IEOR 6711: Stochastic Models I

Professor Whitt

Solutions to Homework Assignment 1

The assignment consists of the following ten problems from Chapter 1: Problems 1.1-1.4, 1.8, 1.11, 1.12, 1.15, 1.22 and 1.37. You need not turn in problems with answers in the back. (Here those are 1.22 and 1.37.) Since you may not have the textbook yet, the problems are given again here and expanded upon. (However, there are plenty of copies of the textbook in the Columbia bookstore.) You are to do the extra parts added in this expansion, as well as the problems from the book. The expansion illustrates that there may be more going on than you at first think.

1. Elaboration on Problem 1.1 on p. 46.

(a) Application

This problem is about the *tail-integral formula for expected value*,

$$EX = \int_0^\infty P(X > t)\, dt = \int_0^\infty \bar{F}(t)\, dt \ ,$$

where $F \equiv F_X$ is the *cumulative distribution function* (cdf) of the random variable X, i.e., $F(t) \equiv P(X \le t)$, under the assumption that X is a nonnegative random variable. Here $\equiv$ denotes "equality by definition." The function $\bar{F} \equiv 1 - F$ is the *complementary cumulative distribution function* (ccdf) or tail-probability function.

The tail-integral formula for expected value is an alternative to the standard formula (usual definition)

$$EX = \int_0^\infty x f(x)\, dx \ .$$

The tail-integral formula for expected value can be proved in at least two ways: (i) by converting it to an iterated double integral and changing the order of integration, and (ii) by integration by parts. Before considering the proof, let us see why the formula is interesting and useful.

Apply the tail integral formula for the expected value to compute the expected value of an exponential random variable with rate λ and mean $1/\lambda$, i.e., for the random variable with tail probabilities (ccdf)

$$\bar{F}(t) \equiv P(X > t) \equiv e^{-\lambda t}, \quad t \ge 0 \ ,$$

and density (probability density function or pdf)

$$f(t) \equiv f_X(t) \equiv \lambda e^{-\lambda t}, \quad t \ge 0 \ .$$

You are supposed to see how easy it is this way:

$$EX = \int_0^\infty \bar{F}(t)\, dt = \int_0^\infty e^{-\lambda t}\, dt = \frac{1}{\lambda} \ .$$

(b) Alternative Approaches

Compute the expected value of the exponential distribution above in two other ways:

(i) Exploit the structure of the *gamma distribution*: The gamma density with *scale parameter* λ and *shape parameter* ν is

$$f_{\lambda,\nu}(t) \equiv \frac{1}{\Gamma(\nu)}\lambda^{\nu}t^{\nu-1}e^{-\lambda t}, \quad t \ge 0,$$

where Γ is the *gamma function*, i.e.,

$$\Gamma(t) \equiv \int_0^{\infty} x^{t-1}e^{-x}\, dx \ ,$$

which reduces to a factorial at integer arguments: $\Gamma(n+1) = n!$. The important point for the proof here is that the gamma density is a proper probability density, and so integrates to 1.

This way is also easy, but it exploits special knowledge about the gamma distributions. We just write down the standard definition of the expected value, and then see the connection to the gamma density for $\nu = 2$. (The exponential distribution is the gamma distribution with $\nu = 1$.)

$$EX = \int_0^{\infty} tf(t)\, dt = \int_0^{\infty} t\lambda e^{-\lambda t}\, dt = \int_0^{\infty} (1/\lambda) f_{\lambda,2}(t)\, dt = \frac{1}{\lambda}$$

because the gamma density $f_{\lambda,2}$ integrates to 1.

(ii) Use *integration by parts*, e.g., see Feller II (1971), p. 150: Suppose that u is bounded and has continuous derivative u'. Suppose that f is a pdf of a nonnegative random variable with associated ccdf $\bar{F}$. (The derivative of $\bar{F}$ is $-f$.) Then, for any b with $0 < b < \infty$,

$$\int_0^b u(t)f(t)\, dt = -u(b)\bar{F}(b) + u(0)\bar{F}(0) + \int_0^b u'(t)\bar{F}(t)\, dt \ .$$

This is the straightforward approach, but it is somewhat complicated. First,

$$EX = \int_0^{\infty} tf(t)\, dt = \lim_{b\to\infty} \int_0^b tf(t)\, dt \ .$$

Next apply integration by parts, in the form above, to get

$$\begin{aligned}
\int_0^b tf(t)\, dt &= -u(b)\bar{F}(b) + u(0)\bar{F}(0) + \int_0^b u'(t)\bar{F}(t)\, dt \\
&= -be^{-\lambda b} + 0e^{-\lambda 0} + \int_0^b 1e^{-\lambda t}\, dt \\
&= -be^{-\lambda b} + \frac{1}{\lambda}(1 - e^{-\lambda b}) \ .
\end{aligned}$$

Now let $b \to \infty$ to get

$$EX = \lim_{b\to\infty} \int_0^b tf(t)\, dt = \lim_{b\to\infty} \{-be^{-\lambda b} + \frac{1}{\lambda}(1 - e^{-\lambda b})\} = \frac{1}{\lambda} \ .$$

The same approach produces a proof of the general formulas.

(c) Interchanging the order of integrals and sums.

Relatively simple proofs of the results to be proved in Problem 1.1 of Ross follow from interchanging the order of integrals and sums. We want to use the following relations:

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{i,j} = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{i,j}$$

and

$$\int_0^{\infty}\int_0^{\infty} f(x,y)\,dx\,dy = \int_0^{\infty}\int_0^{\infty} f(x,y)\,dy\,dx$$

It is important to know that these relations are *usually valid.* It is also important to know that these relations are *not always valid*: In general, there are *regularity conditions.* The complication has to do with infinity and limits. There is no problem at all for finite sums:

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j} = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{i,j} \ .$$

That is always true. Infinite sums involve limits, and integrals are defined as limits of sums.

The interchange property is often used with expectations. In particular, it is used to show that the expectation can be taken inside sums and integrals:

$$E[\sum_{i=1}^{\infty} X_i] = \sum_{i=1}^{\infty} EX_i$$

and

$$E[\int_0^{\infty} X(s)\,ds] = \int_0^{\infty} E[X(s)]\,ds \ .$$

These relations are of the same form because the expectation itself can be expressed as a sum or an integral. However, there are regularity conditions, as noted above.

(i) Compute the two iterated sums $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{i,j}$ and $\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{i,j}$ for

$$a_{i,j} = 2 - 2^{-i} \quad \text{for} \quad i = j, \quad a_{i,j} = -2 + 2^{-i} \quad \text{for} \quad i = j+1, \quad j \geq 1 \ ,$$

and $a_{i,j} = 0$ otherwise. What does this example show?

In this example the two iterated sums are not equal. This example shows that regularity conditions are needed. It would suffice to have either: (i) $a_{i,j} \geq 0$ for all i and j (Tonelli) or (ii) One of the iterated sums, or the double sum, be finite when $a_{i,j}$ is replaced by its absolute value $|a_{i,j}|$ for all i and j (Fubini).

In particular, here

$$\sum_{i=1}^{\infty} a_{i,j} = -2^{-(j+1)} \ ,$$

so that

$$\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{i,j} = -1/2 \ .$$

On the other hand,

$$\sum_{j=1}^{\infty} a_{i,j} = 0 \quad \text{for} \quad i \ge 2 \quad \text{and} \quad \sum_{j=1}^{\infty} a_{1,j} = 2 - 2^{-1} = 3/2 \ .$$

Hence,

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{i,j} = 3/2 \quad \text{while} \quad \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{i,j} = -1/2 \ .$$

(ii) Look up Tonelli's theorem and Fubini's theorem in a book on measure theory and integration, such as *Real Analysis* by H. L. Royden. What do they say about this problem?

Tonelli's theorem says that the interchange is valid for nonnegative summands and (measurable) integrands. There are also conclusions about measurability, but those are relatively minor technical issues. Fubini's theorem says that the interchange is valid if the summands and integrands are summable or integrable, which essentially (aside from measurability issues) means that the absolute values of the summands and integrands are summable and integrable, respectively. Neither condition is satisfied in the example above, so that the overall sum depends on the order in which we perform the summation. We could even get $\infty - \infty$. (There are infinitely many $+2's$ and infinitely many $-2's$.)

(d) Do the three parts to Problem 1.1 in Ross (specified here below).

Hint: Do the proofs by interchanging the order of sums and integrals. The first step is a bit tricky. We need to get an iterated sum or integral; i.e., we need to insert the second sum or integral. To do so for sums, write:

$$EN = \sum_{i=1}^{\infty} iP(N = i) = \sum_{i=1}^{\infty}[\sum_{j=1}^{i} 1]P(N = i) \ .$$

(i) For a nonnegative integer-valued random variable N, show that

$$EN = \sum_{i=1}^{\infty} P(N \ge i) \ .$$

(ii) For a nonnegative random variable X with cdf F, show that

$$EX = \int_0^{\infty} P(X > t)\, dt = \int_0^{\infty} \bar{F}(t)\, dt \ .$$

(iii) For a nonnegative random variable X with cdf F, show that

$$E[X^n] = \int_0^{\infty} nt^{n-1}\bar{F}(t)\, dt \ .$$

(i) Once we have the double sum, we can change the order of summation. The interchange is justified because the summand is nonnegative; we can apply Tonelli's theorem. A key practical step is to get the range of summation right:

$$\begin{aligned} EN &= \sum_{i=1}^{\infty} iP(N=i) = \sum_{i=1}^{\infty}[\sum_{j=1}^{i} 1]P(N=i) \\ &= \sum_{j=1}^{\infty}\sum_{i=j}^{\infty} 1P(N=i) \\ &= \sum_{j=1}^{\infty} P(N \geq j) \ . \end{aligned}$$

(ii) Now consider the first integral. Then

$$\begin{aligned} EX &= \int_0^{\infty} xf(x)\,dx = \int_0^{\infty}[\int_0^{x} 1\,dy]f(x)\,dx \\ &= \int_0^{\infty} 1[\int_y^{\infty} f(x)\,dx]\,dy \\ &= \int_0^{\infty} \bar{F}(y)\,dy \ . \end{aligned}$$

(iii) Finally, turning to the general case,

$$\begin{aligned} E[X^n] &= \int_0^{\infty} x^n f(x)\,dx = E[\int_0^{\infty}[\int_0^{x} ny^{n-1}\,dy]f(x)\,dx \\ &= \int_0^{\infty} ny^{n-1}[\int_y^{\infty} f(x)\,dx]\,dy \\ &= \int_0^{\infty} ny^{n-1}[\bar{F}(y)]\,dy \ . \end{aligned}$$

(e) What regularity property justifies the interchanges used in part (d)?

The summands and integrands are all nonnegative, so we can apply Tonelli's theorem.

(f) What happens to the tail-integral formula for expected value

$$EX = \int_0^{\infty} P(X > t)\,dt = \int_0^{\infty} \bar{F}(t)\,dt \ ,$$

when the random variable X is integer valued?

It reduces to the first formula given in Ross:

$$EN = \sum_{j=1}^{\infty} P(N \geq j) \ .$$

---

2. Elaboration of Problems 1.2 on p. 46 and 1.15 on p. 49.

(a) Application to simulation

(i) Suppose that you know how to generate on the computer a random variable U that is uniformly distributed on the interval $[0, 1]$. (That tends to be easy to do, allowing for the necessary approximation due to discreteness.) How can you use U to generate a random variable X with an cdf F that is an arbitrary continuous and strictly increasing function? (Hint: Use the inverse F^{-1} of the function F.)

Since the function F is continuous and strictly increasing, it has an inverse, say F^{-1}, with the properties

$$F^{-1}(F(x)) = x \quad \text{and} \quad F(F^{-1}(t)) = t$$

for all x and t with $0 < t < 1$. Hence, we can use the random variable $F^{-1}(U)$ because, for all x,

$$P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x) \ .$$

(ii) Suppose that you know how to generate on the computer a random variable U that is uniformly distributed on the interval $[0, 1]$. How can you use U to generate a random variable X with an exponential distribution with mean $1/\lambda$?

We are given the cdf $F(t) = 1 - e^{-\lambda t}$, $t \geq 0$. We want to find the inverse. For t given, $0 < t < 1$, it suffices to find the value of x such that $1 - e^{-\lambda x} = t$. That occurs when $1 - t = e^{-\lambda x}$. Taking logarithms, we find that

$$F^{-1}(t) = -\frac{1}{\lambda} \log(1 - t) \ ,$$

where log is the natural (base e) logarithm. Hence, $-\frac{1}{\lambda} \log(1 - U)$ has the given exponential distribution. Since $1 - U$ has the same distribution as U, the random variable $-\frac{1}{\lambda} \log(U)$ has the same exponential distribution. (incidentally, this is Problem 1.15 (b).)

(b) Problem 1.2 for a cdf F having a positive density f.

Do the two parts of Problem 1.2 in Ross (specified here below) under the assumption that the random variable X concentrates on the interval (a, b) for $-\infty \leq a < b \leq +\infty$ and has a strictly positive density f there. That is, $P(a \leq X \leq b) = 1$,

$$F(x) \equiv P(X \leq x) = \int_a^x f(y)\, dy \ ,$$

where $f(x) > 0$ for all x such that $a \leq x \leq b$.

The two parts of Problem 1.2 are:

(i) If X is a random variable with a continuous cdf F, show that $F(X)$ is uniformly distributed on the interval $[0, 1]$.

The positive density condition implies that the cdf is continuous and strictly increasing. Hence we are in the setting of the previous part: the cdf F has an inverse F^{-1}. We can exploit the inverse property to write

$$P(F(X) \leq t) = P(X \leq F^{-1}(t)) = F(F^{-1}(t)) = t \ ,$$

which implies that $F(X)$ is distributed the same as U, i.e., uniformly on $(0, 1)$.

(ii) If U is uniformly distributed on $(0, 1)$, then $F^{-1}(U)$ has cdf F.

Given the inverse property, this part is repeating the Problem 2. (a) (i) above.

(c) Right continuity.

A cdf F is a right-continuous nondecreasing function of a real variable such that $F(x) \to 1$ as $x \to \infty$ and $F(x) \to 0$ as $x \to -\infty$.

A real-valued function g of a real variable x is *right-continuous* if

$$lim_{y \downarrow x} g(y) = g(x)$$

for all real x, where $\downarrow$ means the limit from above (or from the right). A function g is *left-continuous* if

$$lim_{y \uparrow x} g(y) = g(x)$$

for all real x, where $\uparrow$ means the limit from below (or from the left).

A function g has a limit from the right at x if the limit $lim_{y \downarrow x} g(y)$ exists. Suppose that a function g has limits everywhere from the left and right. Then the *right-continuous version* of g, say g_+ is defined by

$$g_+(x) \equiv \lim_{y \downarrow x} g(y)$$

for all x. The *left-continuous version* of g, g_- is defined similarly.

Suppose that $P(X = 1) = 1/3 = 1 - P(X = 3)$. Let F be the cdf of X. What are $F(1)$, $F(2)$, $F(3)$ and $F(4)$? What are the values of the left-continuous version F_- at the arguments 1, 2, 3 and 4?

Right-continuity means that $F(x) = P(X \leq x)$, while left-continuity means that $F_-(x) = P(X < x)$. Hence,

$$F(1) = 1/3 = F(2), \quad \text{and} \quad F(3) = 1 = F(4) \ ,$$

while

$$F_-(1) = 0, \quad F_-(2) = 1/3 = F_-(3) \quad \text{and} \quad F_-(4) = 1 \ .$$

(d) The left-continuous inverse of a cdf F.

Given a (right-continuous) cdf F, let

$$F^{\leftarrow}(t) \equiv \inf\{x : F(x) \geq t\}, \quad 0 < t < 1 \ .$$

Fact: $F^{\leftarrow}$ is a left-continuous function on the interval $(0, 1)$.

Fact: In general, $F^{\leftarrow}$ need not be right-continuous.

Fact:

$$F^{\leftarrow}(t) \leq x \quad \text{if and only if} \quad F(x) \geq t$$

for all t and x with $0 < t < 1$.

Suppose that X is a random variable with a continuous cdf F. (We now do not assume that F has a positive density.) Show that $F(X)$ is uniformly distributed on the interval $(0, 1)$. (Hint: use the fact that $F(F^{\leftarrow}(t)) = t$ for all t with $0 < t < 1$ when F is continuous.)

For all t, $0 < t < 1$,

$$P(F(X) \leq t) = P(X \leq F^{\leftarrow}(t)) = F(F^{\leftarrow}(t)) = t \ ,$$

using the continuity of F in the last step.

(e) The right-continuous inverse of a cdf F.

Given a (right-continuous) cdf F, let

$$F^{-1}(t) \equiv \inf\{x : F(x) > t\}, \quad 0 < t < 1 \ .$$

Fact: It can be shown that F^{-1} is right-continuous, but in general is not left-continuous.

Show that the relation

$$F^{-1}(t) \leq x \quad \text{if and only if} \quad F(x) \geq t$$

for all t and x with $0 < t < 1$ does *not hold* in general.

We give a counterexample: Let $P(X = 1) = P(X = 2) = 1/2$. Then $F(1) = 1/2$ and $F(2) = 1$. Note that $F(1) \geq 1/2$, but $F^{-1}(1/2) = 2$, which is not less than or equal to 1.

(f) Inverses. Let X be the discrete random variable defined above with $P(X = 1) = 1/3 = 1 - P(X = 3)$. Draw pictures (graphs) of the cdf F of X and the two inverses $F^{\leftarrow}$ and F^{-1}.

If there is a jump in F at x, $F(x)$ assumes the higher value, after the jump; that makes F right-continuous. The graph of $F^{\leftarrow}$ is just the graph of F with the x and y axes switched. That makes $F^{\leftarrow}$ left-continuous. Then F^{-1} is the right-continuous version of $F^{\leftarrow}$.

(g) Suppose that you know how to generate on the computer a random variable U that is uniformly distributed on the interval $[0, 1]$. How can you use U to generate a random variable X with an *arbitrary* cdf F?

Let $X = F^{\leftarrow}(U)$. Then, for any x,

$$P(X \leq x) = P(F^{\leftarrow}(U) \leq x) = P(U \leq F(x)) = F(x) \ ,$$

as desired.

(h) It can be shown that $P(F^{-1}(U) = F^{\leftarrow}(U)) = 1$. Given that result, what is the distribution of $F^{-1}(U)$?

The random variable $F^{-1}(U)$ also has cdf F.

---

3. Problem 1.3 in Ross:

Let the random variable X_n have a binomial distribution with parameters n and p_n, i.e.,

$$P(X_n = k) = \frac{n!}{k!(n-k)!} p_n^k (1 - p_n)^{n-k} \ .$$

Show that

$$P(X_n = k) \to \frac{e^{-\lambda}\lambda^k}{k!} \quad \text{as} \quad n \to \infty$$

if $np_n \to \lambda$ as $n \to \infty$. Hint: Use the fact that $(1 + (c_n/n))^n \to e^c$ as $n \to \infty$ if $c_n \to c$ as $n \to \infty$.

This limit is sometimes referred to as the *law of small numbers.*

$$\begin{aligned} P(X_n = k) &= \frac{n!}{k!(n-k)!} p_n^k (1 - p_n)^{n-k} \\ &= \frac{n!}{k!(n-k)!n^k} (np_n)^k \frac{(1 - (np_n/n)^n}{(1 - p_n)^k} \\ &\to \frac{\lambda^k e^{-\lambda}}{k!} \end{aligned}$$

because $(np_n)^k \to \lambda^k$, $n!/(n-k)!n^k \to 1$, $(1 - p_n)^k \to 1$ and $(1 - (np_n/n)^n \to e^{-\lambda}$.

The limit above is an example of convergence in distribution. We say that random variables X_n with cdf's F_n *converge in distribution* to a random variable X with cdf F, and write $X_n \Rightarrow X$ or $F_n \Rightarrow F$, if

$$F_n(x) \to F(x) \quad \text{as} \quad n \to \infty \quad \text{for all} \quad x \quad \text{that are continuity points of } F \ .$$

A point x is a continuity point of F if F is continuous at x.

(b) Show that the limit above is indeed an example of convergence in distribution. What is the limiting distribution?

The limiting distribution is the Poisson distribution. We have shown that $P(X_n = k) \to P(X = k)$ for all k. We need to show that $P(X_n \le x) \to P(X \le x)$ for all x that are not continuity points of the limiting cdf $P(X \le x)$. That means for all x that are not nonnegative integers, but because all random variables are integer valued, it suffices to show that

$$P(X_n \le k) \to P(X \le k) \quad \text{for all} \quad k \ .$$

(All the cdf's remain unchanged between integer arguments; e.g., $F(x) = F(\lfloor x \rfloor)$, where $\lfloor x \rfloor$ is the greatest integer less than or equal to x.) But that limit for the cdf's follows by induction from the property that the limit of a sum is the sum of the limits: If $a_n \to a$ and $b_n \to b$ as $n \to \infty$, then $a_n + b_n \to a + b$ as $n \to \infty$. (The property remains true for random variables defined on all integers, negative and positive, but that involves an extra argument.)

(c) Suppose that $P(X_n = 1 + n^{-1}) = P(X = 1) = 1$ for all n. Show that $X_n \Rightarrow X$ as $n \to \infty$ and that the continuity-point condition is needed here.

The main point here is that the continuity-point condition is needed, because $F_n(1) \equiv P(X_n \le 1) = 0$ for all n, while $F(1) \equiv P(X \le 1) = 1$. Since 1 is not a continuity point of F, the missing convergence does not cause a problem. For any $x > 1$, $F(x) = 1$. Moreover, there is an n_0 such that $1 + n_0^{-1} < x$, so that $F_n(x) = 1$ for all $n > n_0$. On the other hand, for any $x < 1$, $F_n(x) = F(x) = 0$. Hence, $F_n \Rightarrow F$.

