambda^2}{2/\lambda} = \frac{1}{\lambda} \text{ w.p.1.}$$

Recall that for a Poisson process with arrival rate λ, given any time $t \geq 0$, the expected time until the next arrival is given by $1/\lambda$. In this case, the time average matches the ensemble average.

To find the time-average second moment of $Y(t)$, consider the function $R(t) = Y(t)^2$ and note that $R(t)$ is a valid reward function since its value is uniquely determined by the position of t within the current inter-renewal interval. Since $R(t) \geq 0$ for all t and $\mathrm{E}(X) < \infty$, by Theorem 3.6 in the notes, it follows that

$$\lim_{t\to\infty} \frac{\int_{\tau=0}^{\infty} R(\tau)\mathrm{d}\tau}{t} = \frac{\mathrm{E}(R_n)}{\mathrm{E}(X)} \quad \text{w.p.1.}$$

Supposing that the inter-renewal interval containing t has length x, the total accumulated reward becomes $R_n = \int_0^X y^2 \mathrm{d}y = x^3/3$. It follows that $\mathrm{E}(R_n) = \mathrm{E}_X(\mathrm{E}_{R_n}(R_n \mid X = x)) = \mathrm{E}(X^3)/3$, where the subscripts indicate the random variable with respect to which the expecations are computed. It follows that

$$\lim_{t\to\infty} \frac{\int_{\tau=0}^{\infty} R(\tau)\mathrm{d}\tau}{t} = \frac{\mathrm{E}(X^3)}{3\mathrm{E}(X)} = \frac{3!/\lambda^3}{3/\lambda} = \frac{2}{\lambda^2} \quad \text{w.p.1.}$$

which matches once again the ensemble average of $Y(t)^2$ for a Poisson process, computed as

$$\mathrm{E}(Y(t)^2) = \mathrm{E}(X^2) = \frac{2}{\lambda^2},$$

where the first equality follows by the memoryless property of the exponential random variable.

Letting instead $f_X(x) = 3/(x+1)^4$ for $x \geq 0$, the first three moments of X are the following:

$$\mathrm{E}(X) = \int_0^{\infty} \frac{3x}{(1+x)^4}\mathrm{d}x = \frac{1}{2} \qquad \mathrm{E}(X^2) = \int_0^{\infty} \frac{3x^2}{(1+x)^4}\mathrm{d}x = 1$$

$$E(X^3) = \int_0^{\infty} \frac{3x^3}{(1+x)^4}\mathrm{d}x = \infty$$

Following the same reasoning as previously, the first two time-averaged moments are given by

$$\lim_{t\to\infty} \frac{\int_{\tau=0}^{\infty} Y(\tau)\mathrm{d}\tau}{t} = \frac{\mathrm{E}(X^2)}{2\mathrm{E}(X)} = \frac{1}{2/2} = 1 \quad \text{w.p.1}$$

and

$$\lim_{t\to\infty} \frac{\int_{\tau=0}^{\infty} Y^2(\tau)\mathrm{d}\tau}{t} = \frac{\mathrm{E}(X^3)}{3\mathrm{E}(X)} = \infty \quad \text{w.p.1.}$$

Note that the the time-average residual life of this process is less than that of the Poisson process with rate $\lambda = 1$. However, the squared residual life is infinite, while that of any Poisson process is finite.

Exercise 3.19

Assuming that the process is in the steady state (i.e. $Y(t), Z(t), X(t)$ depend only on t and not on where the actual inter-renewal period started), we use the fact that

$$f_{Z(t),X(t)}(z,x) = \frac{f_X(x)}{\mathrm{E}(X)} \quad \text{for} \quad x \geq z,$$

given as Eqn. 3.38 in the notes. The above also yields,

$$f_{Z(t)}(z) = \frac{1 - F_X(z)}{\mathrm{E}(X)}, \;\; z \geq 0 \quad \text{and} \quad f_{X(t)}(x) = \frac{x f_X(x)}{\mathrm{E}(X)}, \;\; x \geq 0.$$

a To compute $f_{Y(t)|Z(t+s/2)=s}(y)$ for given $s > 0$, it suffices to note that

$$\{Y(t) = y \mid Z(t{+}s/2) = s\} = \{X(t) = y{+}s/2 \mid Z(t{+}s/2) = s\} = \{X(t{+}s/2) = y{+}s/2 \mid Z(t{+}s/2) = s\}.$$