(d) Use the techniques of Problem 2 above to prove the following *representation theorem*:

Theorem 0.1 *If $X_n \Rightarrow X$, then there exist random variables $\tilde{X}_n$ and $\tilde{X}$ defined on a common probability space such that $\tilde{X}_n$ has the same distribution F_n as X_n for all n, $\tilde{X}$ has the same distribution F as X, and*

$$P(\tilde{X}_n \to \tilde{X} \quad as \quad n \to \infty) = 1 \ .$$

Start with a random variable U uniformly distributed on $[0, 1]$. Let $\tilde{X}_n = F_n^{\leftarrow}(U)$ for all n and let $\tilde{X} = F^{\leftarrow}(U)$. By Problem 2, $\tilde{X}_n$ has the same distribution as X_n for all n and $\tilde{X}$ has the same distribution as X. It remains to show the w.p.1 convergence. Note that the same random variable U is used throughout. The assumed convergence means that $F_n(x) \to F(x)$ for all x that are continuity points of F. We need three technical lemmas at this point:

Lemma 0.1 *The convergence*

$$F_n^{\leftarrow}(t) \to F^{\leftarrow}(t) \quad as \quad n \to \infty$$

holds for for all t in $(0, 1)$ that are continuity points of $F^{\leftarrow}$ if and only if

$$F_n(x) \to F(x) \quad as \quad n \to \infty$$

holds for for all x in $\mathbb{R}$ that are continuity points of F.

This first lemma is tedious to prove, but it can be done. Supposing that $F_n \Rightarrow F$, you pick t such that t is a continuity point of $F^{\leftarrow}$. Then we show that indeed $F_n^{\leftarrow}(t) \to F^{\leftarrow}(t)$ as $n \to \infty$. That is a tedious ϵ and δ argument (or, at least, that is how I would proceed)..

Lemma 0.2 *If g is a nondecreasing real-valued function of a real variable, the number of discontinuity points of g is a countably infinite set.*

Lemma 0.3 *If A is a countably infinite subset of the real line $\mathbb{R}$ and X is a random variable with a probability density function, i.e., if*

$$F(x) = \int_{-\infty}^{x} f(y)\, dy, \quad -\infty < x < \infty \ ,$$

then

$$P(X \in A) = \int_A f(x)\, dx = 0 \ .$$

As a consequence of the three technical lemmas, the assumed convergence $F_n \to F$ implies that $F_n^{\leftarrow}(t) \to F^{\leftarrow}(t)$ as $n \to \infty$ for all t in $(0, 1)$ except for a set of t of measure 0 under U. (We omit a full demonstration of this technical point.) Thus

$$P(F_n^{\leftarrow}(U) \to F^{\leftarrow}(U) \quad \text{as} \quad n \to \infty) = 1 \ .$$

---

4. Problem 1.4 in Ross:

Derive the mean and variance of a binomial random variable with parameters n and p. Hint: Use the relation between Bernoulli and binomial random variables.

Following the hint, use the fact that the binomial random variable is distributed as $S_n \equiv X_1 + \cdots + X_n$, where X_i are IID Bernoulli random variables, with $P(X_1 = 1) = p = 1 - P(X_1 = 0)$. Clearly,

$$EX_1 = p \quad \text{and} \quad Var(X_1) = p(1 - p) \ .$$

Since the mean of a sum is the sum of the means (always) and the variance of the sum is the sum of the variances (because of the independence), we have

$$ES_n = np \quad \text{and} \quad Var(S_n) = np(1 - p) \ .$$

See formula (1.3.4) on p. 10 of Ross for the formula of the variance of a sum of random variables in the general case in which independence need not hold.

---

5. Problem 1.8 in Ross:

Let X_1 and X_2 be independent Poisson random variables with means λ_1 and λ_2, respectively.

(a) Find the distribution of the sum $X_1 + X_2$.

The sum of independent Poisson random variables is again Poisson. That is easy to show with moment generating functions:

$$\phi_{X_i}(t) \equiv E[e^{tX_i}] = e^{\lambda_i(e^t - 1)} \ .$$

Because of the independence,

$$\phi_{X_1+X_2}(t) = \phi_{X_1}(t)\phi_{X_2}(t) = e^{\lambda_1(e^t-1)}e^{\lambda_2(e^t-1)} = e^{(\lambda_1+\lambda_2)(e^t-1)} .$$

The result is also not too hard to derive directly. We again exploit the independence and use the convolution formula for sums:

$$\begin{aligned}
P(X_1 + X_2 = k) &= \sum_{j=0}^{k} P(X_1 = j)P(X_2 = k-j) \\
&= \sum_{j=0}^{k} \frac{e^{-\lambda_1}\lambda_1^j}{j!}\frac{e^{-\lambda_2}\lambda_2^{k-j}}{(k-j)!} \\
&= \sum_{j=0}^{k} \frac{e^{-(\lambda_1+\lambda_2)}\lambda_1^j\lambda_2^{k-j}}{j!(k-j)!} \\
&= e^{-(\lambda_1+\lambda_2)} \sum_{j=0}^{k} \frac{\lambda_1^j\lambda_2^{k-j}}{j!(k-j)!} \\
&= e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^k \sum_{j=0}^{k} \frac{\frac{\lambda_1^j}{(\lambda_1+\lambda_2)^j}\frac{\lambda_2^{k-j}}{(\lambda_1+\lambda_2)^{(k-j)}}}{j!(k-j)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^k}{k!} k! \sum_{j=0}^{k} \frac{\frac{\lambda_1^j}{(\lambda_1+\lambda_2)^j}\frac{\lambda_2^{k-j}}{(\lambda_1+\lambda_2)^{(k-j)}}}{j!(k-j)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^k}{k!} \sum_{j=0}^{k} k! \frac{\frac{\lambda_1^j}{(\lambda_1+\lambda_2)^j}\frac{\lambda_2^{k-j}}{(\lambda_1+\lambda_2)^{(k-j)}}}{j!(k-j)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^k}{k!}
\end{aligned}$$

using the law of total probabilities for the binomial distribution in the last step.

(b) Compute the conditional distribution of X_1 given that $X_1 + X_2 = n$.

Use the definition of conditional probability and apply the necessary algebra:

$$\begin{aligned}
P(X_1 = k | X_1 + X_2 = n) &= \frac{P(X_1 = k \quad \text{and} \quad X_1 + X_2 = n)}{P(X_1 + X_2 = n)} \\
&= \frac{P(X_1 = k \quad \text{and} \quad X_2 = n-k)}{P(X_1 + X_2 = n)} \\
&= \frac{P(X_1 = k)P(X_2 = n-k)}{P(X_1 + X_2 = n)} \\
&= \frac{\frac{e^{-\lambda_1}\lambda_1^k}{k!}\frac{e^{-\lambda_2}\lambda_2^{n-k}}{(n-k)!}}{\frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!}} \\
&= \frac{n!}{k!(n-k)!}(\lambda_1/(\lambda_1+\lambda_2))^k(\lambda_2/(\lambda_1+\lambda_2))^{n-k}
\end{aligned}$$

So that $(X_1|X_1+X_2=n)$ has a binomial distribution with parameters n and $p=(\lambda_1/(\lambda_1+\lambda_2)$.

---

6. Problem 1.11 in Ross: generating functions.

Let X be a nonnegative integer-valued random variable. Then the generating function of X is

$$\hat{P}(z) \equiv E[z^X] = \sum_{j=0}^{\infty} z^j P(X=j) \; .$$

(a) Show that the k^{th} derivative of $\hat{P}(z)$ evaluated at $z=0$ is $k!P(X=k)$.

Use mathematical induction to justify the formula for all k. Differentiate term by term to get

$$\begin{aligned}
\frac{d^k \hat{P}(z)}{dz^k} &= \sum_{j=k}^{\infty} j(j-1)(j-2)\cdots(j-k+1)z^{j-k}P(X=j) \\
&= k!P(X=k) + \sum_{j=k+1}^{\infty} j(j-1)(j-2)\cdots(j-k+1)z^{j-k}P(X=j)
\end{aligned}$$

Evaluating at $z=0$ gives the desired result.

(b) Show that (with 0 being considered even)

$$P(X \text{ is even}) = \frac{\hat{P}(1)+\hat{P}(-1)}{2} \; .$$

$$\hat{P}(1)+\hat{P}(-1) = \sum_j P(X=j) + \sum_{j:even} P(X=j) - \sum_{j:odd} P(X=j) = 2\sum_{j:even} P(X=j) \; .$$

(c) Calculate $P(X$ is even $)$ when

(i) X is binomial (n,p);

When X is binomial(n,p),

$$\hat{P}(z) = (zp+1-p)^n$$

so that

$$P(X \text{ even}) = \frac{1^n+(1-2p)^n}{2} = \frac{1+(1-2p)^n}{2}$$

(ii) X is Poisson with mean λ;

When X is Poisson(λ),

$$\hat{P}(z) = e^{\lambda(z-1)}$$

so that

$$P(X \text{ even}) = \frac{e^0 + e^{-2\lambda}}{2} = \frac{1 + e^{-2\lambda}}{2}$$

(iii) X is geometric with parameter p, i.e., $P(X = k) = p(1-p)^{k-1}$ for $k \geq 1$.

When X is geometric(p),

$$\hat{P}(z) = \frac{zp}{(1 - z + zp)}$$

so that

$$P(X \text{ even}) = \frac{1 + \frac{-p}{2-p}}{2} = \frac{1-p}{2-p}$$

---

7. Problem 1.12 in Ross:

If $P(0 \leq X \leq a) = 1$, show that

$$Var(X) \leq a^2/4.$$

To get your intuition going, consider the obvious candidate distribution on the interval $[0, a]$ with large variance: Put mass $1/2$ on 0 and a. The variance of that two-point distribution is $a^2/4$, which is our desired bound. So it looks right, to get started.

It is perhaps easier to bound $E[X^2]$ above, given EX: Note that $X^2 \leq aX$ under the condition. Then take expected values, getting

$$E[X^2] \leq aEX \ .$$

Then

$$Var(X) = E[X^2] - (EX)^2 \leq aEX - (EX)^2 = EX(a - EX) \ ,$$

but EX can be anything in the interval $[0, a]$, so let $EX = pa$ for some p with $0 \leq p \leq 1$. Then

$$Var(X) \leq EX(a - EX) = pa(a - pa) = a^2(p(1-p)) \leq \frac{a^2}{4} \ ,$$

because, by calculus, $p(1-p)$ is maximized at $p = 1/2$.

---

8. Problem 1.22 in Ross:

The *conditional variance* of X given Y is defined as

$$Var(X|Y) \equiv E[(X - E[X|Y])^2|Y] \ .$$

Prove the *conditional variance formula*:

$$Var(X) = E[Var(X|Y)] + Var(E[X|Y]) \ .$$

A key relation is $EX = E[E[X|Y]]$ for any random variables X and Y. It is easier to start with the two pieces and put them together. First,

$$\begin{aligned} E[Var(X|Y)] &= E[E[(X - E[X|Y])^2|Y]] \\ &= E[E[X^2 - 2XE[X|Y] + E[X|Y]^2|Y]] \\ &= E[E[X^2|Y] - 2E[X|Y]^2 + E[X|Y]^2] \\ &= E[E[X^2|Y] - E[X|Y]^2] \\ &= E[E[X^2|Y]] - E[E[X|Y]^2]] \\ &= E[X^2] - E[E[X|Y]^2]] \ . \end{aligned}$$

Next,

$$\begin{aligned} Var(E[X|Y]) &= E[E[X|Y]^2] - (E[E[X|Y]])^2 \\ &= E[E[X|Y]^2] - (E[X])^2 \ . \end{aligned}$$

Therefore,

$$E[Var(X|Y)]+Var(E[X|Y]) = (E[X^2]-E[E[X|Y]^2])+(E[E[X|Y]^2]-(E[X])^2) = E[X^2]-(E[X])^2 = Var(X) \ .$$

---

9. Problem 1.37 in Ross:

Let $\{X_n : n \geq 1\}$ be a sequence of independent and identically distributed (IID) random variables with a probability density function f. Say that a peak occurs at index n if $X_{n-1} < X_n > X_{n+1}$. (We stipulate that, then, a peak cannot occur for $n = 1$.) Show that the long-run proportion of indices at which a peak occurs is, with probability 1, equal to $1/3$. (Hint: Use the strong law of large numbers for partial sums of IID random variables.)

Since the random variables X_n are IID with a probability density function, the probability any two random variables are identically equal is 0. So we do not have to worry about ties. For $n \geq 2$, each of the three variables X_{n-1}, X_n and X_{n+1} is equally likely to be the largest of these three. So the probability that a peak occurs at time (index) n is $1/3$. But X_n and X_{n+1} are dependent. If n is a peak, then $n+1$ cannot be a peak.

Let $Y_n = 1$ if a peak occurs at time n and 0 otherwise. We know that $EY_n = P(Y_n = 1) = 1/3$ for all $n \geq 2$, but the successive Y_n variables are dependent. However, $\{Y_2, Y_5, Y_8, \ldots\}$, $\{Y_3, Y_6, Y_9, \ldots\}$ and $\{Y_4, Y_7, Y_{10}, \ldots\}$ are three identically distributed sequences of IID random variables. For each one, we can apply the strong law of large numbers for IID random variables and deduce that the long-run proportion within each sequence is $1/3$ with probability 1. For example, we have

$$n^{-1} \sum_{k=1}^{n} Y_{2+3k} \to \frac{1}{3} \quad \text{as} \quad n \to \infty \text{ w.p.1} \ .$$

Those three separate limits then imply that

$$n^{-1} \sum_{k=1}^{n} Y_k \rightarrow \frac{1}{3} \quad \text{as} \quad n \rightarrow \infty \text{ w.p.1 } .$$

IEOR 6711: Stochastic Models I

Professor Whitt

Solutions to Homework Assignment 2

Problem 1.6 Answer in back.

Problem 1.7 Imagine that there are k buckets and we choose k balls and fill each bucket at the same time. Set

$$I_i = \begin{cases} 1 & \text{if } i\text{-th bucket contains a white ball} \\ 0 & \text{if } i\text{-th bucket contains a black ball.} \end{cases}$$

Then $X = \sum_{i=1}^{k} I_i$. Observe that for $i \neq j$,

$$\begin{aligned} \mathbf{E}[I_i] &= \mathbf{E}[I_i^2] = \mathbf{P}(I_i = 1) = \frac{n}{n+m}, \\ \mathbf{E}[I_i|I_j = 1] &= \mathbf{P}(I_i = 1|I_j = 1) = \frac{n-1}{n+m-1} \text{ and} \\ \mathbf{E}[I_i I_j] &= \mathbf{E}[\mathbf{E}[I_i I_j|I_j]] = \mathbf{E}[I_j \mathbf{E}[I_i|I_j]] = \mathbf{P}(I_j = 1)\mathbf{E}[I_i|I_j = 1] = \frac{n}{n+m} \times \frac{n-1}{n+m-1} \; . \end{aligned}$$

Using these, we get

$$\begin{aligned} \mathbf{E}[X] &= k\mathbf{P}(I_1 = 1) = \frac{kn}{n+m} \; , \\ \mathbf{E}[X^2] &= k\mathbf{E}[I_1^2] + k(k-1)\mathbf{E}[I_1 I_2] \\ &= \frac{kn}{n+m} + \frac{k(k-1)n(n-1)}{(n+m)(n+m-1)} \; , \text{ and} \\ \mathbf{V}(X) &= \mathbf{E}[X^2] - \mathbf{E}[X]^2 = \frac{knm(n+m-k)}{(n+m)^2(n+m-1)}. \end{aligned}$$

Problem 1.9 Assume that the results of each pairing are independent with each of the players being equally likely to win. For each permutation $i_1, \cdots, i_n$ of $1, 2, \cdots, n$ define an indicator variable $I_{(i_1,\cdots,i_n)}$ equal to 1 if that permutation is a Hamiltonian and 0 if it is not. Then

$$\begin{aligned} \mathbf{E}[\text{Number of hamiltonians}] &= \mathbf{E}[\textstyle\sum I_{(i_1,\cdots,i_n)}] \\ &= n!\mathbf{E}[I_{(1,2,\cdots,n)}] \\ &= \frac{n!}{2^{n-1}} \; . \end{aligned}$$

Hence, for at least one outcome the number of Hamiltonians must be at least $\frac{n!}{2^{n-1}}$.

Problem 1.14 Answer in back.

(a) Some explanation might help. The solutions in the back are somewhat confusing. There is an extra "is 1" in the third line. It first is good to observe that

$$X_1 = \sum_{i=1}^{10} Y_i \ ,$$

where Y_i are IID. (Even Y_1 is well defined this way.) Then a step is left out. Let p be the probability a 1 occurs before an even number. We develop an equation for p:

$$p = \frac{1}{6} + \frac{2}{6}p \ ,$$

so that

$$p = \frac{1}{4}.$$

Then use p to calculate $E[Y_i]$ by developing another equation. That step is given in the answers.

Problem 1.17 Answer in back.

Problem 1.18 Let N be the number of flips that are made until a string of r heads in a row. Define T as the the number of trials until the first tails. Then we have

$$\mathbf{E}[N|T=k] = \begin{cases} k + \mathbf{E}[N] & \text{if } k \le r \\ r & \text{if } k > r \ . \end{cases}$$

Using the fact that T has geometric distribution,

$$\begin{aligned}
\mathbf{E}[N] &= \mathbf{E}[\mathbf{E}[N|T]] = \sum_{k=1}^{\infty} \mathbf{E}[N|T=k]\mathbf{P}(T=k) \\
&= \sum_{k=1}^{r}(k+\mathbf{E}[N])(1-p)p^{k-1} + \sum_{k=r+1}^{\infty} r(1-p)p^{k-1} \\
&= (1-p)\sum_{k=1}^{r}(k+\mathbf{E}[N])p^{k-1} + r(1-p)\sum_{k=r+1}^{\infty} p^{k-1} \\
&= (1-p)\sum_{k=1}^{r} kp^{k-1} + \mathbf{E}[N](1-p)\sum_{k=1}^{r} p^{k-1} + rp^r \\
&= \frac{1-(r+1)p^r + rp^{r+1}}{1-p} + \mathbf{E}[N](1-p^r) + rp^r \\
&= \frac{1-p^r}{1-p} + \mathbf{E}[N](1-p^r) \\
&= \frac{1-p^r}{(1-p)p^r} \ .
\end{aligned}$$

Problem 1.20 Let L be the left hand point of the first interval. Note that $\{N(x)|L = y\} = \{1 + N(y) + N(x - y - 1)\}$. If $x > 1$,

$$\begin{aligned}
M(x) &= \mathbf{E}[N(x)] = \mathbf{E}[\mathbf{E}[N(x)|L]] \\
&= \int_0^{x-1} \mathbf{E}[N(x)|L = y] \frac{dy}{x-1} \\
&= \frac{1}{x-1} \int_0^{x-1} \mathbf{E}[1 + N(y) + N(x - y - 1)] dy \\
&= 1 + \frac{1}{x-1} \int_0^{x-1} \mathbf{E}[N(y) + N(x - y - 1)] dy \\
&= 1 + \frac{1}{x-1} \int_0^{x-1} (M(y) + M(x - y - 1)) dy \\
&= 1 + \frac{2}{x-1} \int_0^{x-1} M(y) dy.
\end{aligned}$$

Problem 1.28 The MGF is given by $\phi(t) = \frac{\lambda}{\lambda - t}$. So

$$\phi'(t) = \frac{\lambda}{(\lambda - t)^2} \quad \text{and} \quad \phi''(t) = \frac{2\lambda}{(\lambda - t)^3} .$$

Hence,

$$\begin{aligned}
\mathbf{E}[X] &= \phi'(0) = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda} \\
\mathbf{V}(X) &= \mathbf{E}[X^2] - \mathbf{E}[X]^2 \\
&= \phi''(0) - \frac{1}{\lambda^2} = \frac{2\lambda}{\lambda^3} - \frac{1}{\lambda^2} \\
&= \frac{1}{\lambda^2} .
\end{aligned}$$

Problem 1.29 The MGF of an exponential random variable, X, is $\phi_X(t) = \frac{\lambda}{\lambda - t}$. Then

$$\begin{aligned}
\phi_{\sum_n X_i}(t) &= \mathbf{E}[e^{t \sum_n X_i}] \\
&= \prod_{i=1}^{n} \mathbf{E}[e^{tX_i}] = \phi_X(t)^n \\
&= \left(\frac{\lambda}{\lambda - t} \right)^n
\end{aligned}$$

which is an MGF of an Gamma distribution with parameter (n, λ). Hence the result follows from the uniqueness of MGF.

Problem 1.31

$$\mathbf{P}(\min\{X, Y\} > a | \min\{X, Y\} = X) = \mathbf{P}(X > a | X < Y) = \frac{\mathbf{P}(a < X, X < Y)}{\mathbf{P}(X < Y)} .$$

$$\begin{aligned}\mathbf{P}(a < X, X < Y) &= \int_a^\infty \mathbf{P}(Y > X|X = x)\lambda_1 e^{-\lambda_1 x}dx = \int_a^\infty e^{-\lambda_2 x}\lambda_1 e^{-\lambda_1 x}dx \\ &= \lambda_1 \int_a^\infty e^{-(\lambda_1+\lambda_2)x}dx = \frac{\lambda_1}{\lambda_1+\lambda_2}e^{-(\lambda_1+\lambda_2}a \ ,\end{aligned}$$

$$\begin{aligned}\mathbf{P}(X < Y) &= \int_0^\infty \mathbf{P}(Y > X|Y = y)\lambda_2 e^{-\lambda_2 y}dy = \int_0^\infty (1 - e^{-\lambda_1 y})\lambda_2 e^{-\lambda_2 y}dy \\ &= 1 - \frac{\lambda_2}{\lambda_1+\lambda_2} = \frac{\lambda_1}{\lambda_1+\lambda_2} \ .\end{aligned}$$

Hence,

$$\mathbf{P}(\min\{X,Y\} > a|\min\{X,Y\} = X) = e^{-(\lambda_1+\lambda_2)a} \ .$$

Problem 1.34 Answer in back.

Problem 1.43

$$\mathbf{P}(X \geq a) = \mathbf{P}(X^t \geq a^t) \leq \frac{\mathbf{E}[X^t]}{a^t}$$

with the inequality following from the Markov inequality. Let X be exponential with rate 1, and let $a = t = n$ in the preceding, to obtain that

$$e^{-n} \leq \frac{n!}{n^n} \ .$$

(Of course, the above inequality could also be shown by noting that it is equivalent to the statement that $\mathbf{P}(Y = n) \leq 1$ where Y is Poisson with mean n.

Problem 1.23 Let $i \to j$ be the event that the particle moves from i to j in one step. Let $i \Rightarrow j$ be the event that the particle ever reaches j starting i. Conditioning on the random variable denoting the first movements of the particle, I,

(a)

$$\begin{aligned}\alpha &= \mathbf{P}(0 \Rightarrow 1) \\ &= \mathbf{E}[\mathbf{P}(0 \Rightarrow 1|I)] \\ &= \mathbf{P}(0 \to 1)\mathbf{P}(1 \Rightarrow 1) + \mathbf{P}(0 \to -1)\mathbf{P}(-1 \Rightarrow 1) \\ &= p \times 1 + (1-p)\mathbf{P}(-1 \Rightarrow 1) \\ &= p + (1-p)\mathbf{P}(-1 \Rightarrow 0, 0 \Rightarrow 1) \\ &= p + (1-p)\mathbf{P}(-1 \Rightarrow 0)\mathbf{P}(0 \Rightarrow 1) \\ &= p + (1-p)\mathbf{P}(0 \Rightarrow 1)^2 \\ &= p + (1-p)\alpha^2 \ .\end{aligned}$$

(b) Solving the previous quadratic equation, we get two solutions, 1 and $\frac{p}{1-p}$. The condition $\frac{p}{1-p} < 1$ implies $p < 1/2$. Hence if $p \geq 1/2$, α should be 1. For $p < 1/2$, the *strong law of large numbers* says that the particle ever goes to the negative infinity with probability 1. If $\alpha = 1$, then the starting position would be reached infinitely often, which contradicts to the *strong law of large numbers*. Hence

$$\alpha = \begin{cases} 1 & \text{if } p \geq 1/2 \\ \frac{p}{1-p} & \text{if } p < 1/2 \ . \end{cases}$$

(c)

$$\begin{aligned} \mathbf{P}(0 \Rightarrow n) &= \mathbf{P}(0 \Rightarrow 1) \times \cdots \times \mathbf{P}(n-1 \Rightarrow n) \\ &= \mathbf{P}(0 \Rightarrow 1) \times \cdots \times \mathbf{P}(0 \Rightarrow 1) \\ &= \mathbf{P}(0 \Rightarrow 1)^n \\ &= \alpha^n \ . \end{aligned}$$

(d)

$$\begin{aligned} \mathbf{P}(i \to i+1 | i \Rightarrow n) &= \frac{\mathbf{P}(i \to i+1, i \Rightarrow n)}{\mathbf{P}(i \Rightarrow n)} \\ &= \frac{\mathbf{P}(i \Rightarrow n | i \to i+1)\mathbf{P}(i \to i+1)}{\mathbf{P}(i \Rightarrow n)} \\ &= \frac{\mathbf{P}(i+1 \Rightarrow n)p}{\mathbf{P}(i \Rightarrow n)} \\ &= \frac{\alpha^{n-i-1}p}{\alpha^{n-i}} \\ &= \frac{p}{\alpha} \\ &= 1-p \ . \end{aligned}$$

Problem 1.24 Let $T_{(i \Rightarrow j)}$ the number of steps to reach j first time starting i. Then we have an apparent arithmetic like $T_{(-1 \Rightarrow 1)} = T_{(-1 \Rightarrow 0)} + T_{(0 \Rightarrow 1)}$ and a distributional identity like $T_{(-1 \Rightarrow 0)} \stackrel{d}{=} T_{(0 \Rightarrow 1)}$. We also know that $T_{(-1 \Rightarrow 0)}$ and $T_{(0 \Rightarrow 1)}$ are independent because of the independence of every transition. That is, $T_{(-1 \Rightarrow 0)}$ and $T_{(0 \Rightarrow 1)}$ are iid. Using the notation in the book, $T \equiv T_{(0 \Rightarrow 1)}$,

$$\begin{aligned} \mathbf{E}[T_{(-1 \Rightarrow 1)}] &= 2\mathbf{E}[T] \ , \\ \mathbf{V}(T_{(-1 \Rightarrow 1)}) &= 2\mathbf{V}(T) \ . \end{aligned}$$

Let the random variable X denote the particle's location after the first move.

(a) Conditioning on X gives

$$\begin{aligned}
\mathbf{E}[T] &= \mathbf{E}[\mathbf{E}[T|X]] \\
&= \mathbf{E}[T|X=1]\mathbf{P}(X=1) + \mathbf{E}[T|X=-1]\mathbf{P}(X=-1) \\
&= 1 \times p + (1 + \mathbf{E}[T_{(-1 \Rightarrow 1)}])(1-p) \\
&= 1 + 2(1-p)\mathbf{E}[T] \ .
\end{aligned}$$

Hence, $\mathbf{E}[T] = \infty$ if $p \leq 1/2$. If we can show that $\mathbf{E}[T] < \infty$ when $p > 1/2$, we obtain in this case that

$$\mathbf{E}[T] = \frac{1}{2p-1} \ .$$

Now let's show that $\mathbf{E}[T] < \infty$ if $p > 1/2$: Let $p^{(n)}$ denotes the probability that the particle reaches 1 by n-transitions starting 0. Then n should be odd. That is, only $p^{(2n+1)}$ is nonzero. Now we have an upper bound on this probability:

$$p^{(2n+1)} \leq \binom{2n}{n} p[p(1-p)]^n \sim p\frac{[4p(1-p)]^n}{\sqrt{\pi n}}$$

using an approximation, due to Stirling, which asserts that

$$n! \sim n^{n+1/2}e^{-n}\sqrt{2\pi}$$

where $a_n \sim b_n$ denotes $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$. Now it is easy to verify that if $a_n \sim b_n$, then $\sum_n a_n < \infty$ if, and only if, $\sum_n b_n < \infty$. Hence $\mathbf{E}[T] < \infty$ if

$$\sum_{n=0}^{\infty} (2n+1)p\frac{[4p(1-p)]^n}{\sqrt{\pi n}} < \infty$$

which is true if $4p(1-p) < 1$ or $p \neq 1/2$.

(b) Noting that $T|\{X=1\} = 1$ and $T|\{X=-1\} = 1 + T_{(-1\Rightarrow 1)}$ which is 1 plus the *convolution* of two independent random variables both having the distribution of T. Therefore,

$$\mathbf{E}[T|X=1] = 1, \quad \mathbf{E}[T|X=-1] = 1 + 2\mathbf{E}[T]$$
$$\mathbf{V}(T|X=1) = 0, \quad \mathbf{V}(T|X=-1) = 2\mathbf{V}(T)$$

and thus

$$\begin{aligned}
\mathbf{V}(\mathbf{E}[T|X]) &= \mathbf{V}(\mathbf{E}[T|X]-1) = 4\mathbf{E}[T]^2 p(1-p) = \frac{4p(1-p)}{(2p-1)^2} \\
\mathbf{E}[\mathbf{V}(T|X)] &= 2(1-p)\mathbf{V}(T) \ .
\end{aligned}$$

By the conditional variance formula

$$\mathbf{V}(T) = 2(1-p)\mathbf{V}(T) + \frac{4p(1-p)}{(2p-1)^2}$$

which gives the result.

(c) $T_{(0\Rightarrow n)} = T_{(0\Rightarrow 1)} + \cdots + T_{(n-1\Rightarrow n)} = \sum_{i=1}^{n} T_i$ where T_i are iid having distribution of T. Hence

$$\mathbf{E}[T_{(0\Rightarrow n)}] = n\mathbf{E}[T] \ .$$

(d) By the same reasoning as in (c),

$$\mathbf{V}(T_{(0\Rightarrow n)}) = n\mathbf{V}(T) \ .$$

IEOR 6711: Stochastic Models I

Professor Whitt

Solutions to Homework Assignment 3

Problem 2.1 The conditions (i) and (ii) of definition 2.1.2 are apparent from the definition 2.1.1. Hence it is sufficient to show that definition 2.1.1 implies the last two conditions of definition 2.1.2.

- $\mathbf{P}(N(h)=1)=\lambda h+o(h)$:

$$\lim_{h\to 0}\frac{\mathbf{P}(N(h)=1)-\lambda h}{h}=\lim_{h\to 0}\frac{e^{-\lambda h}\lambda h-\lambda h}{h}=\lim_{h\to 0}\left(e^{-\lambda h}-1\right)\lambda=0\ .$$

- $\mathbf{P}(N(h)\geq 2)=o(h)$:

$$\begin{aligned}&\lim_{h\to 0}\frac{\mathbf{P}(N(h)\geq 2)}{h}=\lim_{h\to 0}\frac{1-e^{-\lambda h}-e^{-\lambda h}\lambda h}{h}=\lim_{h\to 0}\frac{1-e^{-\lambda h}}{h}-\lim_{h\to 0}e^{\lambda h}\lambda\\ =\;&\frac{1-(1-\lambda h+o(h))}{h}-(1-o(h))\lambda=\lambda+\frac{o(h)}{h}-\lambda+o(h)\to 0.\end{aligned}$$

Or using $e^{ax}=1+ax+o(x)$, $o(x)\times o(x)=o(x)$, and $f(x)\times o(x)=o(x)$ for any $f(x)$ satisfying $\lim_{x\to 0}f(x)$ is finite,

- $\mathbf{P}(N(h)=1)=e^{-\lambda h}\lambda h=(1-\lambda h+o(h))\lambda h=\lambda h+o(h)$.
- $\mathbf{P}(N(h)\geq 2)=1-e^{-\lambda h}-e^{-\lambda h}\lambda h=1-(1+\lambda h)(1-\lambda h+o(h))=\lambda^2h^2+o(h)=o(h)$.

Problem 2.2 For $s<t$,

$$\begin{aligned}\mathbf{P}(N(s)=k|N(t)=n)&=\frac{\mathbf{P}(N(s)=k,N(t)=n)}{\mathbf{P}(N(t)=n)}=\frac{\mathbf{P}(N(s)=k,N(t)-N(s)=n-k)}{\mathbf{P}(N(t)=n)}\\ &=\frac{\mathbf{P}(N(s)=k)\mathbf{P}(N(t)-N(s)=n-k)}{\mathbf{P}(N(t)=n)}\\ &=\frac{\mathbf{P}(N(s)=k)\mathbf{P}(N(t-s)=n-k)}{\mathbf{P}(N(t)=n)}\\ &=\left(\frac{e^{-\lambda s}(\lambda s)^k}{k!}\right)\left(\frac{e^{-\lambda(t-s)}(\lambda(t-s))^{(n-k)}}{(n-k)!}\right)\left(\frac{e^{-\lambda t}(\lambda t)^n}{n!}\right)^{-1}\\ &=\frac{n!}{k!(n-k)!}\frac{s^k(t-s)^{n-k}}{t^n}\\ &=\binom{n}{k}\left(\frac{s}{t}\right)^k\left(1-\frac{s}{t}\right)^{n-k}\ .\end{aligned}$$

Problem 2.4 Let $\{X(t) : t \geq 0\}$ be a stochastic process having stationary independent increments and $X(0) = 0$. (People call it *Levy* process.) Two typical *Levy* processes are *Poisson* and *Brownian motion* processes. They are representatives of purely discrete and purely continuous continuous time stochastic processes, respectively. Furthermore, it is not easy to find any non-trivial Levy process except them. Now let's try to express $\mathbf{E}[X(t)X(t+s)]$ by moments of X using only the properties of Levy process.

$$\begin{aligned}
\mathbf{E}[X(t)X(t+s)] &= \mathbf{E}[X(t)(X(t+s) - X(t) + X(t))] \\
&= \mathbf{E}[X(t)(X(t+s) - X(t)) + X(t)^2] \\
&= \mathbf{E}[X(t)(X(t+s) - X(t))] + \mathbf{E}[X(t)^2] \\
&= \mathbf{E}[X(t)]\mathbf{E}[(X(t+s) - X(t))] + \mathbf{E}[X(t)^2] \quad \text{by independent increment} \\
&= \mathbf{E}[X(t)]\mathbf{E}[(X(s))] + \mathbf{E}[X(t)^2] \quad \text{by stationary increment}
\end{aligned}$$

Now return to our original process, Poisson process. By substituting $\mathbf{E}[N(t)] = \lambda t$, $\mathbf{E}[N(t)^2] = \lambda t + (\lambda t)^2$,

$$\mathbf{E}[N(t)N(t+s)] = \lambda^2 st + \lambda t + \lambda^2 t^2 \ .$$

A digression : if $X(t) \sim \text{Normal}(0, t)$, what is the result? This is the Brownian motion case.

Problem 2.5

- $\{N_1(t) + N_2(t), t \geq 0\}$ is a Poisson process with rate $\lambda_1 + \lambda_2$.
  Axioms (i) and (ii) of definition e.1.2 easily follow. Letting $N(t) = N_1(t) + N_2(t)$,

  $$\begin{aligned}
  \mathbf{P}(N(h) = 1) &= \mathbf{P}(N_1(h) = 1, N_2(h) = 0) + \mathbf{P}(N_1(h) = 0, N_2(h) = 1) \\
  &= \lambda_1 h(1 - \lambda_2 h) + \lambda_2 h(1 - \lambda_1 h) + o(h) \\
  &= (\lambda_1 + \lambda_2)h + o(h)
  \end{aligned}$$

  and

  $$\begin{aligned}
  \mathbf{P}(N(h) = 2) &= \mathbf{P}(N_1(h) = 1, N_2(h) = 1) \\
  &= (\lambda_1 h + o(h))(\lambda_2 h + o(h)) \\
  &= \lambda_1 \lambda_2 h^2 + o(h) = o(h) \ .
  \end{aligned}$$

- The probability that the first event of the combined process comes from $\{N_1(t), t \geq 0\}$ is $\lambda_1/(\lambda_1 + \lambda_2)$, independently of the time of the event. Let X_i and Y_i are the i-th inter arrival times of N_1 and N_2, respectively. Then

  $$\begin{aligned}
  \mathbf{P}(\text{first from } N_1 | \text{first at } t) &= \mathbf{P}(X_1 < Y_1 | \min\{X_1, Y_1\} = t) \\
  &= \frac{\lambda_1}{\lambda_1 + \lambda_2}
  \end{aligned}$$

  where the last equality comes from our old homework 1.1.34.

Problems 2.6–2.9 Answers in back of the book.

Problem 2.10 **(a)** First note that the time until next bus arrival follows exponential distribution with rate λ by the *memoryless* property of exponential distribution. Let X be the time until next bus arrival. Then T, the random variable representing the time spent to reach home is

$$\begin{aligned} T &= \begin{cases} X+R & \text{if } X \le s \,, \\ s+W & \text{if } X > s \end{cases} \\ &= (X+R)\mathbf{1}_{\{X\le s\}} + (s+W)\mathbf{1}_{\{X>s\}} \,. \end{aligned}$$

Hence

$$\begin{aligned} \mathbf{E}[T] &= \mathbf{E}[(X+R)\mathbf{1}_{\{X\le s\}}] + \mathbf{E}[(s+W)\mathbf{1}_{\{X>s\}}] \\ &= \mathbf{E}[X\mathbf{1}_{\{X\le s\}}] + R\mathbf{E}[\mathbf{1}_{\{X\le s\}}] + (s+W)\mathbf{E}[\mathbf{1}_{\{X>s\}}] \\ &= \int_0^s x\lambda e^{-\lambda x}dx + R\mathbf{P}(X\le s) + (s+W)\mathbf{P}(X>s) \\ &= \int_0^s x\lambda e^{-\lambda x}dx + R(1-e^{-\lambda s}) + (s+W)e^{-\lambda s} \\ &= -xe^{-\lambda x}\Big|_0^s - \frac{1}{\lambda}e^{-\lambda x}\Big|_0^s + R + (s+W-R)e^{-\lambda s} \\ &= \frac{1}{\lambda}(1-e^{-\lambda s}) - se^{-\lambda s} + R + (s+W-R)e^{-\lambda s} \\ &= \frac{1}{\lambda} + R + \left(W - R - \frac{1}{\lambda}\right)e^{-\lambda s}. \end{aligned}$$

(b) Considering

$$\frac{d}{ds}\mathbf{E}[T] = (1-\lambda(W-R))e^{-\lambda s} \begin{cases} >0 & \text{if } W < \frac{1}{\lambda}+R \,, \\ =0 & \text{if } W = \frac{1}{\lambda}+R \,, \\ <0 & \text{if } W > \frac{1}{\lambda}+R \,, \end{cases}$$

we get

$$\text{argmin}_{0\le s<\infty}\mathbf{E}[T] = \begin{cases} 0 & \text{if } W < \frac{1}{\lambda}+R \,, \\ \text{any number } \in [0,\infty) & \text{if } W = \frac{1}{\lambda}+R \,, \\ \infty & \text{if } W > \frac{1}{\lambda}+R \,, \end{cases}$$

(c) Since the time until the bus arrives is exponential, it follows by the *memoryless* property that if it is optimal to wait any time then one should always continue to wait for the bus.

Problem 2.11 Conditioning on the time of the next car yields

$$\mathbf{E}[\text{wait}] = \int_0^\infty \mathbf{E}[\text{wait}|\text{car at } x]\lambda e^{-\lambda x}dx \,.$$

Now,

$$\mathbf{E}[\text{wait}|\text{car at } x] = \begin{cases} x + \mathbf{E}[\text{wait}] & \text{if } x < T \ , \\ 0 & \text{if } x \geq T \end{cases}$$

and so

$$\mathbf{E}[\text{wait}] = \int_0^T x\lambda e^{-\lambda x} dx + \mathbf{E}[\text{wait}](1 - e^{-\lambda T})$$

or

$$\mathbf{E}[\text{wait}] = \frac{1}{\lambda} e^{-\lambda T} - T - \frac{1}{\lambda} \ .$$

Problem 2.13 First note that T is (unconditionally) exponential with rate λp, N is geometric with parameter p, and the distribution of T given that $N = n$ is gamma with parameter n and λ, we obtain

$$\begin{aligned} \mathbf{P}(N = n | t - \epsilon < T \leq t) &= \frac{\mathbf{P}(t - \epsilon < T \leq t | N = n)}{\mathbf{P}(t - \epsilon < T \leq t)} \\ &\simeq \frac{\epsilon \lambda e^{-\lambda t} (\lambda t)^{n-1}}{(n-1)!} \frac{p(1-p)^{n-1}}{\epsilon \ \lambda p \ e^{-\lambda p t}} \\ &= \frac{e^{-\lambda t(1-p)} [\lambda t(1-p)]^{n-1}}{(n-1)!} \ . \end{aligned}$$

Hence, given that $T = t$, N has the distribution of $X + 1$, where X is a Poisson random variable with mean $\lambda t(1-p)$. A simpler argument is to note that the occurrences of failure causing shocks and non-failure causing shocks are independent Poisson processes. Hence, the number of non-failure causing shocks by time t is Poisson with mean $\lambda(1-p)t$, independent of the event that the first failure shock occurred at that time.

IEOR 6711: Stochastic Models I, Professor Whitt

Solutions to Homework Assignment 4.

Problem 2.14 Let us use D_j instead of O_j for the number of people getting off at floor j. Let $D_{i,j}$ denote the number of people that get on at floor i and get off at floor j. First, $D_{i,j}$ is an independent thinning of N_i with

$$N_i = \sum_{j=i+1}^{n} D_{i,j}$$

so $D_{i,j}$ for different j are independent Poisson random variables. But then N_i are independent for different i. Consequently $D_{i,j}$, as i and j both vary, are independent Poisson random variables, with mean $\lambda_i p_{i,j}$.

(a)-(c) Clearly,

$$D_j = \sum_{i=0}^{j-1} D_{i,j}$$

so that D_j is the sum of independent Poisson random variables, so itself must have a Poisson distribution with mean

$$E[D_j] = \sum_{i=0}^{j-1} E[D_{i,j}] = \sum_{i=0}^{j-1} \lambda_i p_{i,j} \ .$$

Moreover, Since D_{j_1} and D_{j_2} for $j_1 \neq j_2$ have no variables in common in the sums, these two Poisson random variables are independent Poisson random variables.