It follows that

$$\begin{aligned} f_{Y(t)|Z(t+s/2)=s}(y) &= f_{X(t+s/2)|Z(t+s/2)=s}(y + s/2) = \frac{f_{X(t+s/2),Z(t+s/2)}(y + s/2, s)}{f_{Z(t+s/2)}(s)} \\ &= \frac{f_X(y + s/2)/\mathrm{E}(X)}{(1 - F_X(s))/\mathrm{E}(X)} = \frac{f_X(y + s/2)}{(1 - F_X(s))} \quad \text{for } \; y \geq s/2 \end{aligned}$$

(Can you see why this is a valid density, that is, why it is non-negative and integrates to 1?)

b Since $\{Y(t) = y, Z(t) = z\} = \{X(t) = y + z, Z(t) = z\}$, it follows that

$$f_{Y(t),Z(t)}(y, z) = f_{X(t),Z(t)}(y + z, z) = \frac{f_X(y + z)}{\mathrm{E}(X)}, \qquad y, z \geq 0.$$

(Can you see how integrating out $Y(t)$ yields the familiar density for $Z(t)$?)

c Since

$$\{Y(t) = y \mid X(t) = x\} = \{Z(t) = x - y \mid X(t) = x\},$$

we have that

$$\begin{aligned} f_{Y(t)|X(t)=x}(y) &= f_{Z(t)|X(t)=x}(x - y) = \frac{f_{Z(t),X(t)}(x - y, x)}{f_{X(t)}(x)} = \frac{f_X(x)/\mathrm{E}(X)}{x f_X(x)/\mathrm{E}(X)} \\ &= \frac{1}{x}, \quad 0 \leq y \leq x. \end{aligned}$$

(Why was this to be expected? Hint: Given $X(t) = x$, $Z(t)$ and $Y(t)$ are identically distributed.)

d Similarly to part a),

$$\{Z(t) = z \mid Y(t{-}s/2) = s\} = \{X(t) = z{+}s/2 \mid Y(t{-}s/2) = s\} = \{X(t{-}s/2) = z{+}s/2 \mid Y(t{-}s/2) = s\},$$

from which it follows that

$$f_{Z(t)|Y(t-s/2)=s}(z) = f_{X(t-s/2)|Y(t-s/2)=s}(z + s/2) = \frac{f_{X(t-s/2),Y(t+s/2)}(z + s/2, s)}{f_{Y(t-s/2)}(s)}.$$

Note that for any t such that the process has reached steady state,

$$f_{X(t),Y(t)}(x, y) = f_{X(t),Z(t)}(x, x - y) = \frac{1}{x} \;, \;\; 0 \leq y \leq x$$

and, similarly, $f_{Y(t)}(y) = (1 - F_X(y))/\mathrm{E}(X)$ for $y \geq 0$. It follows that

$$f_{Z(t)|Y(t-s/2)=s}(z) = \frac{f_X(z + s/2)/\mathrm{E}(X)}{(1 - F_X(s))/\mathrm{E}(X)} = \frac{f_X(z + s/2)}{(1 - F_X(s))} \;, \;\; z \geq s/2.$$

(Again, why was this to be expected?)

e Finally, we have

$$f_{Y(t)|Z(t+s/2)\geq s}(y) = \frac{\int_s^\infty f_{Y(t),Z(t+s/2)}(y,a)\mathrm{d}a}{\int_s^\infty f_{Z(t+s/2)}(a)\mathrm{d}a}.$$

But note that,

$$\{Y(t) = y, Z(t+s/2) = a\} = \{Y(t+s/2) = y - s/2, Z(t+s/2) = a\}.$$

Applying the result of b),

$$f_{Y(t),Z(t+s/2)}(y,a) = f_{Y(t+s/2),Z(t+s/2)}(y-s/2,a) = \frac{f_X(y-s/2+a)}{\mathrm{E}(X)} \quad , \quad a \geq 0,\ y \geq s/2,$$

It follows that

$$f_{Y(t)|Z(t+s/2)\geq s}(y) = \frac{\int_s^\infty f_X(y+a-s/2)\mathrm{d}a/\mathrm{E}(X)}{\int_s^\infty (1-F_X(a))\mathrm{d}a/\mathrm{E}(X)} \quad , \quad y \geq s/2.$$

But, $\int_s^\infty f_X(y+a-s/2)\mathrm{d}a = 1 - F_X(y+s/2)$, yielding

$$f_{Y(t)|Z(t+s/2)\geq s}(y) = \frac{1-F_X(y+s/2)}{\int_s^\infty (1-F_X(a))\mathrm{d}a} \quad , \quad y \geq s/2.$$