Problem 2.16 For fixed j, let

$$I_i = \begin{cases} 1 & \text{if outcome } i \text{ occurs } j \text{ times,} \\ 0 & \text{otherwise ,} \end{cases}$$

and note that $I_i, i = 1, \cdots, n$ are independent since the number of type i outcomes, $i = 1, \cdots, n$, will be independent. (If we think that there is a Poisson process with rate λ and we count on $[0, 1]$, then the n-types of events are independent by proposition 2.3.2 and hence so are I_i's.) Writing

$$X_j = \sum_{i=1}^{n} I_j$$

we have

$$\mathbf{E}[X_j] = \sum_{i=1}^{n} \mathbf{P}(I_i = 1)$$

and

$$\mathbf{V}ar[X_j] = \sum_{i=1}^{n} \mathbf{P}(I_i = 1)(1 - \mathbf{P}(I_i = 1)) \ .$$

As the number of times outcome i results is Poisson with mean λP_i we have that

$$\mathbf{P}(I_i = 1) = \frac{e^{-\lambda P_i}(\lambda P_i)^j}{j!}$$

and so

$$\mathbf{E}[X_j] = \sum_{i=1}^{n} \frac{e^{-\lambda P_i}(\lambda P_i)^j}{j!} \ ,$$

$$\mathbf{V}ar[X_j] = \mathbf{E}[X_j] - \sum_{i=1}^{n} \frac{e^{-2\lambda P_i}(\lambda P_i)^{2j}}{(j!)^2} \ .$$

Problem 2.17 **(a)** $\{X_{(i)} = x\}$ implies $i-1$ X_j's are less than x and $n-i$ X_j's are greater than x and one is equal to x. Hence

$$\begin{aligned} f_{X_{(i)}}(x) &= \frac{\mathbf{P}(X_{(i)} \in (x, x+dx))}{dx} \\ &= \frac{n!}{(i-1)!1!(n-i)!} \frac{(F(x))^{i-1} f(x) dx (\bar{F}(x+dx))^{n-i}}{dx} \\ &= \frac{n!}{(i-1)!(n-i)!}(F(x))^{i-1} f(x) (\bar{F}(x))^{n-i} \ . \end{aligned}$$

(b) At least i.

(c)

$$\begin{aligned} \mathbf{P}(X_{(i)} \leq x) &= \mathbf{P}(i \text{ or more } X_j's \text{ are less than or equal to } x) \\ &= \sum_{k=i}^{n} \binom{n}{k} (F(x))^k (\bar{F}(x))^{n-k} \ . \end{aligned}$$

(d)

$$\begin{aligned} \mathbf{P}(X_{(i)} \leq x) &= \sum_{k=i}^{n} \binom{n}{k} (F(x))^k (\bar{F}(x))^{n-k} \\ &= \int_0^x f_{X_{(i)}}(t) dt = \int_0^x \frac{n!}{(i-1)!(n-i)!}(F(t))^{i-1} f(t) (\bar{F}(t))^{n-i} dt \\ &= \int_0^x \frac{n!}{(i-1)!(n-i)!}(F(t))^{i-1} (\bar{F}(t))^{n-i} dF(t) \end{aligned}$$

and substituting $F(x)$ by y gives

$$\sum_{k=i}^{n} \binom{n}{k} y^k (1-y)^{n-k} = \int_0^y \frac{n!}{(i-1)!(n-i)!} x^{i-1} (1-x)^{n-i} dx \ .$$

(e) First note that $S_{(i)}|\{N(t)=n\} \sim \text{Uniform}(0,t)$ if $i \le n$. Hence from (a),

$$\begin{aligned}
\mathbf{E}[X_{(i)}] &= \int_0^t x f_{X_{(i)}}(x)dx \\
&= \int_0^t x \frac{n!}{(i-1)!(n-i)!}\left(\frac{x}{t}\right)^{i-1}\left(1-\frac{x}{t}\right)^{n-i}\frac{dx}{t} \\
&= \frac{i}{n+1}t\int_0^t \frac{(n+1)!}{i!(n-i)!}\left(\frac{x}{t}\right)^{i}\left(1-\frac{x}{t}\right)^{n-i}\frac{dx}{t} \\
&= \frac{i}{n+1}t
\end{aligned}$$

if $i \le n$. For $i > n$, using memoryless property

$$\begin{aligned}
\mathbf{E}[S_i|N(t)=n] &= t+\mathbf{E}[S_i-t|N(t)=n] = t+\mathbf{E}[X_{n+1}+\cdots+X_i] \\
&= t+\frac{i-n}{\lambda}\ .
\end{aligned}$$

Hence

$$\mathbf{E}[S_i|N(t)=n] = \begin{cases} \frac{i}{n+1}t & \text{if } i \le n, \\ t+\frac{i-n}{\lambda} & \text{if } i > n\ . \end{cases}$$

Problem 2.18 As the joint density of $U_{(1)}, U_{(2)}, \cdots, U_{(n)}$ is $f(u_1,\cdots,u_n) = 1/n!$, $0 < u_1 < \cdots < u_n < 1$, the conditional density is

$$\begin{aligned}
f(u_1,\cdots,u_{n-1},y|U_{(n)}=y) &= \frac{f(u_1,\cdots,u_{n-1},y|U_{(n)}=y)}{f_{U_{(n)}(y)}} \\
&= \frac{n!}{ny^{n-1}} \\
&= \frac{(n-1)!}{y^{n-1}}\ , \quad 0 < u_1 < \cdots < u_{n-1} < y
\end{aligned}$$

which proves the result.

Problem 2.19 Observe that this is essentially an $M/G/\infty$ queue problem, so we can apply the "Physics" paper. We could also easily generalize the results to a nonhomogeneous Poisson arrival process.

It is not clear if the service times apply to each bus or to each customer. We assume the service times apply to the buses, rather than the customers, but the overall mean is unaffected by that result. Presumably the service times are meant to be IID. We need to assume that too.

(a) We first want to apply the splitting or thinning property. We can split the original Poisson process according to the number of customers on the bus. A bus is of type j if the bus contains j customers. Thus, by the splitting property, the overall arrival process is the superposition of infinitely many independent Poisson processes. Poisson process j has arrival rate $\lambda\alpha_j$. Thus the whole system behaves as infinitely many independent $M/G/\infty$ queues. Arrival process j has arrival rate $\lambda\alpha_j$.

By the basic $M/G/\infty$ theory, the number of buses of type j to depart from the system has a Poisson distribution with mean $m_j(t) = \lambda\alpha_j \int_0^t G(s)\,ds, \quad t \geq 0$. Thus the total number of *buses* to depart in $[0,t]$ also has a Poisson distribution with a mean equal to the sum of the means, i.e., $m(t) = \sum_{j=1}^{\infty} m_j(t)$.

However, we are asked about the number of *customers*. The number of customers on a bus of type j (which contains exactly j customers) is j times the Poisson random variable with mean $m_j(t)$. Thus the overall mean is

$$E[X(t)] = \sum_{j=1}^{\infty} j m_j(t) = \lambda \int_0^t G(s)\,ds \sum_{j=1}^{\infty} j\alpha_j \ . \tag{1}$$

(b) If the service times apply to buses, then batches of customers depart together, so that the departure process of customers cannot be Poisson. Even if the service times are associated with customers, the departure process is not Poisson. That is easy to show if the service time distribution G is in fact deterministic. Then the customers that arrive together on the same bus will also depart together. Otherwise the non-Poisson character of the departure process is harder to show. But since the arrival process is a batch Poisson process, we should not expect the departure process to be a Poisson process.

We now give an alternative direct derivation of part (a): Let N_i be the number of customers in $i-$th busload. Then $\mathbf{E}[N_i] = \sum_{k=1}^{\infty} k\alpha_k$ since $\mathbf{P}(N_i = k) = \alpha_k$. Let the indicator, $I_{i,j}$ for $1 \leq j \leq N_i$, denote whether the $j - th$ customer in $i-$th busload finishes his/her service at time t. Then $X(t) = \sum_{i=0}^{N(t)} \sum_{j=1}^{N_i} I_{i,j}$.

(a) First note that $\mathbf{E}[\mathbf{E}[X|Y,Z]|Z] = \mathbf{E}[X|Z]$ and $\mathbf{E}[\mathbf{E}[X|Z]|Y,Z] = \mathbf{E}[X|Z]$ which mean *smaller information wins always in double conditioning!* (You might prove it right now, or may consult the equation (6.1.2) in page 296.) Also, we have

$$\mathbf{E}[I_{i,j}|N(t) = n] = \int_0^t G(t-s)\frac{1}{t}ds = \frac{1}{t}\int_0^t G(s)ds \equiv p \ .$$

$$\begin{aligned}
\mathbf{E}[X(t)] &= \mathbf{E}\left[\sum_{i=0}^{N(t)} \sum_{j=1}^{N_i} I_{i,j}\right] \\
&= \mathbf{E}\left[\mathbf{E}\left[\sum_{i=0}^{N(t)} \sum_{j=1}^{N_i} I_{i,j} \middle| N(t)\right]\right] \\
&= \mathbf{E}\left[\sum_{i=0}^{N(t)} \mathbf{E}\left[\sum_{j=1}^{N_i} I_{i,j} \middle| N(t)\right]\right]
\end{aligned}$$

$$
\begin{aligned}
&= \mathbf{E}\left[\sum_{i=0}^{N(t)} \mathbf{E}\left[\mathbf{E}\left[\sum_{j=1}^{N_i} I_{i,j}\middle| N(t), N_i\right]\middle| N(t)\right]\right] \quad \text{(smaller information wins)} \\
&= \mathbf{E}\left[\sum_{i=0}^{N(t)} \mathbf{E}\left[\sum_{j=1}^{N_i} \mathbf{E}\left[I_{i,j}\middle| N(t), N_i\right]\middle| N(t)\right]\right] \\
&= \mathbf{E}\left[\sum_{i=0}^{N(t)} \mathbf{E}\left[N_i\mathbf{E}\left[I_{i,j}\middle| N(t)\right]\middle| N(t)\right]\right] \quad (I_{i,j} \text{ are independent of } N_i) \\
&= \mathbf{E}\left[\sum_{i=0}^{N(t)} p\mathbf{E}\left[N_i\middle| N(t)\right]\right] \\
&= \mathbf{E}\left[\sum_{i=0}^{N(t)} p\mathbf{E}\left[N_i\right]\right] \quad (N(t) \text{ is independent of } N_i) \\
&= p\mathbf{E}\left[N_i\right]\mathbf{E}\left[N(t)\right] \\
&= \lambda t p\mathbf{E}\left[N_i\right] \\
&= \lambda \sum_{j=1}^{\infty} j\alpha_j \int_0^t G(s)ds \ .
\end{aligned}
$$

Problem 2.20 The key thing here is to apply the conditional distribution of arrival times given a Poisson number of events in an interval. Under the conditioning, the unordered arrival times are distributed as IID uniform random variables. We apply this representation in the first step.

Assume that $n = \sum_{i=1}^{k} n_i$. Note that given $N(t) = n$ the unordered set of arrival times are independent uniform $(0, t)$. The probability that an arbitrary event is type i is thus

$$
p_i \equiv \frac{1}{t}\int_0^t P_i(x)dx \ .
$$

$$
\begin{aligned}
\mathbf{P}(N_i(t) = n_i, i = 1, \cdots, k) &= \mathbf{P}(N_i(t) = n_i, i = 1, \cdots, k | N(t) = n)\frac{e^{-\lambda t}(\lambda t)^n}{n!} \\
&= \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k} \frac{e^{-\lambda t}(\lambda t)^n}{n!} \quad \text{(multinomial)} \\
&= \frac{e^{-\lambda p_1 t}(\lambda p_1 t)^{n_1}}{n_1!} \cdots \frac{e^{-\lambda p_k t}(\lambda p_k t)^{n_k}}{n_k!} \ .
\end{aligned}
$$

Hence we're done.(Why?)

Problem 2.21 The key idea here is to apply Problem 2.20. In this problem, the state of an individual varies, which might cause trouble to put into the setting of Problem 2.20. But, if we fix the observation time t, then the probability that an individual who arrived at time s is in state i at time t is just $\alpha_i(t-s)$ and if we define $P_i(s) \equiv \alpha_i(t-s)$, then we can apply

Problem 2.20. Furthermore,

$$\begin{aligned}
\mathbf{E}[N_i(t)] &= \lambda \int_0^t P_i(x)dx \\
&= \lambda \int_0^t \alpha_i(t-x)dx \\
&= \lambda \int_0^t \alpha_i(y)dy \\
&= \lambda \int_0^t \mathbf{E}[I_i(y)]dy \\
&= \lambda \mathbf{E}\left[\int_0^t I_i(y)dy\right]
\end{aligned}$$

which leads to the desired interpretation of the mean.

Now, are we done? Unless you are an advanced reader, not yet. The problem doesn't mention on the number of states i. Hence it is possible that there are countably many states and the random variables, $N_i(t)$, are countably many. Then what is the definition of independence among countably many random variables. One general fact in mathematics and probability is that when we define a property among countably many objects, we require that the property holds among any finitely many ones. (Recall the definition of basis in infinite dimensional vector space.) So, to prove the independence of $N_i(t)$ it is sufficient to show it under finite states.

IEOR 6711: Stochastic Models I, Professor Whitt

Solutions to Homework Assignment 5, not to be turned in.

Problem 2.25 Just as in Problem 2.20, the key thing here is to apply the conditional distribution of arrival times given a Poisson number of events in an interval. Under the conditioning, the unordered arrival times are distributed as IID uniform random variables. We apply this representation.

Let X be a contribution by an event occurred at random time distributed by uniform$(0,t)$. If we denote the occurrence time of the event by U (which is uniform $(0,t)$), then

$$\begin{aligned}
\mathbf{P}(X \le x) &= \mathbf{E}[\mathbf{P}(X \le x|U)] \\
&= \mathbf{E}[F_U(x)] \\
&= \int_0^t F_u(x)\frac{1}{t}du \\
&= \frac{1}{t}\int_0^t F_u(x)du \ .
\end{aligned}$$

Hence, if we define that X_T is a contribution of an event occurred at random time T and $U_1, \cdots, U_k, \cdots$ are random samples from uniform$(0,t)$, then $X \stackrel{d}{=} X_{U_i}$ and

$$\begin{aligned}
W &\equiv \sum_{i=1}^{N(t)} X_{U_i} \\
&= \sum_{i=1}^{N} X_{U_i} \\
&= \sum_{i=1}^{N} X_i
\end{aligned}$$

where N is a Poisson random variable with mean λt and X_i are IID with distribution given above.

Problem 2.30 **(a)** No. It should be intuitively clear that the answer is indeed no, but providing a proof is not so easy. It is not hard to make a proof in special cases. For example, when $\lambda(t)$ is a monotonically (strictly) decreasing function. Then the larger T_1, the larger T_2. However, the variables are *never* independent if the arrival process is nonhomogeneous. That is harder to prove. We will not really try. We will be impressed if you work that out.

(b) No. In the special case above, T_1 should be larger than T_2.

(c)

$$\begin{aligned}
\mathbf{P}(T_1 \ge t) &= \mathbf{P}(N(t) = 0) \\
&= e^{-\int_0^t \lambda(s)ds}
\end{aligned}$$

(d)

$$
\begin{aligned}
\mathbf{P}(T_2 \geq t) &= \mathbf{E}[\mathbf{P}(T_2 \geq t | T_1)] \\
&= \int_0^\infty \mathbf{P}(T_2 \geq t | T_1 = s) f_{T_1}(s) ds \\
&= \int_0^\infty e^{-\int_s^{t+s} \lambda(u)du} \lambda(s) e^{-\int_0^s \lambda(u)du} ds \\
&= \int_0^\infty e^{-\int_0^{t+s} \lambda(u)du} \lambda(s) ds \\
&= \int_0^\infty e^{-m(t+s)} \lambda(s) ds \ .
\end{aligned}
$$

Problem 2.31 A key initial step is to observe that the mean function $m(t)$ is a strictly increasing continuous function, by virtue of the assumptions. (The assumption $\lambda(\cdot) > 0$ is important.) Then m has a unique well defined inverse m^{-1}, which itself has a unique well defined inverse m.

For the rest of the way, we check the parts of Definition 2.1.1 one by one.

(i) $N^*(0) = N(m^{-1}(0)) = N(0) = 0$.

(ii) For $t > 0$ and $s > 0$, $m^{-1}(t+s) > m^{-1}(t) > 0$. $N^*(t+s) - N^*(t) = N(m^{-1}(t+s)) - N(m^{-1}(t))$ and $N(m^{-1}(t)) = N^*(t)$ are independent since $(m^{-1}(t), m^{-1}(t+s)]$ and $[0, m^{-1}(t)]$ are non-overlapping. So $N^*(\cdot)$ increases independently.

(iii) For $s, t \geq 0$,

$$
\begin{aligned}
& \mathbf{P}(N^*(t+s) - N^*(s) = n) \\
= \ & \mathbf{P}(N(m^{-1}(t+s)) - N(m^{-1}(s)) = n)) \\
= \ & e^{-m(m^{-1}(t+s)) + m(m^{-1}(s))} \frac{[m(m^{-1}(t+s)) - m(m^{-1}(s))]^n}{n!} \\
= \ & \frac{e^{-t} t^n}{n!}
\end{aligned}
$$

Problem 2.32 (a) The easy way to proceed is to apply Problem 2.31 above. To do so, we need to assume that $m(t)$ is continuous and strictly increasing. So assume that is the case. Then, from Problem 2.31, $N^*(m(t)) = N(t)$ and $N^*(\cdot)$ is a Poisson process with rate 1. Hence if we set the unordered arrival times of $N(\cdot)$ by $V_1, \cdots, V_n$, then $U_1 = m(V_1), \cdots, U_n = m(V_n)$ are those of $N^*(\cdot)$. But we know that the unordered arrival times $U_1, \cdots, U_n$ of $N^*(\cdot)$ given that $N^*(m(t)) = n$ are random samples from uniform$(0, m(t))$.

$$
\begin{aligned}
\mathbf{P}(V_i \leq x) &= \mathbf{P}(m(V_i) \leq m(x)) \\
&= \mathbf{P}(U_i \leq m(x)) \\
&= \frac{\min\{m(x), m(t)\}}{m(t)} \ .
\end{aligned}
$$

We can treat the general case of m as the limit of strictly increasing m. We thus obtain the formula above, by taking the limit.

(b) Note this question is exactly an $M_t/G/\infty$ queue question, so we can apply Theorem 1 in the "Physics" paper. The number of workers out has a Poisson distribution with the mean given by the mean formula $Mean(t) \equiv p = \int_0^t \lambda(x)\bar{F}(t-x)\,dx$. The Poisson distribution property implies that the variance equals the mean.

Below we give a direct derivation: Let $N(t)$ denote the number of accidents by time t. Let I be the indicator representing that an injured person is out of work at time t. Let V be the time of accident.

$$\begin{aligned}
\mathbf{P}(I=1|N(t)) &= \mathbf{E}[\mathbf{P}(I=1|N(t),V)|N(t)] \\
&= \int_0^t \bar{F}(t-x)\frac{dm(x)}{m(t)} \\
&= \frac{1}{m(t)}\int_0^t \bar{F}(t-x)\lambda(x)dx \\
&= \frac{p}{m(t)}\ .
\end{aligned}$$

Now $X(t) = \sum_{i=1}^{N(t)} I_i$ and $X(t)|N(t)$ is a binomial$(N(t), \frac{p}{m(t)})$. Hence

$$\mathbf{E}[X(t)|N(t)] = N(t)\frac{p}{m(t)}$$

and

$$\mathbf{V}ar[X(t)|N(t)] = N(t)\frac{p}{m(t)}\left(1-\frac{p}{m(t)}\right)\ .$$

Therefore

$$\mathbf{E}(X(t)) = \mathbf{E}[N(t)]\frac{p}{m(t)} = p$$

and

$$\begin{aligned}
\mathbf{V}ar[X(t)] &= \mathbf{V}ar(\mathbf{E}[X(t)|N(t)]) + \mathbf{E}[\mathbf{V}ar(X(t)|N(t))] \\
&= \mathbf{V}ar(N(t))\left(\frac{p}{m(t)}\right)^2 + \mathbf{E}[N(t)]\frac{p}{m(t)}\left(1-\frac{p}{m(t)}\right) \\
&= \frac{p^2}{m(t)} + p\left(1-\frac{p}{m(t)}\right) \\
&= p\ .
\end{aligned}$$

Problem 2.33 Here we have a Poisson random measure problem.

Let the reference point be x and $B(x,r)$ represent the circular region of radius r with center at x.

(a)

$$\begin{aligned}\mathbf{P}(X > t) &= \mathbf{P}(\text{no event in } B(x,t)) \\ &= e^{-\lambda\pi t^2} .\end{aligned}$$

(b)

$$\begin{aligned}\mathbf{E}[X] &= \int_0^\infty \mathbf{P}(X > t)dt \\ &= \int_0^\infty e^{-\lambda\pi t^2} dt \\ &= \sqrt{\frac{1}{\lambda}} \int_0^\infty \frac{1}{\sqrt{2\pi\frac{1}{2\lambda\pi}}} e^{-\lambda\pi t^2} dt \\ &= \sqrt{\frac{1}{\lambda}} \times \frac{1}{2} \\ &= \frac{1}{2\sqrt{\lambda}} .\end{aligned}$$

(c) Since $\pi R_i^2 - \pi R_{i-1}^2$ is the area of the region between the circle of radius R_i and the one of radius of R_{i-1}, we have

$$\begin{aligned}\mathbf{P}(\pi R_i^2 - \pi R_{i-1}^2 > a | R_j, j < i) &= \mathbf{P}(\text{no event in the area a}) \\ &= e^{-\lambda a} .\end{aligned}$$

Problem 2.35 **(a)** No, because knowing the number of events in some interval is informative about the value of τ, which gives information about the distribution of the number of events in a non-overlapping interval.

(b) Yes, because the occurrences in any non-overlapping intervals are independent.

Problem 2.39 $s < t$. Let $X(t) = \sum_{i=1}^{N(t)} Y_i$. First note that $X(t)$ possesses stationary independent increments.

$$\begin{aligned}\mathbf{E}[X(s)X(t)] &= \mathbf{E}[X(s)(X(t) - X(s) + X(s))] \\ &= \mathbf{E}[X(s)(X(t) - X(s)) + \mathbf{E}[X(s)^2] \\ &= \mathbf{E}[X(s)]\mathbf{E}[X(t-s)] + \mathbf{E}[X(s)^2] \\ &= \lambda s\mathbf{E}[Y]\lambda(t-s)\mathbf{E}[Y] + \lambda s E[Y^2] + (\lambda s\mathbf{E}[Y])2 \\ &= \lambda^2 st\mathbf{E}[Y]^2 + \lambda s E[Y^2] .\end{aligned}$$

Hence

$$\begin{aligned}\mathbf{C}ov(X(s), X(t)) &= \mathbf{E}[X(s)X(t)] - \mathbf{E}[X(s)]\mathbf{E}[X(t)] \\ &= \lambda^2 st\mathbf{E}[Y]^2 + \lambda s E[Y^2] - \lambda s\mathbf{E}[Y]\lambda t\mathbf{E}[Y] \\ &= \lambda s E[Y^2] .\end{aligned}$$

IEOR 6711: Stochastic Models I, Professor Whitt

Solutions to Homework Assignment 6

Problem 3.1 **(a)** Yes since it is a contrapositive statement of (3.2.1).

(b) No. Note that we can have $X_{n+1} = 0$, so that $S_n = S_{n+k} = 0$ for some $k \geq 1$. Hence We can have $S_n = t$, but $N(t) > n$.

(c) No. Follows from (b) above.

Problem 3.2

$$\mathsf{P}(\text{exactly } n \text{ renewals}) = F(\infty)^n (1 - F(\infty)).$$

Problem 3.3 One might simply say that it is the length of the renewal interval that contains t, but that is not quite correct. Or, at least that answer is at best only partially correct. More precisely, it is the length of the first renewal interval that ends *after* time t. Suppose that $S_n = t < S_{n+1}$. Then the interval $(S_{n-1}, S_n]$ contains t. It is natural to define the renewal intervals as open on the left and closed on the right. Then $N(t) = n$. Then $X_{N(t)+1}$ is the length of the interval $(S_n, S_{n+1}]$, which does not contain t. The point $S_{N(t)+1}$ is the first renewal after ($>$) t. But we can only say that $S_{N(t)} \leq t$; we could have $S_{N(t)} = t$. When we write $X_{N(t)+1}$, we only discuss the length of that interval; the nature of the endpoints does not come up; so it is ambiguous.

$$\begin{aligned}
\mathsf{P}(A(t) \leq x) &= \begin{cases} 1 & \text{if } x \geq t \\ \mathsf{P}(\text{At least one event in } [t-x, t]) & \text{if } x < t \end{cases} \\
&= \begin{cases} 1 & \text{if } x \geq t \\ 1 - e^{-\lambda x} & \text{if } x < t \end{cases}
\end{aligned}$$

which is a *truncated* exponential distribution. The momoryless property implies that $Y(t)$ is an exponential distribution with rate λ and so the distribution of $X_{N(t)+1}$ is the convolution of $A(t)$ and an exponential with rate λ.

Problem 3.4

$$\mathsf{E}[N(t)] = \int_0^\infty \mathsf{E}[N(t) | X_1 = x] dF(x) \ .$$

As

$$\mathsf{E}[N(t) | X_1 = x] = \begin{cases} 1 + \mathsf{E}[N(t-x)] & \text{if } x \leq t \\ 0 & \text{if } x > t \end{cases}$$

the result follows.

Or following purely analytic way:

$$\begin{aligned} m(t) &= \sum_{n=1}^{\infty} F_n(t) \\ &= F_1(t) + \sum_{n=2}^{\infty} F_n(t) \\ &= F(t) + \sum_{n=1}^{\infty} \int_0^{\infty} F_n(t-x) dF(x) \\ &= F(t) + \int_0^{\infty} \left(\sum_{n=1}^{\infty} F_n(t-x) \right) dF(x) \\ &= F(t) + \int_0^{\infty} m(t-x) dF(x) \\ &= F(t) + \int_0^{t} m(t-x) dF(x) \end{aligned}$$

where the interchange of integral and sum is allowed by the non-negativity of the integrand and $m(t) = 0$ if $t \leq 0$.

Problem 3.5 There is a simple derivation here (to be presented below), but there is a source of confusion: We need to properly distinguish between the Laplace transform (LT) of a function and the Laplace-Stieltjes transform (LST) of an increasing function (or measure). The LT of the cdf F is

$$\hat{F}(s) = \int_0^{\infty} e^{-sx} F(x)\, dx \ ,$$

while the LST of F is

$$\hat{f}(s) = \int_0^{\infty} e^{-sx} dF(x) \ .$$

If the cdf F has a pdf (density) f, then the LST of F is the LT of f:

$$\hat{f}(s) = \int_0^{\infty} e^{-sx} dF(x) = \int_0^{\infty} e^{-sx} f(x)\, dx \ .$$

The same distinction applies to the renewal function $m(t)$, but the notation is confusing, because $m(t)$ plays the role of the cdf F, being increasing. We call $\hat{m}(s)$ the LT of m:

$$\hat{m}(s) = \int_0^{\infty} e^{-sx} m(x)\, dx \ ,$$

But we could consider the LST of m.

It is next important to be able to relate the LST and the LT. They are related by

$$\hat{F}(s) = \frac{\hat{f}(s)}{s} \ .$$

Note that this is a straightforward relation between LT's if the cdf F has a pdf f, but it applies more generally. Then, using properties of the Laplace transform of a convolution, we know that if $f_n(\cdot)$ represents the density of the sum of n IID random variables, then

$$\hat{f}_n(s) = (\hat{f}(s))^n ,$$

so that

$$\hat{F}_n(s) = \frac{\hat{f}_n(s)}{s} = \frac{(\hat{f}(s))^n}{s} .$$

Now using the identity

$$m(t) = \sum_{n=1}^{\infty} F_n(t) ,$$

take Laplace transform to obtain

$$\begin{aligned}
\hat{m}(s) &= \sum_{n=1}^{\infty} \hat{F}_n(s) \\
&= \sum_{n=1}^{\infty} \frac{\hat{f}(s)^n}{s} \\
&= \frac{\hat{f}(s)}{s(1-\hat{f}(s))}
\end{aligned}$$

or, equivalently,

$$\hat{F}(s) = \frac{\hat{f}(s)}{s} = \frac{\hat{m}(s)}{1 + s\ \hat{m}(s)} .$$

Since, a function is determined by its Laplace transform, this implies the result.

Or we may derive it by applying the Laplace transform to the result of Problem 3.3.4:

$$\begin{aligned}
\hat{m}(s) &= \hat{F}(s) + \hat{m}(s)\hat{f}(s) \\
&= \frac{\hat{f}(s)}{s} + \hat{m}(s)\hat{f}(s) .
\end{aligned}$$

Are you sure above derivations? For the first one, we need $|\hat{f}(s)| < 1$ and for the second one, $\hat{f}(s) \neq 1$.
$s = a + ib$ with $a > 0$ implies

$$\begin{aligned}
|\hat{f}(s)| &\leq \int_0^{\infty} |e^{-ax-ibx}| dF(x) \\
&= \int_0^{\infty} e^{-ax} dF(x) \\
&= \int_0^{x^*} e^{-ax} dF(x) + \int_{x^*}^{\infty} e^{-ax} dF(x) \\
&\leq \int_0^{x^*} dF(x) + e^{-ax^*} \int_{x^*}^{\infty} dF(x) \\
&= F(x^*) + e^{-ax^*}(1 - F(x^*)) \\
&< 1
\end{aligned}$$

for any $x^* > 0$ satisfying $F(x^*) < 1$, which supports our previous derivations.

Added Problem (a)

$$\hat{f}_{X+Y}(s) = \hat{f}_X(s)\hat{f}_Y(s) = \frac{1/2}{(s+1/2)}\frac{1/3}{(s+1/3)} = \frac{1}{(2s+1)(3s+1)} .$$

Hence

$$\hat{m}(s) = \frac{\hat{f}(s)}{s(1-\hat{f}(s))} = \frac{1}{s^2(6s+5)} .$$

(b) $m(10) \simeq 1.7601$, $m(20) \simeq 3.7600$.

(c) Adopting the assumption $m(t) \le c + dt$, let's start at the equation (11) of Whitt's 1995 paper:

$$\begin{aligned}
e_d &= \sum_{k=1}^{\infty} e^{-kA} m((2k+1)t) \\
&\le \sum_{k=1}^{\infty} e^{-kA}(c + d(2k+1)t) \\
&= (c+dt)\sum_{k=1}^{\infty} e^{-kA} + 2dt\sum_{k=1}^{\infty} ke^{-kA} \\
&= (c+dt)\frac{e^{-A}}{1-e^{-A}} + 2dt \lim_{n\to\infty} \frac{(n+1)e^{-(n+1)A} - ne^{-nA} + 1}{(1-e^{-A})^2} e^{-A} \\
&= (c+dt)\frac{e^{-A}}{1-e^{-A}} + 2dt\frac{e^{-A}}{(1-e^{-A})^2} \\
&\simeq (c+dt)e^{-A} + 2dte^{-A} \\
&= (c+3dt)e^{-A} .
\end{aligned}$$

Hence to achieve $\epsilon = 10^{-8}$ precision,

$$\begin{aligned}
A &\ge -\ln\left(\frac{\epsilon}{c+3dt}\right) \\
&= 8\ln 10 + \ln(c+3dt) .
\end{aligned}$$

$$d = \frac{1}{\mu} = \frac{1}{2+3} = \frac{1}{5} = 0.2$$

$$\begin{aligned}
c &\simeq \frac{\mathsf{E}[X^2]}{2\mu^2} - 1 = \frac{2^2+3^2+5^2 2\times 5^2}{2\times 2\times 5^2} \\
&= \frac{4+9-25}{50} = -\frac{6}{25} .
\end{aligned}$$

So

$$\begin{aligned}
A &\ge 8\ln 10 + \ln\left(-\frac{6}{25} + 0.6t\right) \\
&= 8\ln 10 + \ln(5.76 \text{ or } 11.76) \\
&= 20.1716 \text{ or } 20.8854 \\
&\simeq 21 .
\end{aligned}$$

With this new parameter, no change of the result up to four decimal digits.

Problem 3.6 For $s \le t$,

$$\mathsf{E}[N(s)|N(t)] = \frac{s}{t}N(t)$$

and so

$$\mathsf{E}\left[\frac{N(s)}{s}\right] = \mathsf{E}\left[\frac{N(t)}{t}\right]$$

implying $\mathsf{E}[N(t)/t]$ is a constant. That is $\mathsf{E}[N(t)] = ct$. But since this is also the renewal function for a Poisson process with rate c, and the renewal function uniquely determines the interarrival distribution (by Problem 3.3.5), we conclude that the renewal process is a Poisson process.

Problem 3.9 Set $\mu = \int_0^\infty \bar{G}(t)dt$. The underlying model is M/G/1/1. The first step is converting the model into an alternating renewal process. We may define the alternating cycles like (server is idle, server is busy), which is (exponential interarrival time, service time with $G(\cdot)$) in distributional sense. This is a renewal process since the arrival follows Poisson or the interarrival time doesn't have any memory. Then is is easy to get the followings:

(a)

$$\left(\frac{1}{\lambda} + \mu\right)^{-1}$$

(b)

$$\frac{\frac{1}{\lambda}}{\frac{1}{\lambda} + \mu} = \frac{1}{1 + \lambda\mu}$$

(c)

$$\frac{\mu}{\frac{1}{\lambda} + \mu}$$

Problem 3.10 First I will do the computations, without thinking carefully about justifying the steps:

(a)

$$\frac{S_1 + \cdots + S_m}{N_1 + \cdots + N_m} = \frac{\sum_{i=1}^{N_1 + \cdots + N_m}}{N_1 + \cdots + N_m} \underset{m\to\infty}{\longrightarrow} \mathsf{E}[X] .$$

(b)

$$\frac{S_1 + \cdots + S_m}{m} \underset{m\to\infty}{\longrightarrow} \mathsf{E}[S] = \mathsf{E}\left[\sum_1^{N_1} X_i\right] ,$$

$$\frac{N_1 + \cdots + N_m}{m} \underset{m\to\infty}{\longrightarrow} \mathsf{E}[N] = \mathsf{E}[N_1]$$

imply

$$\frac{S_1 + \cdots + S_m}{N_1 + \cdots + N_m} = \frac{S_1 + \cdots + S_m}{b} \frac{m}{N_1 + \cdots + N_m} \underset{m\to\infty}{\longrightarrow} \frac{\mathsf{E}\left[\sum_1^{N_1} X_i\right]}{\mathsf{E}[N_1]} .$$

(c)

$$\frac{\mathsf{E}\left[\sum_1^{N_1} X_i\right]}{\mathsf{E}[N_1]} = \mathsf{E}[X] .$$

Now where did we make bold computations? Are S_i IID? Why? If they are IID, that means

$$X_{N_1+i} \overset{d}{=} X_i$$

which looks similar to the famous *strong Markov* property. First noting that

$$\begin{aligned}
\mathsf{P}(X_{N_1+i} \le x | N_1 = n) &= \mathsf{P}(X_{n+i} \le x | N_1 = n) \\
&= \mathsf{P}(X_{n+i} \le x) \quad \text{since } N_1 \text{ is a stopping time.} \\
&= \mathsf{P}(X_i \le x) ,
\end{aligned}$$

$$\begin{aligned}
\mathsf{P}(X_{N_1+i} \le x) &= \mathsf{E}[\mathsf{P}(X_{N_1+i} \le x | N_1)] \\
&= \mathsf{E}[\mathsf{P}(X_i \le x)] \\
&= \mathsf{P}(X_i \le x) .
\end{aligned}$$

We can also conclude that S_i are IID using similar arguments.

Problem 3.11 **(a)** X_i is the travel time on the ith choice; N is the number of choices until freedom is reached.

(b) $\mathsf{E}[T] = \mathsf{E}[N]\mathsf{E}[X]$, $\mathsf{E}[X] = \frac{1}{3}(2+4+8) = \frac{14}{3}$, $\mathsf{E}[N] = 3$ since N is geometric with $p = 1/3$. Hence $\mathsf{E}[T] = 14$.

(c)

$$\begin{aligned}
\mathsf{E}\left[\sum_1^N X_i \middle| N = n\right] &= \sum_1^n \mathsf{E}\left[X_i | N = n\right] \\
&= \sum_1^{n-1} \mathsf{E}[X_i|\text{Incorrect door}] + \mathsf{E}[X_n|\text{Correct door}] \\
&= \sum_1^{n-1} \frac{1}{2}(4+8) + 2 \\
&= 6n - 4 \\
&\ne \frac{14}{3} n \\
&= \mathsf{E}\left[\sum_1^n X_i\right] .
\end{aligned}$$

(c)

$$
\begin{aligned}
\mathsf{E}[T] &= \mathsf{E}\left[\sum_1^N X_i\right] \\
&= \mathsf{E}\left[\left[\sum_1^N X_i \middle| N\right]\right] \\
&= \mathsf{E}[6N-4] \\
&= 6\mathsf{E}[N] - 4 = 18 - 4 = 14 \ .
\end{aligned}
$$

IEOR 6711: Stochastic Models I, Professor Whitt

SOLUTIONS to Homework Assignment 7

Problem 3.12 Take $h(t) = \mathbf{1}_{(0,a]}(t)$.

Problem 3.13 Since the state circulates the state space $\{1, 2, \cdots, n\}$ in the same order. Hence we can define the alternating renewal process by *on* when it is in the state i and *off* when it is among $\{1, \cdots, i-1, i+1, \cdots, n\}$. Define $\mu_i = \int \bar{F}_i(t)dt$. Then

$$\mathsf{P}(\text{process is in } i) \to \frac{\mathsf{E}[on]}{\mathsf{E}[on] + \mathsf{E}[off]} = \frac{\mu_i}{\sum_{j=1}^{n} \mu_j} \ .$$

Problem 3.14 **(a)** $[t-x, t]$

(b) $[t, t+x]$

(c) $\mathsf{P}(Y(t) > x) = \mathsf{P}(A(t+x) > x)$

(d) See Problem 3.3.

Problem 3.15 **(a)**

$$\begin{aligned} \mathsf{P}(Y(t) > x | A(t) = s) &= \mathsf{P}(X_{N(t)+1} > x + s | \text{time at } t \text{ since the last renewal} = s) \\ &= \frac{\bar{F}(x+s)}{\bar{F}(s)} \ . \end{aligned}$$

(b) Using (a),

$$\mathsf{P}(Y(t) > x | A(t + x/2) = s) = \begin{cases} 0 & \text{if } s < \frac{x}{2} \\ \frac{\bar{F}(s+x/2)}{\bar{F}(s)} & \text{if } s \geq \frac{x}{2} \end{cases} \ .$$

(c)

$$\mathsf{P}(Y(t) > x | A(t+x) > s) = \begin{cases} 1 & \text{if } s \geq x \\ \mathsf{P}(\text{no events in } [t, t+x-s]) = e^{-\lambda(x-s)} & \text{if } s < x \end{cases} \ .$$

(d)

$$\mathsf{P}(Y(t) > x,\ A(t) > y) = \mathsf{P}(Y(t-y) > x + y) = \mathsf{P}(A(t+x) > x + y) \ .$$

(e)

$$\frac{A(t)}{t} = \frac{t - S_{N(t)}}{t} = 1 - \frac{S_{N(t)}}{N(t)} \frac{N(t)}{t} \to 1 - \mu \frac{1}{\mu} = 0 \ .$$

Problem 3.16

$$\begin{aligned}\lim_{t\to\infty} \mathsf{E}[Y(t)] &= \frac{\mathsf{E}[X^2]}{2\mu} \\ &= \frac{n\left(\frac{1}{\lambda}\right)^2 + \left(\frac{n}{\lambda}\right)^2}{2\frac{n}{\lambda}} \\ &= \frac{1+n}{2\lambda}\ .\end{aligned}$$

To get it without any computations, consider a Poisson process with rate λ and say the a *renewal* occurs at the Poisson events numbered $n, 2n, \cdots$. Now at time t, t large, it is equally likely that the most recent event was an event of the form $i+kn$, $i = 0, 1, 2, \cdots, n-1$. That is, modulo n, the number of the most recent Poisson event is equally likely to be $n, 1, \cdots, n-1$. Conditioning on the value of this quantity gives that for the renewal process

$$\lim_{t\to\infty} \mathsf{E}[Y(t)] = \frac{1}{n}\left(\frac{1}{\lambda} + \cdots + \frac{n}{\lambda}\right) = \frac{n+1}{2\lambda}\ .$$

Problem 3.18 **(a)** Delayed renewal process.

(b) Neither.

If F is exponential,

(a) Delayed renewal process.

(b) Renewal process.

Problem 3.21 Let X_i equal 1 if the gambler wins bet i, and let it be 0 otherwise. Also, let N denote the first time the gambler has won k consecutive bets. Then $X = \sum_{i=1}^{N} X_i$ is equal to the number of bets that he wins, and $X - (N - X) = 2X - N$ is his winnings. By Wald's equation

$$\mathsf{E}[X] = p\mathsf{E}[N] = p\sum_{i=1}^{k} p^{-i}\ .$$

Thus

(a) $\mathsf{E}[2X - N] = 2\mathsf{E}[X] - \mathsf{E}[N] = (2p-1)\mathsf{E}[N] = (2p-1)\sum_{i=1}^{k} p^{-i}$

(b) $\mathsf{E}[X] = p\sum_{i=1}^{k} p^{-i}$

Problem 3.22 **(a)**

$$\begin{aligned}\mathsf{E}[T_{HHTTHH}] &= \mathsf{E}[T_{HH}] + p^{-4}(1-p)^{-2} \\ &= \mathsf{E}[T_H] + p^{-2} + p^{-4}(1-p)^{-2} \\ &= p^{-1} + p^{-2} + p^{-4}(1-p)^{-2}\end{aligned}$$

(b) $\mathsf{E}[T_{HTHTT}] = p^{-2}(1-p)^{-3}$

$\mathsf{E}[N_{B|A}] = \mathsf{E}[N_{HTHTT|H}] = \mathsf{E}[N_{HTHTT}] - \mathsf{E}[N_H] = 32 - 2 = 30$, $\mathsf{E}[N_{A|B}] = \mathsf{E}[N_A] = 64 + 4 + 2 = 70$ and $\mathsf{E}[N_B] = 32$.

(c) $P_A = (32 + 70 - 70)/(30 + 70) = 0.32$

(d) $\mathsf{E}[M] = 32 - 30(0.32) = 22.4$

Problem 3.23 Let H denote the first k flips and Ω is the set of all possible H. Conditioning on H gives:

$$\begin{aligned}\mathsf{E}[\text{number until repeat}] &= \sum_{H\in\Omega} \mathsf{E}[\text{number until repeat}|H]\mathsf{P}(H) \\ &= \sum_{H\in\Omega} \frac{1}{\mathsf{P}(H)}\mathsf{P}(H) = |\Omega| = 2^k\end{aligned}$$

Problem 3.25 **(a)** First note that

$$\mathsf{E}[N_D(t)|X_1 = x] = \begin{cases} 1 + \mathsf{E}[N(t-x)] & \text{if } x \le t \\ 0 & \text{if } x > t \end{cases}.$$

$$\begin{aligned}m_D(t) = \mathsf{E}[N_D(t)] &= \int_0^\infty \mathsf{E}[N_D(t)|X_1 = x]dG(x) \\ &= \int_0^t (1 + \mathsf{E}[N(t-x)])dG(x) \\ &= G(t) + \int_0^t m(t-x)dG(x)\end{aligned}$$

(b)

$$\begin{aligned}\mathsf{E}[A_D(t)] &= \mathsf{E}[A_D(t)|S_{N_D(t)} = 0]\bar{G}(t) + \int_0^t \mathsf{E}[A_D(t)|S_{N_D(t)} = s]\bar{F}(t-s)dm_D(s) \\ &= t\bar{G}(t) + \int_0^t (t-s)\bar{F}(t-s)dm_D(s) \\ &\overset{t\to\infty}{\longrightarrow} \frac{1}{\mu}\int_0^\infty t\bar{F}(t)dt \quad \text{By key renewal theorem (Proposition 3.5.1(v)} \\ &= \frac{1}{\mu}\int_0^\infty t\int_t^\infty dF(s)dt \\ &= \frac{\int_0^\infty s^2 dF(s)}{2\int_0^\infty s dF(s)}\end{aligned}$$

(c) $t\bar{G}(t) = t\int_t^\infty dG(x) \le \int_t^\infty s\ dG(s) \overset{t\to\infty}{\longrightarrow} 0$ since $\int_0^\infty s\ dG(s) < \infty$.
(Here we used the so-called *dominated convergence theorem.*

$$\begin{aligned}\int_n^\infty s\ dG(s) &= \int_0^\infty s\mathbf{1}_{[n,\infty)}(s)\ dG(s) \\ &\underset{\text{dominated convergence thorem}}{\overset{n\to\infty}{\longrightarrow}} \int_0^\infty \lim_{n\to\infty} s\mathbf{1}_{[n,\infty)}(s)\ dG(s) = \int_0^\infty 0\ dG(s)\end{aligned}$$

since $s\mathbf{1}_{[n,\infty)}(s) \le s$ and s is integrable with respect to $G(\cdot)$ from $\int_0^\infty s\, dG(s) < \infty$ and $s\mathbf{1}_{[n,\infty)}(s) \to 0$ for each s in *pointwise* sense. (Check the conditions for the dominated convergence theorem.) Now we extend n to t using monotonicity of the integral. Wow! This is a good example showing that if you are familiar with a little rigorous *analysis*, then it's O.K. with only one line. But if not, you should practice the underlying logic whenever you encounter them.)

Problem 3.28 Using the *uniformity* of each Poisson arrival under given $N(t)$,

$$\mathsf{E}[\text{Cost of a cycle}|N(T)] = K + N(T) \times c \times \frac{T}{2}$$

and so

$$\frac{\mathsf{E}[\text{Cost}]}{\mathsf{E}[\text{Time}]} = \frac{K + \lambda c T^2/2}{T} = \frac{K}{T} + \frac{\lambda c T}{2}$$

which is minimized at $T^* = \sqrt{2K/\lambda c}$ and minimal average cost is thus $\sqrt{2\lambda K c}$. On the other hand the optimal value of N is (using calculus) $N^* = \sqrt{2\lambda K/c}$ and the minimal average cost is $\sqrt{2\lambda c K} - \frac{c}{2}$.

IEOR 6711: Stochastic Models I, Professor Whitt

SOLUTIONS to Homework Assignment 8.

Problem 3.19

$$\begin{aligned}
P\{S_{N(t)} \le s\} &= \sum_{n=0}^{\infty} P\{S_n \le s, S_{n+1} > t\} \\
&= \bar{G}(t) + \sum_{n=1}^{\infty} P\{S_n \le s, S_{n+1} > t\} \\
&= \bar{G}(t) + \sum_{n=1}^{\infty} \int_0^{\infty} P\{S_n \le s, S_{n+1} > t | S_n = y\} d(G * F_{n-1})(s) \\
&= \bar{G}(t) + \int_0^s \bar{F}(t-y) dm_D(y)
\end{aligned}$$

Problem 3.20

(a) Say that a renewal occurs when that pattern appears. By Blackwell's theorem for renewal processes we obtain

$$E[time] = \frac{1}{(1/2)^7} = 2^7$$

(b) By Blackwell's theorem

E[time between HHTT renewals] $= E$[time between HTHT renewals] $= 16$

But the HHTT renewal process is an ordinary one and so the mean time until HHTT occurs is 16 whereas the HTHT process is a delayed renewal process and so the mean time until HTHT occurs is greater than 16.

Problem 3.29

Let L denote the lifetime of a car with distribution function $F(\cdot)$.

(a) Under the policy of replacements at A,

$$\text{Cost of cycle} = \begin{cases} C_1 + C_2 & \text{if } L \le A \\ C_1 - R(A) & \text{if } L > A \end{cases}$$

and

$$\text{Length of cycle} = \begin{cases} L & \text{if } L \le A \\ A & \text{if } L > A \end{cases} .$$

Then

$$\frac{\mathsf{E}[\text{Cost}]}{\mathsf{E}[\text{Time}]} = \frac{C_1 + C_2 F(A) - R(A)\bar{F}(A)}{\int_0^A x dF(x) + A\bar{F}(A)}$$

(Validate the final formula by yourself. If you are confusing, utilize the *indicator* to combine the *if*-clauses into one function as I said in the first recitation.)

(b) Condition on the life of the initial car.

$$\begin{aligned}
\mathsf{E}[\text{Length of cycle}] &= \int_0^\infty \mathsf{E}[\text{Length}|L=x]dF(x) \\
&= \int_0^A x dF(x) + \int_A^\infty (A + \mathsf{E}[\text{Length}])dF(x) \\
&= \int_0^A x dF(x) + (A + \mathsf{E}[\text{Length}])\bar{F}(A) \\
&= \frac{\int_0^A x dF(x) + A\bar{F}(A)}{F(A)}
\end{aligned}$$

and similarly

$$\begin{aligned}
\mathsf{E}[\text{Cost of cycle}] &= \int_0^\infty \mathsf{E}[\text{Cost}|L=x]dF(x) \\
&= \int_0^A (C_1 + C_2)dF(x) + (C_1 - R(A) + \mathsf{E}[Cost])\bar{F}(A) \\
&= \frac{C_1 + C_2 F(A) - R(A)\bar{F}(A)}{F(A)}\ .
\end{aligned}$$

Then

$$\frac{\mathsf{E}[\text{Cost}]}{\mathsf{E}[\text{Time}]} = \text{same as in (a).}$$

Problem 3.31

Let μ_i and ν_i denote the means of F_i and G_i, respectively for $i = 1, 2, 3, 4$. Then,

$$\lim P\{\text{i is working at t}\} = \mu_i/(\mu_i + \nu_i), i = 1, 2, 3, 4$$

Now, if p_i is the probability that component i is working, then

$$P\{\text{system works}\} = (p_1 + p_2 - p_1 p_2)(p_3 + p_4 - p_3 p_4)$$

Hence $\lim P\{\text{system works at t}\}$ is equal to the preceding expression with $p_i = \mu_i/(\mu_i + \nu_i), i = 1, 2, 3, 4$

Problem 3.32

(a) $1 - P_0 =$ average number in service $= \lambda\mu$

(b) By alternating renewal processes

$$P_0 = \text{proportion of time empty} = \frac{E[I]}{E[I] + E[B]}$$

where I is an idle period and B a busy period. But clearly I is exponential with rate λ and so

$$1 - \lambda\mu = \frac{1/\lambda}{1/\lambda + E[B]} \text{ or } E[B] = \frac{\mu}{1 - \lambda\mu}$$

(c) Let C denote the number of customers served in a busy period B and let S_i denote the service time of the i-th customer, $i \geq 1$. Then

$$B = \sum_{i=1}^{C} S_i$$

and by Wald's equation

$$E[C] = \frac{E[B]}{E[S]} = \frac{1}{1 - \lambda\mu}$$

IEOR 6711: Stochastic Models I

Solutions to Homework Assignment 9

Problem 4.1 Let D_n be the random demand of time period n. Clearly D_n is i.i.d. and independent of all X_k for $k < n$. Then we can represent $X_n + 1$ by

$$X_{n+1} = \max\{0, X_n \cdot \mathbf{1}_{[s,\infty)}(X_n) + S \cdot \mathbf{1}_{[0,s)}(X_n) - D_{n+1}\}$$

which depends only on X_n since D_{n+1} is independent of all history. Hence $\{X_n, n \geq 1\}$ is a Markov chain. It is easy to see assuming $\alpha_k = 0$ for $k < 0$,

$$P_{ij} = \begin{cases} \alpha_{S-j} & \text{if } i < s, j > 0 \\ \sum_{k=S}^{\infty} \alpha_k & \text{if } i < s, j = 0 \\ \alpha_{i-j} & \text{if } i \geq s, j > 0 \\ \sum_{k=i}^{\infty} \alpha_k & \text{if } i \geq s, j = 0 \end{cases}$$

The following three problems (4.2, 4.4, 4.5) needs a fact:

$$\mathsf{P}(A \cap B|C) = \mathsf{P}(A|B \cap C)\mathsf{P}(B|C)$$

which requires a proof to use. Try to prove it by yourself.

Problem 4.2 Let $\mathcal{S}$ be the state space. First we show that

$$\mathsf{P}(X_{n_k+1} = j|X_{n_1} = i_1, \cdots, X_{n_k} = i_k) \quad = \quad \mathsf{P}(X_{n_k+1} = j|X_{n_k} = i_k)$$

by the following : Let $A = \{X_{n_k+1} = j\}$, $B = \{X_{n_1} = i_1, \cdots, X_{n_k} = i_k\}$ and $B_b, b \in \mathcal{I}$ are elements of $\{(X_l, l \leq n_k, l \neq n_1, \cdots, l \neq n_k) : X_l \in \mathcal{S}\}$.

$$\begin{aligned} \mathsf{P}(A|B) &= \sum_{b\in\mathcal{I}} \mathsf{P}(A \cap B_b|B) \\ &= \sum_{b\in\mathcal{I}} \mathsf{P}(A|B_b \cap B)\mathsf{P}(B_b|B) \\ &= \sum_{b\in\mathcal{I}} \mathsf{P}(A|X_{n_k} = i_k)\mathsf{P}(B_b|B) \\ &= \mathsf{P}(A|X_{n_k} = i_k) \sum_{b\in\mathcal{I}} \mathsf{P}(B_b|B) \\ &= \mathsf{P}(A|X_{n_k} = i_k)\mathsf{P}(\Omega|B) \\ &= \mathsf{P}(X_{n_k+1} = j|X_{n_k} = i_k) \ . \end{aligned}$$

We consider the mathematical induction on $l \equiv n-m$. For $l=1$, we just showed. Now assume that the statement is true for all $l \le l^*$ and consider $l = l^*+1$:

$$\begin{aligned}
& \mathsf{P}(X_n = j | X_{n_1} = i_1, \cdots, X_{n_k} = i_k) \\
= & \sum_{i \in \mathcal{S}} \mathsf{P}(X_n = j, X_{n-1} = i | X_{n_1} = i_1, \cdots, X_{n_k} = i_k) \\
= & \sum_{i \in \mathcal{S}} \mathsf{P}(X_n = j | X_{n-1} = i, X_{n_1} = i_1, \cdots, X_{n_k} = i_k)\mathsf{P}(X_{n-1} = i | X_{n_1} = i_1, \cdots, X_{n_k} = i_k) \\
= & \sum_{i \in \mathcal{S}} \mathsf{P}(X_n = j | X_{n-1} = i)\mathsf{P}(X_{n-1} = i | X_{n_k} = i_k) \quad \text{By } l \le l^* \text{ cases} \\
= & \sum_{i \in \mathcal{S}} \mathsf{P}(X_n = j | X_{n-1} = i, X_{n_k} = i_k)\mathsf{P}(X_{n-1} = i | X_{n_k} = i_k) \\
= & \sum_{i \in \mathcal{S}} \mathsf{P}(X_n = j, X_{n-1} = i | X_{n_k} = i_k) \\
= & \mathsf{P}(X_n = j | X_{n_k} = i_k)
\end{aligned}$$

which completes the proof for $l = l^*+1$ case.

Problem 4.3 Simply by *Pigeon hole principle* which saying that if n pigeons return to their $m(< n)$ home (through hole), then at least one home contains more than one pigeon. Consider any path of states $i_0 = i, i_1, \cdots, i_n = j$ such that $P_{i_k, i_{k+1}} > 0$. Call this a path from i to j. If j can be reached from i, then there must be a path from i to j. Let $i_0. \cdots, i_n$ be such a path. If all of values $i_0, \cdots, i_n$ are not distinct, then there must be a subpath from i to j having fewer elements (for instance, if $i, 1, 2, 4, 1, 3, j$ is a path, then so is $i, 1, 3, j$). Hence, if a path exists, there must be one with all distinct states.

Problem 4.4 Let Y be the first passage time to the state j starting the state i at time 0.

$$\begin{aligned}
P_{ij}^n &= \mathsf{P}(X_n = j | X_0 = i) \\
&= \sum_{k=0}^{n} \mathsf{P}(X_n = j, Y = k | X_0 = i) \\
&= \sum_{k=0}^{n} \mathsf{P}(X_n = j | Y = k, X_0 = i)\mathsf{P}(Y = k | X_0 = i) \\
&= \sum_{k=0}^{n} \mathsf{P}(X_n = j | X_k = j)\mathsf{P}(Y = k | X_0 = i) \\
&= \sum_{k=0}^{n} P_{jj}^{n-k} f_{ij}^k
\end{aligned}$$

Problem 4.5 **(a)** The probability that the chain, starting in state i, will be in state j at time n without ever having made a transition into state k.

(b) Let Y be the last time leaving the state i before first reaching to the state j starting the state i at time 0.

$$\begin{aligned} P_{ij}^n &= \mathsf{P}(X_n = j | X_0 = i) \\ &= \sum_{k=0}^{n} \mathsf{P}(X_n = j, Y = k | X_0 = i) \\ &= \sum_{k=0}^{n} \mathsf{P}(X_n = j, Y = k, X_k = i | X_0 = i) \\ &= \sum_{k=0}^{n} \mathsf{P}(X_n = j, Y = k | X_k = i, X_0 = i) \mathsf{P}(X_k = i | X_0 = i) \\ &= \sum_{k=0}^{n} \mathsf{P}(X_n = j, Y = k | X_k = i) P_{ii}^k \\ &= \sum_{k=0}^{n} \mathsf{P}(X_n = j, X_l \neq i, l = k+1, \cdots, n-1 | X_k = i) P_{ii}^k \\ &= \sum_{k=0}^{n} P_{ij/i}^{n-k} P_{ii}^k \end{aligned}$$

Problem 4.7

(a) ∞

Here is an argument: Let x be the expected number of steps required to return to the initial state (the origin). Let y be the expected number of steps to move to the left 2 steps, which is the same as the expected number of steps required to move to the right 2 steps. Note that the expected number of steps required to go to the left 4 steps is clearly $2y$, because you first need to go to the left 2 steps, and from there you need to go to the left 2 steps again. Then, consider what happens in successive pairs of steps: Using symmetry, we get

$$x = 2 + (0 \times (1/2) + y \times (1/2) = 2 + y/2$$

and

$$y = 2 + (0 \times (1/4) + y \times (1/2) + (2 * y) \times (1/4)$$

If we subtract y from both sides, this last equation yields

$$2 = 0 \ .$$

Hence there is no finite solution. The quantity y must be infinite; a finite value cannot solve the equation.

(b) Note that the expected number of returns in $2n$ steps is the sum of the probabilities of returning in $2k$ steps for k from 1 to n, each term of which is binomial. Thus, we have

$$E[N_{2n}] = \sum_{k=1}^{n} \frac{(2k)!}{k!k!}(1/2)^{2k} \ ,$$

which can be shown to be equal to the given expression by mathematical induction.

(c) We say that $f(n) \sim g(n)$ as $n \to \infty$ if

$$f(n)/g(n) \to 1 \quad \text{as} \quad n \to \infty \ .$$

By Stirling's approximation,

$$(2n+1)\frac{(2n)!}{n!n!}(1/2)^2 n \sim 2\sqrt{n/\pi} \ ,$$

so that

$$E[N_n] \sim \sqrt{2n/\pi} \quad \text{as} \quad n \to \infty \ .$$

Problem 4.8 (a)

$$P_{ij} = \frac{\alpha_j}{\sum_{k=i+1}^{\infty} \alpha_k} \ , \quad j > i$$

(b) $\{T_i, i \geq 1\}$ is not a Markov chain - the distribution of T_i does depend on R_i. $\{(R_{i+1}, T_i), i \geq 1\}$ is a Markov chain.

$$\begin{aligned} \mathsf{P}(R_{i+1} = j, T_i = n | R_i = l, T_{i-1} = m) &= \frac{\alpha_j}{\sum_{k=l+1}^{\infty} \alpha_k} \left(\sum_{k=0}^{l} \alpha_k \right)^{n-1} \sum_{k=l+1}^{\infty} \alpha_k \\ &= \alpha_j \left(\sum_{k=0}^{l} \alpha_k \right)^{n-1} \ , \quad j > l \end{aligned}$$

(c) If $S_n = j$ then the $(n+1)^{st}$ record occurred at time j. However, knowledge of when these $n+1$ records occurred does not yield any information about the set of values $\{X_1, \cdots, X_j\}$. Hence, the probability that the next record occurs at time k, $k > j$, is the probability that both $\max\{X_1, \cdots, X_j\} = \max\{X_1, \cdots, X_{k-1}\}$ and that $X_k = \max\{X_1, \cdots, X_k\}$. Therefore, we see that $\{S_n\}$ is a Markov chain with

$$P_{jk} = \frac{j}{k-1}\frac{1}{k} \ , \quad k > j \ .$$

Problem 4.11 (a)

$$\begin{aligned} \sum_{n=1}^{\infty} P_{ij}^n &= \mathsf{E}[\text{number of visits to } j | X_0 = i] \\ &= \mathsf{E}[\text{number of visits to } j | \text{ever visit } j, X_0 = i] f_{ij} \\ &= (1 + \mathsf{E}[\text{number of visits to } j | X_0 = j]) f_{ij} \\ &= \frac{f_{ij}}{1 - f_{jj}} < \infty \ . \end{aligned}$$

since 1 + number of visits to $j | X_0 = j$ is geometric with mean $\frac{1}{1-f_{jj}}$.

(b) Follows from above since

$$\begin{aligned}\frac{1}{1-f_{jj}} &= 1+\mathsf{E}[\text{number of visits to } j|X_0=j] \\ &= 1+\sum_{n=1}^{\infty} P_{jj}^{n}\ .\end{aligned}$$

Problem 4.12 If we add the irreducibility of $\mathbf{P}$, it is easy to see that $\boldsymbol{\pi} = \frac{1}{n}\mathbf{1}$ is a (and the unique) limiting probability.

IEOR 6711: Stochastic Models I, Professor Whitt

Solutions to Homework Assignment 10

Numerical Problems 1.(a) $\pi_5 = 0.1667$

1.(b) Yes, because the Markov chain is irreducicle and has a finite state space. The stationary probability of being in state 5 is $\pi_5 = 0.1667$. The stationary probability vector is π such that $\pi = \pi P$. However, there is no limiting probability (i.e., we do not have a limit for P^n as $n \to \infty$), because the chain is periodic, with period 2.

1.(c) For large n, $P^{2n+1}_{1,5} = 0$ and $P^{2n}_{1,5} \simeq 2\pi_5 = 0.3334$

1.(d) $1/\pi_5 = 6$

2.(a) $M_1 = 14.26303$

2.(b) $N_{1,5} = 2.21054$

2.(c) $B_{1,10} = 0.3684$

Problem 4.18 Let $a_j = e^{-\lambda}\lambda^j/j!\ , \quad j \geq 0$.

(a)

$$P_{0,j} = a_j\ ,\ j < N, \quad P_{0,N} = 1 - \sum_{j=0}^{N-1} a_j$$

$$\text{For } i > 0,\ P_{i,j} = a_{j-i+1}\ ,\ j = i-1, \cdots, N-1,\ P_{i,N} = 1 - \sum_{j=0}^{N-i} a_j\ .$$

(b) Yes, because it is a finite, irreducible Markov chain.

(c) As one of the equations is redundant, we can write them as follows :

$$\begin{aligned} \pi_j &= \pi_0 a_j + \sum_{i=1}^{j+1} \pi_i a_{j-i+1}\ , \quad j = 0, \cdots, N-1 \\ \sum_{j=0}^{N} \pi_j &= 1\ . \end{aligned}$$

Problem 4.19 (a) are from state i to state j.

(b) go from a state in A to one in A^c.

(c) This follows because between any two transitions that go from a state in A to one in A^c there must be a transition from a state in A^c to one in A, and vice-versa.

(d) It follows from (c) that the long-run proportion of transitions that are from a state in A to one in A^c must equal the long-run proportion of transitions that go from a state in A^c to one in A; and that is what (d) asserts.

Problem 4.31 Let the states be

0 : spider and fly at same location

1 : spider at location 1 and fly at 2

2 : spider at 2 and fly at 1

$$P = \begin{bmatrix} 1 & 0 & 0 \\ .54 & .28 & .18 \\ .54 & .18 & .28 \end{bmatrix}$$

(a)

$$P_{11}^n = (0.46)^n \left[\frac{1}{2} + \frac{1}{2}\left(\frac{28}{23} - 1\right)^n\right]$$

which is obtained by first conditioning on the event that 0 is not entered and then using the fact that for the

$$\begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$$

chain $P_{00}^n = \frac{1}{2} + \frac{1}{2}(2p-1)^n$.

More generally, we can find explicit analytical expressions for n-step transition probabilities by applying the spectral representation of the sub-probability transition matrix

$$Q = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$$

(The same argument applies without that special structure. See the Appendix of Karlin and Taylor for a textbook review of this part of basic linear algebra.) We want to find constants λ such that

$$xQ = \lambda x \ . \tag{1}$$

Those are the eigenvalues of Q. To find the eigenvalues, we solve the equation

$$det(Q - \lambda I) = 0 \ ,$$

where *det* is the determinant. Here the equation is

$$(a - \lambda)^2 - b^2 = 0 \ ,$$

which yields two solutions: $a + b$ and $a - b$. We then find the left eigenvectors of Q. A row vector x is a left eigenvector of Q associated with the eigenvalue

λ if equation (1) hold. Similarly, the transpose of x, denoted by x^T, is a right eigenvector of Q associated with eigenvalue λ if

$$Qx^T = \lambda x^T \ . \tag{2}$$

We then can find a *spectral representation* for Q:

$$Q = R\Lambda L \ , \tag{3}$$

with the following properties: (i) R and L are square matrices with the same dimension as Q, (ii) the columns of R are right eigenvectors of Q; (iii) the rows of L are left eigenvectors of Q, (iv) $RL = LR = I$, and (v) Λ is a square diagonal matrix with the eigenvalues for its diagonal elements. As a consequence, we have

$$Q^n = R\Lambda^n L \quad \text{for all} \quad n \geq 1 \ , \tag{4}$$

enabling us to compute Q^n, easily because Λ^n is a diagonal matrix with diagonal elements λ^n, where λ is an eigenvector.

Here we get eigenvalues of Q equal to $a+b$ and $a-b$. Here we get eigenvector matrices

$$L = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

and

$$R = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

We obtain one of these by directly solving for the eigenvectors (which are not unique). Given L or R, we can obtain the other by inverting the matrix, i.e., $L = R^{-1}$.
Hence, equation (4) holds

$$Q^n = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \times \begin{bmatrix} (a+b)^n & 0 \\ 0 & (a-b)^n \end{bmatrix} \times \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

Thus, in general,

$$Q^n_{1,1} = \frac{(a+b)^n}{2} + \frac{(a-b)^n}{2}$$

and, in particular,

$$Q^n_{1,1} = \frac{(0.46)^n}{2} + \frac{(0.10)^n}{2}$$

(b) $\mathsf{E}[N] = \frac{1}{.54}$ since N is geometric (on the positive integers, not including 0) with $p = 0.54$.

IEOR 6711: Stochastic Models I, Professor Whitt

Solutions to Homework Assignment 11 on DTMC's

Problem 4.40 Consider a segment of a sample path beginning and ending in state i, with no visit to i in between, i.e, the vector $(i, j_1, j_2, j_3, \ldots, j_{n-1}, j_n = i)$, where $j_k \neq i$ for the non-end states j_k. Going forward in time, the probability of this segment is

$$\pi_i P_{i,j_1} P_{j_1,j_2} P_{j_2,j_3} \cdots P_{j_{n-1},i}.$$

The probability, say p, of the reversed sequence $(i, j_{n-1}, j_{n-2}, j_{n-3}, \ldots, j_1, j_0 = i)$ under the reverse DTMC with transition matrix

$$\overleftarrow{P}_{i,j} \equiv \frac{\pi_j P_{j,i}}{\pi_i}$$

is

$$p = \pi_i \overleftarrow{P}_{i,j_{n-1}} \overleftarrow{P}_{j_{n-1},j_{n-2}} \overleftarrow{P}_{j_{n-2},j_{n-3}} \cdots \overleftarrow{P}_{j_1,i}.$$

However, successively substituting in the reverse-chain transition probabilities, we get

$$\begin{aligned}
p &= \pi_i \frac{\pi_{j_{n-1}} P_{j_{n-1},i}}{\pi_i} \overleftarrow{P}_{j_{n-1},j_{n-2}} \overleftarrow{P}_{j_{n-2},j_{n-3}} \cdots \overleftarrow{P}_{j_1,i} \\
&= P_{j_{n-1},i} \pi_{j_{n-1}} \overleftarrow{P}_{j_{n-1},j_{n-2}} \overleftarrow{P}_{j_{n-2},j_{n-3}} \cdots \overleftarrow{P}_{j_1,i} \\
&= P_{j_{n-1},j_{n-2}} \pi_{j_{n-1}} \frac{\pi_{j_{n-2}} P_{j_{n-2},j_{n-1}}}{\pi_{j_{n-1}}} \overleftarrow{P}_{j_{n-1},j_{n-2}} \overleftarrow{P}_{j_{n-2},j_{n-3}} \cdots \overleftarrow{P}_{j_1,i} \\
&= P_{j_{n-1},i} P_{j_{n-2},j_{n-1}} P_{j_{n-3},j_{n-2}} \ldots P_{j_1,j_2} \pi_{j_1} \overleftarrow{P}_{j_1,i} \\
&= P_{j_{n-1},i} P_{j_{n-2},j_{n-1}} P_{j_{n-3},j_{n-2}} \ldots P_{j_1,j_2} P_{i,j_1} \pi_i \\
&= \pi_i P_{i,j_1} P_{j_1,j_2} P_{j_2,j_3} \cdots P_{j_{n-1},i}.
\end{aligned}$$

Problem 4.41 (a) The reverse time chain has transition matrix

$$\overleftarrow{P}_{i,j} \equiv \frac{\pi_j P_{j,i}}{\pi_i}$$

To find it, we need to first find the stationary vector π. By symmetry (or by noting that the chain is doubly stochastic), $\pi_j = 1/n,\ j = 1, \cdots, n$. Hence,

$$P^*_{ij} = \pi_j P_{ji}/\pi_i = P_{ji} = \begin{cases} p & \text{if } j = i-1 \\ 1-p & \text{if } j = i+1 \end{cases}$$

(b) In general, the DTMC is not time reversible. It is in the special case $p = 1/2$. Otherwise, the probabilities of clockwise and counterclockwise motion are reversed.

Problem 4.42 Imagine that there are edges between each of the pair of nodes i and $i+1$, $i = 0, \cdots, n-1$, and let the weight on edge $(i, i+1)$ be w_i, where

$$\begin{aligned} w_0 &= 1 \\ w_i &= \prod_{j=1}^{i} \frac{p_j}{q_j} \ , \quad i \geq 1 \end{aligned}$$

where $q_j = 1 - p_j$. As a check, note that with these weights

$$P_{i,i+1} = \frac{w_i}{w_{i-1} + w_i} = \frac{p_i/q_i}{1 + p_i/q_i} = p_i \ , \quad 0 < i < n \ .$$

Since the sum of the weights on edges out of node i is $w_{i-1} + w_i$, $i = 1, \cdots, n-1$, it follows that

$$\begin{aligned} \pi_0 &= c \\ \pi_i = c \left[\prod_{j=1}^{i-1} \frac{p_j}{q_j} + \prod_{j=1}^{i} \frac{p_j}{q_j} \right] &= \frac{c}{q_i} \prod_{j=1}^{i-1} \frac{p_j}{q_j} \ , \ 0 < i < n \\ \pi_n = c \prod_{j=1}^{n-1} \frac{p_j}{q_j} & \end{aligned}$$

where c is chosen to make $\sum_{j=0}^{n} \pi_j = 1$.

Problem 4.46 (a) Yes, it is a Markov chain. It suffices to construct the transition matrix and verify that the process has the Markov property. Let P^* be the new transition matrix. Then we have, for $0 \leq i \leq N$ and $0 \leq j \leq N$,

$$P^*_{i,j} = P_{i,j} + \sum_{k=N+1}^{\infty} P_{i,k} B^{(N)}_{k,j} ,$$

where $B^{(N)}_{k,j}$ is the probability of absorption into the absorbing state j in the absorbing Markov chain, where the states $N+1, N+2, \ldots$ are the transient states, while the state $1, 2, \ldots N$ are the N absorbing states. In other words, $B^{(N)}_{k,j}$ is the probability that the next state with index in the set $\{1, 2, \ldots, N\}$ visited by the Markov chain, starting with $k > N$ is in fact j. It is easy to see that the markov property is still present.

(b) The proportion of time in j is $\pi_j / \sum_{i=1}^{N} \pi_i$.

(c) Let $\pi_i(N)$ be the steady-state probabilities for the chain, only counting to visits among the states in the subset $\{1, 2, \ldots, N\}$. (This chain is necessarily positive recurrent.) By renewal theory,

$$\pi_i(N) = (E[\text{Number of}\, Y - \text{transitions between}\, Y - \text{ visits to}\, i)^{-1}$$

and

$$\begin{aligned}\pi_j(N) &= \frac{E[\text{No. } Y\text{-transitions to } j \text{ between } Y \text{ visits to } i]}{E[\text{No. } Y\text{-transitions to } i \text{ between } Y \text{ visits to } i]} \\ &= \frac{E[\text{No. } X\text{-transitions to } j \text{ between } X \text{ visits to } i]}{1/\pi_i(N)}\end{aligned}$$

(d) For the symmetric random walk, the new MC is doubly stochastic, so $\pi_i(N) = 1/(N+1)$ for all i. By part (c), we have the conclusion.

(e) It suffices to show that

$$\pi_i(N)P^*_{i,j} = \pi_j(N)P^*_{j,i}$$

for all i and j with $i \le N$ and $j \le N$. However, by above,

$$\pi_i(N)P^*_{i,j} = \pi_i(N)P_{i,j} + \pi_i(N)\sum_{k=N+1}^{\infty} P_{i,k}B^{(N)}_{k,j},$$

and

$$\pi_j(N)P^*_{j,i} = \pi_j(N)P_{j,i} + \pi_j(N)\sum_{k=N+1}^{\infty} P_{j,k}B^{(N)}_{k,i},$$

The two terms on the right are equal in these two displays. First, by the original reversibility, we have

$$\pi_i(N)P_{i,j} = \pi_j(N)P_{j,i}.$$

Second, by Theorem 4.7.2, we have

$$\pi_j(N)\sum_{k=N+1}^{\infty} P_{j,k}B^{(N)}_{k,i} = \pi_i(N)\sum_{k=N+1}^{\infty} P_{i,k}B^{(N)}_{k,j}.$$

We see that by expanding into the individual paths, and seeing that there is a reverse path.

Problem 4.47 Intuitively, in steady state each ball is equally likely to be in any of the urns and the positions of the balls are independent. Hence it seems intuitive that

$$\pi(\underline{n}) = \frac{M!}{n_1! \cdots n_m!}\left(\frac{1}{m}\right)^M .$$

To check the above and simultaneously establish time reversibility let

$$\underline{n}' = (n_1, \cdots, n_{i-1}, n_i - 1, n_{i+1}, \cdots, n_{j-1}, n_j + 1, n_{j+1}, \cdots, n_m)$$

and note that

$$
\begin{aligned}
\pi(\underline{n})P(\underline{n},\underline{n}') &= \frac{M!}{n_1!\cdots n_m!}\left(\frac{1}{m}\right)^M \frac{n_i}{M}\frac{1}{m-1} \\
&= \frac{M!}{n_1!\cdots(n_i-1)!\cdots(n_j+1)!\cdots n_m!}\left(\frac{1}{m}\right)^M \frac{n_j+1}{M}\frac{1}{m-1} \\
&= \pi(\underline{n}')P(\underline{n}',\underline{n})\ .
\end{aligned}
$$

Problem 4.48 **(a)** Each transition into i begins a new cycle. A reward of 1 is earned if state visited from i is j. Hence average reward per unit time is P_{ij}/μ_{ii}.

(b) Follows from (a) since $1/\mu_{jj}$ is the rate at which transitions into j occur.

(c) Suppose a reward rate of 1 per unit time when in i and heading for j. New cycle whenever enter i. Hence, average reward per unit time is $P_{ij}\eta_{ij}/\mu_{ii}$.

(d) Consider (c) but now only give a reward at rate 1 per unit time when the transition time from i to j is within x time units. Average reward is

$$
\begin{aligned}
\frac{\mathsf{E}[\text{Reward per cycle}]}{\mathsf{E}[\text{Time of cycle}]} &= \frac{P_{ij}\mathsf{E}[\min(X_{ij},x)]}{\mu_{ii}} \\
&= \frac{P_{ij}\int_0^x \bar{F}_{ij}(y)dy}{\mu_{ii}} \\
&= \frac{P_{ij}\eta_{ij}F_{ij}^e(x)}{\mu_{ii}}
\end{aligned}
$$

where $X_{ij} \sim F_{ij}$.

Problem 4.49

$$
\begin{aligned}
\lim_{t\to\infty}\mathsf{P}(S(t)=j|X(t)=i) &= \frac{\lim_{t\to\infty}\mathsf{P}(S(t)=j,X(t)=i)}{\mathsf{P}(X(t)=i)} \\
&= \frac{P_{ij}\int_0^\infty \bar{F}_{ij}(y)dy/\mu_{ii}}{P_i} \quad \text{by Theorem 4.8.4} \\
&= \frac{P_{ij}\eta_{ij}}{\mu_i}
\end{aligned}
$$

Problem 4.50 $\pi = (6,3,5)/14$, $\mu_1 = 25$, $\mu_2 = 80/3$, and $\mu_3 = 30$.

(a)

$$
\begin{aligned}
P_1 &= \frac{6\times 25}{6\times 25 + 3\times\frac{80}{3} + 5\times 30} = \frac{15}{38} \\
P_2 &= \frac{3\times\frac{80}{3}}{6\times 25 + 3\times\frac{80}{3} + 5\times 30} = \frac{8}{38} \\
P_3 &= \frac{5\times 30}{6\times 25 + 3\times\frac{80}{3} + 5\times 30} = \frac{15}{38}
\end{aligned}
$$

(b)

$$\mathsf{P}(\text{heading for 2}) \quad = \quad P_1 \frac{P_{12} t_{12}}{\mu_1} = \frac{15}{38} \times \frac{10}{25} = \frac{3}{19}$$

(c)

$$\text{fraction of time from 2 to 3} \quad = \quad P_2 \frac{P_{23} t_{23}}{\mu_2} = \frac{8}{38} \times \frac{60}{80} = \frac{3}{19}$$

IEOR 6711: Stochastic Models I, Professor Whitt

Solutions to Homework Assignment 12

Problem 5.3 **(a)** Let $N(t)$ denote the number of transitions be t. It is easy to show in this case that

$$\mathsf{P}(N(t) \geq n) \leq \sum_{j=n}^{\infty} e^{-Mt} \frac{(Mt)^j}{j!}$$

and thus $\mathsf{P}(N(t) < \infty) = 1$.

(b) Let X_{n+1} denote the time between the n-th and $(n+1)$-st transition and let J_n denote the n-th state visited. Also if we let

$$N(t) \triangleq \sup\{n : X_1 + \cdots + X_n \leq t\}$$

then $N(t)$ denotes the number of transitions by t. Now let j be the first recurrent state that is reached and suppose it was reached at the n_0-th transition (n_0 must be finite by assumption). Let $n_1, n_2, \cdots$ be the successive integers n at which $J_n = j$. (Such integers exist since j is recurrent.) Set $T_0 = X_1 + \cdots + X_{n_0}$, and

$$T_k \triangleq X_{n_{k-1}+1} + \cdots + X_{n_k} \ .$$

In other words, T_k denote the amount of time between the k-th and $(k+1)$-th visit to j. Therefore it follows that $\{T_k, k \geq 1\}$ forms a renewal process, and so $\sum_{k=1}^{\infty} T_k = \infty$ with probability 1. Since

$$\sum_{n=1}^{\infty} X_n = \sum_{k=0}^{\infty} T_k$$

it follows that $\sum_{n=1}^{\infty} X_n = \infty$.

Problem 5.4 Let T_i denote the time to go from i to $i+1$, $i \geq 0$. Then $\sum_{i=0}^{N-1} T_i$ is the time to go from 0 to N. Now T_i is exponential with rate λ_i and the T_i are independent. Hence

$$\mathsf{E}\left[e^{s\sum_{i=0}^{N-1} T_i}\right] = \prod_{i=0}^{N-1} \frac{\lambda_i}{\lambda_i - s} \ .$$

We may use it to compute the mean and variance or we can do directly and mean $=$ $1/\lambda_0 + \cdots + 1/\lambda_{N-1}$, variance $= 1/\lambda_0^2 + \cdots + 1/\lambda_{N-1}^2$.

Problem 5.9

$$\begin{aligned} P_{ij}(t+s) &= \sum_k \mathsf{P}(X(t+s) = j \mid X_0 = i, X(t) = k)\mathsf{P}(X(t) = k \mid X_0 = i) \\ &= \sum_k P_{kj}(s) P_{ik}(t) \ . \end{aligned}$$

Problem 5.10 (a)

$$\lim_{t\to 0}\frac{1-P(t)}{t} = v_0 \ .$$

(b) The first inequality follows from exercise 5.9 and

$$\begin{aligned} P(t+s) &= \mathsf{P}(X(t+s)=0|X(0)=0, X(s)=0)P(s) \\ &\quad +\mathsf{P}(X(t+s)=0|X(0)=0, X(s)\neq 0)(1-P(s)) \\ &\leq P(t)P(s)+1-P(s) \ . \end{aligned}$$

(c) From (b)

$$P(s)P(t-s) \leq P(t) \leq P(s)P(t-s)+1-P(t-s)$$

or

$$P(s)+P(t-s)-1 \leq P(t) \leq P(s)+1-P(t-s)$$

where the left hand inequality follows from

$$P(s)(1-P(t-s)) \leq 1-P(t-s) \ .$$

$\lim_{s\to t} P(s-t) = 1$ implies the continuity of P.

Problem 5.13

$$\prod_{j=i}^{i+k-1}\frac{\lambda_j}{\lambda_j+\mu_j} \ .$$

Problem 5.15 (a) Birth and death process.

(b) $\lambda_n = n\lambda+\theta$, $\mu_n = n\mu$.

(c) Set $M(t) = \mathsf{E}[X(t)|X(0)=i]$. Then

$$\mathsf{E}[X(t+h)|X(t)] = X(t)+(\lambda X(t)+\theta)h-\mu X(t)h+o(h)$$

and so $M(t+h) = M(t)+(\lambda-\mu)M(t)h+\theta h+o(h)$. Therefore, $M'(t) = (\lambda-\mu)M(t)+\theta$ or $e^{-(\lambda-\mu)t}[M'(t)-(\lambda-\mu)M(t)] = \theta e^{-(\lambda-\mu)t}$. Integrating both sides gives

$$e^{-(\lambda-\mu)t}M(t) = -\frac{\theta}{\lambda-\mu}e^{-(\lambda-\mu)t}+C$$

or

$$M(t) = Ce^{-(\lambda-\mu)t}-\frac{\theta}{\lambda-\mu} \ .$$

As $M(0)=i$ we obtain

$$M(t) = \frac{\theta}{\lambda-\mu}\left(e^{-(\lambda-\mu)t}-1\right)+ie^{-(\lambda-\mu)t} \ .$$

Problem 5.21 With the number of customers in the shop as the state, we get a birth and death process with $\lambda_0 = \lambda_1 = 3$, $\mu_1 = \mu_2 = 4$. Therefore $P_1 = \frac{3}{4}P_0$, $P_2 = \frac{3}{4}P_1 = \left(\frac{3}{4}\right)^2 P_0$. And since $P_0 + P_1 + P_2 = 1$, we get $P_0 = 16/37$.

(a)

$$P_1 + 2P_2 \quad = \quad \left[\frac{3}{4} + 2\left(\frac{3}{4}\right)^2\right] P_0 = \frac{30}{37}$$

(b) The proportion of customers that enter the shop is

$$\frac{\lambda(1-P_2)}{\lambda} = 1 - P_2 = 1 - \frac{9}{37} = \frac{28}{37} \ .$$

(c) $\mu = 8$, and so $P_0 = \frac{64}{97}$. So the proportion of customers who now enter the shop is

$$1 - P_2 = 1 - \left(\frac{3}{8}\right)^2 \frac{64}{97} = \frac{88}{97} \ .$$

The rate of added customers is therefore

$$\lambda\frac{88}{97} - \lambda\frac{28}{37} \simeq 0.45 \ .$$

The business he does would improve by 0.45 customers per hour.

IEOR 6711: Stochastic Models I, Professor Whitt

Solutions to Homework Assignment 13

Problem 5.12 **(a)** Since

$$P_0 = \frac{1/\lambda}{1/\lambda + 1/\mu} = \frac{\mu}{\lambda + \mu}$$

it follows that

$$\lim_{t \to \infty} \frac{N(t)}{t} = \frac{\alpha_0 \mu}{\lambda + \mu} + \frac{\alpha_1 \lambda}{\lambda + \mu} .$$

(b) The expected total time spent in state 0 by t is

$$\int_0^t P_{00}(s) ds = \frac{\mu}{\lambda + \mu} t + \frac{\lambda}{(\lambda + \mu)^2} \left(1 - e^{-(\lambda + \mu)t}\right) .$$

Calling the above $\mathsf{E}[T(t)]$ we have

$$\mathsf{E}[N(t)] = \alpha_0 \mathsf{E}[T(t)] + \alpha_1 (t - \mathsf{E}[T(t)]) .$$

Problem 5.19 Let R_0 be the time required to return to state 0 (starting in state 0), once it has been left (i.e., starting from the moment that the CTMC first leaves state 0). What is the expected value, $E[R_0]$?

The times spent in state 0 and away from state 0 constitute an alternating renewal process. So this can be answered by an elementary application of renewal theory. For that purpose, Let $\nu_0 = \sum_{j, j \neq 0} Q_{0,j} \equiv -Q_{0,0}$ and let α_j be the stationary or steady-state probability of state j. Then

$$\alpha_0 = \frac{1/\nu_0}{E[R_0] + (1/\nu_0)} ,$$

from which we see that

$$E[R_0] = \frac{1 - \alpha_0}{\alpha_0 \nu_0} .$$

Let Y be the time of the first transition, and condition on that time, to get

$$\begin{aligned} E[T] &= te^{-\nu_0 t} + \int_0^t E[T|Y = s] \nu_0 e^{-\nu_0 s} \, ds \\ &= te^{-\nu_0 t} + \int_0^t (s + E[R_0] + E[T]) \nu_0 e^{-\nu_0 s} \, ds . \end{aligned}$$

We can thus solve for $E[T]$, getting

$$E[T] = \frac{te^{-\nu_0 t} + [(1 - \alpha_0)/(\alpha_0 \nu_0)](1 - e^{-\nu_0 t}) + \int_0^t s \nu_0 e^{-\nu_0 s} \, ds}{e^{-\nu_0 t}} .$$

It would be OK to stop there, but you could go on to evaluate the integral. That leads to some simplification. Using the fact that $\nu_0^2 s e^{-\nu_0 s}$ is the density of the Erlang E_2 (special form of gamma, with shape parameter 2) distribution, we have

$$\int_0^t s\nu_0 e^{-\nu_0 s}\, ds = \frac{1-e^{-\nu_0 t}}{\nu_0} - \frac{\nu_0 t e^{-\nu_0 t}}{\nu_0}$$

or

$$\int_0^t s\nu_0 e^{-\nu_0 s}\, ds = \frac{1-e^{-\nu_0 t}}{\nu_0} - te^{-\nu_0 t} \, .$$

When we apply the algebra, we get

$$E[T_0] = \frac{e^{\nu_0 t}-1}{\alpha_0 \nu_0} \, .$$

Problem 5.22 The analysis on page 153-154 with

$$\begin{aligned} \lambda_n &= \quad \lambda \quad n \geq 0 \\ \mu_n &= \begin{cases} n\mu & 1 \leq n \leq s \\ s\mu & n > s \end{cases} \end{aligned} \, .$$

We need $\lambda < s\mu$.

Problem 5.24 Let $X_i(t)$ denote the number of customers at server i, $i = 1, 2$, when there is unlimited waiting room. The, in steady state,

$$\mathsf{P}(n_i \text{ at server } i) = \prod_{i=1}^{2} \left(\frac{\lambda_i}{\mu_i}\right)^{n_i} \left(1 - \frac{\lambda_i}{\mu_i}\right) \, .$$

Now the model under consideration is just a truncation of the above, which is time reversible by problem 5.23. The truncation is to the set

$$A \triangleq \{(n,m) : n = 0, m \leq N+1 \text{ or } m = 0, n \leq N+1 \text{ or } n > 0, m > 0, n+m \leq N+2\} \, .$$

Hence, for $(n, m) \in A$,

$$P(n,m) = C \left(\frac{\lambda_1}{\mu_1}\right)^n \left(1 - \frac{\lambda_1}{\mu_1}\right) \left(\frac{\lambda_2}{\mu_2}\right)^m \left(1 - \frac{\lambda_2}{\mu_2}\right) \, .$$

Problem 5.25 In steady state it has the same probability structure as the arrival process. Hence if we include in the departure process those arrivals that do not enter, then it is a Poisson process.

Problem 5.26 (a) Follows from results of section 6.2 by writing $\underline{n}' = D_j \underline{n}$ and so $\underline{n} = B_{j-1} \underline{n}'$.

(b) A Poisson process by time reversibility. If $D(0) = 0$, it is a nonhomogeneous Poisson process.

Problem 5.28 For $\underline{n} = (n_1, \cdots, n_r)$, let

$$P(\underline{n}) = C \prod_{i=1}^{r} \alpha_i^{n_i} \ .$$

If

$$\underline{n}' = (n_1, \cdots, n_i - 1, \cdots, n_j + 1, \cdots, n_r)$$

with $n_i > 0$ then

$$\begin{aligned} & P(\underline{n})q(\underline{n}, \underline{n}') = P(\underline{n}')q(\underline{n}', \underline{n}) \\ \Leftrightarrow \quad & \alpha_i \frac{\mu_i}{r-1} = \alpha_j \frac{\mu_j}{r-1} \\ \Leftrightarrow \quad & \alpha_i \mu_i = \alpha_j \mu_j \ . \end{aligned}$$

So, setting $\alpha_i = 1/\mu_i$, $i = 1, \cdots, r$ and letting C be such that $\sum_{\underline{n}} P(\underline{n}) = 1$ the time reversibility equations are satisfied.

Problem Extra

```
Program = Calculations for the M/M/s/r=M Model following 1999 Improving paper
WhoWhen = Ward Whitt 11/17/03

Step1 = Model Parameters
s = number of servers, s = 100
lambda = arrival rate, lambda =   100
mu = individual service rate, mu = 1
alpha = individual abandonment rate (exponential abandonments), alpha =    1
r = number of extra waiting spaces, r =   100

Step2 = Key Steady-state Probabilities
ProbNoWait = Probability of not having to wait before beginning service ProbNoWait = 0.48670120172085
ProbAllBusy = Probability that all servers are busy upon arrival ProbAllBusy = 0.51329879827915
ProbWaitandServed = 0.47343780147000
ProbServed = 0.96013900319085
ProbLoss = Probability that an arrival is lost (blocked) at arrival ProbLoss = 4.716970602792618e-019
ProbAban = Probability that a customer eventually abandons ProbAban = 0.03986099680915

Step3 = Mean Values - Counting
MeanInQueue =   3.98609968091471
MeanBusyServers =  96.01390031908522
MeanNumberinSys = 99.99999999999993

Step4 = Second Moments and Variances - Counting
SecondMomInQueue = 51.32987982791486
VarInQueue =  35.44088916172650
SecondMomBusyServers =    9.251450183989135e+003
VarBusyServers = 32.78112950590003
SecondMomNumberInSys =    1.010000000000000e+004
VarNumberInSys =    1.000000000000109e+002

Step5 = Response-Time Moments
MeanRespTime =   0.99750428450446
ConditMeanRespTimeServed =   1.03891653311596
SecMomRespTime = 2.94660609401905
VarRespTime =   1.95159129641430
ConditSecMomRespTimeServed =   3.06893698123555
ConditVarRespTimeServed =   1.98958941845386

Step6 = Waiting-Time Moments
Step6a = Waiting Times for Customers who are Served
MeanWaitServed =   0.03736528131360
ConditMeanWaitServed =   0.03891653311596
```

```
SecMomWaitandServed = 0.00472192207013
VarWaitandServed =   0.00332575782249
ConditSecMomWaitServed =   0.00491795672756
ConditVarWaitServed = 0.00340346017779

Step6b = Waiting Times for Customers who Abandon
MeanWaitAbandon = 0.00249571549554
ConditMeanWaitAban =   0.06261046374463
SecMomWaitAbandon =    2.695089209549688e-004
ConditSecMomWaitAbandon =   0.00676121879855
ConditVarWaitAban =   0.00284114862823

Step6c = Waiting Time for All Customers, Served or Abandoning
MeanWaitTimeAll =   0.03986099680915
SecMomWaitTimeAll =   0.00499143099109
VarWaitTimeAll =   0.00340253192447

Step7 = Conditional Waiting Time CCDF Values For Served Customers
t =   0.05000000000000 ProbOKWaitifServed =   0.70144791424012
t =   0.10000000000000 ProbOKWaitifServed =   0.84743410401854
t =   0.20000000000000 ProbOKWaitifServed =   0.97658087743471
t =   0.40000000000000 ProbOKWaitifServed =   0.99993119489496

Step8 = Conditional Waiting Time CDF Values For Customers Who Abandon
t =   0.05000000000000 ProbOKWaitifAbandon =   0.50953673688750
t =   0.10000000000000 ProbOKWaitifAbandon =   0.79190403053181
t =   0.20000000000000 ProbOKWaitifAbandon =   0.97645578331304
t =   0.40000000000000 ProbOKWaitifAbandon =   0.99995439602867

Step9 = Waiting Time CDF Values For All Customers
t =   0.05000000000000 ProbOKWait =   0.69379814341203
t =   0.10000000000000 ProbOKWait =   0.84522061993647
t =   0.20000000000000 ProbOKWait =   0.97657589105832
t =   0.40000000000000 ProbOKWait =   0.99993211971528
```

(a) MeanInQueue = 3.98609968091471, VarInQueue = 35.44088916172650

(b) arrival rate $\times$ ProbAban = 100 $\times$ 0.03986099680915= 3.986

(c) ProbNoWait = 0.48670120172085

(d) ConditMeanWaitServed + service time = 0.03891653311596 + 1

(e) Conditional Waiting Time CDF Values For Served Customers at $t = 0.1$ = ProbOK-WaitifServed = 0.84743410401854

(f) Since the abandon rate α equals the service rate μ, the model simplifies. But note that the simplification and following solution only works for the special case of $\alpha = \mu$. First, in the view point of *the number of customers in the system*, the system is equivalent to $M/M/s+r/0$ since the departure rate from the system including the abandoned customers is $n\mu$ even if $n > s$ because of $(n-s)\alpha = (n-s)\mu$ abandonment rate. However, we can anticipate that blocking is negligible with such a large waiting space. Thus the $M/M/s+r/0$ model should be essentially equivalent to a $M/M/\infty$ model, for which the steady-state distribution is exactly Poisson, and approximately normal. Hence, the probability all servers are busy in the given $M/M/s/r+M$ model is approximately equal to the probabiity that at least s servers are busy in the $M/M/\infty$ model. Clearly, it is possible to compute the probability that at least s servers are busy in the $M/M/\infty$ mode very quickly. (It might not be regarded as "simple," however. When you read, "it is easy to see that . . . ," in a paper, take that as an invitation to check very carefully.)

Now we describe how to proceed with the more special, but exact, representation in terms of the $M/M/s+r/0$ model: For $n < s+r = 200$

$$\begin{aligned} P_n &= \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n \left(1 + \sum_{k=1}^{s+r} \frac{\lambda_0 \cdots \lambda_{k-1}}{\mu_1 \cdots \mu_k}\right)} \\ &= \frac{\lambda^n}{n! \left(\sum_{k=0}^{s+r} \frac{\lambda^k}{k!}\right)} \\ &= \frac{e^{-\lambda} \frac{\lambda^n}{n!}}{\left(\sum_{k=0}^{s+r} e^{-\lambda} \frac{\lambda^k}{k!}\right)} \end{aligned}$$

and we have (for Poisson random variable N with parameter $\lambda = 100$)

$$\begin{aligned} \mathsf{P}(\text{All servers are busy}) &= \sum_{n=0}^{s-1} P_n \\ &= \frac{\mathsf{P}(N \le 99)}{\mathsf{P}(N \le 200)} \\ &= 0.48670120172087 \quad \text{from MATLAB} \end{aligned}$$

(g) MeanInQueue$\times$ abandon rate per customer in queue
$= 3.98609968091471 \times 1 = 3.98609968091471$

IEOR 6711: Professor Whitt

Notes on Laplace Transforms and Their Inversion

"The shortest path between two truths in the real domain passes through the complex domain;" Jacques Hadamard (1865-1963).

1. Basic Definition

Let X be a nonnegative real-valued random variable with probability density function (pdf) $f \equiv f_X$. Then the Laplace transform of the random variable X, and also the Laplace transform of the pdf f, is

$$E[e^{-sX}] \equiv \hat{f}(s) \equiv \int_0^\infty e^{-st} f(t)\, dt \ , \tag{1}$$

where s is a complex variable with nonnegative real part. (If we write $s = u + vi$, where $i \equiv \sqrt{-1}$ and u and v are real numbers, then u is $Re(s)$ (the real part of s) and v is $Im(s)$ (the imaginary part of s. Google it if this is new to you. We assume that $u \geq 0$, which guarantees that the integral is well defined provided that the function f is itself integrable.)

If we replaced $-s$ by $+s$, where we regard s as real, we would have the moment generating function (mgf), discussed in Ross; see Section 1.4. A problem with the mgf is that it is not always finite. For example if $f(t) = A/(1+t)^p$, for $t \geq 0$, where $A > 0$ and $p > 1$ are parameters, then the mgf is infinite.

But the Laplace transform is well defined for any nonnegative random variable (probability distribution on the positive halfline $[0, \infty)$.

2. Project on Numerical Inversion

You are asked to do a short project on numerical inversion. Associated with that project is a numerical homework. It concerns learning about a method to numerically invert Laplace transforms. The idea is to write your own code, presumably using MATLAB, but you could use another language, C++, say. You will need to work with complex numbers. There is a specific homework assignment for this optional assignment, which involves solving some problems with your code. For this, you can follow the papers on the computational tools web page. In particular, you should look at the 1995 JoC paper.

3. Project Management

We now illustrate how numerical inversion can be applied. We discuss Example 1.1.1 on pages 3-5 of the book chapter by Abate, Choudhury and Whitt (1999), referred to here as ACW, given on the tools web page. Suppose that we are managing a project, consisting of four steps. We want to know the probability distribution of the time T to complete the entire project. Let X_i be the time required to complete step i of the project, $i = 1, \ldots, 4$. Assume that the total time can be represented as the sum of the four separate steps, i.e.,

$$T = X_1 + X_2 + X_3 + X_4 \ . \tag{2}$$

We are thus assuming one task is done after the other. We have to finish task 1 before we start task 2, and so forth. Moreover, suppose that the four step times X_i are **mutually independent random variables**, with known pdf's f_i. We want to compute the cdf of T; i.e., we want to calculate

$$G(t) \equiv P(T \le t), \quad t \ge 0 . \tag{3}$$

Directly, the cdf G can be expressed as a relatively complicated convolution integral, as shown in (1.1) of ACW. That single integral extends to a multiple (three-dimensional) integral when we consider the sum of 4 random variables or, equivalently, the convolution of 4 distributions.

An attractive alternative approach is to exploit Laplace transforms. We can write

$$\hat{g}(s) = \hat{f}_1(s) \times \hat{f}_2(s) \times \hat{f}_3(s) \times \hat{f}_4(s) . \tag{4}$$

where $\hat{f}_i(s)$ is the Laplace transform of the pdf f_i, as defined above in (1) and

$$\hat{g}(s) = E[e^{-sT}] = \int_0^\infty e^{-st} g(t)\, dt , \tag{5}$$

where g is the pdf of T. To see that is so, note that

$$E[e^{-s(X_1+\cdots+X_4)}] = E[e^{-s(X_1)}e^{-s(X_2)}e^{-s(X_3)}e^{-s(X_4)}] = \hat{f}_1(s) \times \hat{f}_2(s) \times \hat{f}_3(s) \times \hat{f}_4(s) . \tag{6}$$

We use independence to have the expectation of the product be equal to the product of the expectations; see p. 8 of Ross. You want more than is given there. You want to know that $E[h_1(X_1)h_2(X_2)] = E[h_1(X_1)]E[h_2(X_2)]$ for all functions h_1 and h_2 if X_1 and X_2 are independent.

The above shows that we can calculate the Laplace transform of T, denoted by $\hat{g}(s)$, simply as the product of the Laplace transforms of X_i. Since the cdf G is the integral of the pdf g, the Laplace transforms are related by

$$\hat{G}(s) \equiv \int_0^\infty e^{-st} G(t)\, dt = \frac{\hat{g}(s)}{s} . \tag{7}$$

(The last formula is a standard formula for Laplace transforms, shown by using integration by parts.) Hence, we can calculate $G(t)$ for any t of interest by numerically inverting its Laplace transform, $\hat{G}(s) = \hat{g}(s)/s$. That can be done quite quickly with an inversion program. See the papers on line for discussion.

Sometimes we are not given the component distributions or transforms, but instead are given only estimates of the first two moments. The mean is easy:

$$E[T] = E[X_1] + E[X_2] + E[X_3] + E[X_4] .$$

Since the random variables are independent, the variances add as well:

$$Var[T] = Var[X_1] + Var[X_2] + Var[X_3] + Var[X_4] .$$

But that does not give us the distribution of T. We could fit a distribution to the first two moments (or mean and variance) of T, given the relations above, but it is often better to fit appropriate distributions to the individual X_i, and then calculate their transforms, and perform the inversion above. For example, we might fit the distribution of X_i to a gamma distribution. The transform of a gamma distribution is

$$\hat{f}(s) = \left(\frac{\lambda}{\lambda + s}\right)^{\nu} ,$$

where λ and ν are positive parameters. The mean is ν/λ and the variance is ν/λ^2. We can thus easily fit the parameters ν and λ to the mean and variance. We then can carry out the inversion described above.

We close by discussing inversion, giving reference to the 1995 JoC paper.

4. A brief recap of the Euler algorithm, with reference to the 1995 JoC paper.

One way to start is with the **Bromwich contour integral**. It gives the desired function value $f(t)$ in terms of the transform $\hat{f}(s)$, see (2) on page 37. In (2) we show that the contour integral is easily represented as an integral of a real-valued function of a real variable over the positive half line $[0, \infty)$. That real integral is given in the last line of (2). One tricky step in getting there is that we must use that the inversion integral yields the value 0 for $t < 0$. That is part of the Bromwich theorem. That is used in the last step in (2), as well as basic properties of even and odd functions. The steps in (2) are better explained in the 1999 survey paper on page 7.

From (2), it suffices to perform the numerical integration, also called numerical quadrature. There are two steps, discretizing to get a series and truncating the infinite series to get a finite sum. In the end we use a convergence-improvement technique instead of simple sum. That produces a finite weighted sum. The numerical-integration technique we use is the **trapezoidal rule**. That is usually thought to be a primitive method, but it turns out to be effective in this context. The trapezoidal rule yields formula (4). A key step is to determine the **discretization error** in applying the trapezoidal rule with step size h. That is addressed by aliasing below. Given (4), we perform a change of variables to force all the cosine terms to be $+1$, -1 or 0. The result is (5). We do that to get an alternating series (more precisely, a series that is likely to be eventually an alternating series), which is convenient for convergence improvement.

An alternative approach to the inversion is via Fourier series. As motivation, we can observe that (4) is in fact a trigonometric series; i.e., it is a Fourier series. We might thus ask what is the function that it is the Fourier series of? You can find notes on Fourier series from Mathworld. We obtain Fourier series expansions of functions on a finite interval or of functions which are periodic. That is, given a periodic function, we can construct its Fourier series. One approach is to determine the function for which (4) is its Fourier series.

However, we need not even know about the Bromwich integral. We could instead start with f and construct a periodic function and take its Fourier series. If done right, expression (4) will be the Fourier series of that periodic function. We can construct a periodic function of an arbitrary given function by **aliasing**, that is, by adding infinitely many translates of the function. Starting from a function g, the associated periodic function is g_p given in (6). However, we must do something to guarantee that the series converges. To get convergence, we first damp our original function by writing

$$g(t) = e^{-bt} f(t), \quad t \geq 0, \quad \text{for} \quad b > 0 \ .$$

Assuming that f is well behaved, e.g., if f is bounded, then the series in the definition of g_p will converge.

We observe that g_p is periodic. Then we construct the complex Fourier series of that periodic function, given in (7). A critical property is that the coefficients of the Fourier series can be expressed in terms of the transform values, which is precisely what we have to work with.. That is shown in (8). The two steps together - aliasing and Fourier series - give us the **Poisson summation formula**, given in (9). That is a classic formula, with many many applications. We observe that $f(t)$ itself is obtained by looking at the single term on the left in (9) corresponding to $k = 0$.

We then perform a change of variables, and obtain (10), which gives us once again formula (5), but with an explicit expression for the discretization error. We get that simple representation because $f(t) = 0$ for $t < 0$. We then get a simple bound on the discretization error in (12), under the assumption that $|f(t) \le 1$ for all t.

The remaining problem is to sum the infinite series in (5). We could do so by simple truncation, but it turns out to be much better to do a **convergence improvement technique**. Having made (5) a nearly alternating series, a natural candidate is Euler summation. We take the weighted average of the last 12 terms (with k ranging from 0 to $m = 11$). Specifically, we take a binomial average of those last 12 terms, as shown in (15). I give you simple exercises in the homework, so that you can see the power of Euler summation. See Section 1.2.3 on p. 17-19 of the Summary paper for more on Euler summation.

In general, the algorithm is given in formulas (13)–(15). There are three parameters A, n and m. We use A to control the discretization error, but we are not too greedy; we avoid roundoff error by not choosing A too large. (Roundoff error occurs in (13) because of the multiplication by $e^{A/2}$. Roundoff error is discussed on page 15 of the 1999 summary.) Reasonable values of the parameters are $A = 19$, $n = 38$ and $m = 11$. There is of course another parameter: t. That is the argument of the function we wish to compute: We are computing $f(t)$. The algorithm is summarized on page 16 of the 1999 summary article. Pseudo-code (actually UBASIC code) appears in the left-hand box on page 41 of the 1995 JoC paper.

5. Other Inversion Methods.

There are other inversion techniques. We may apply the Bromwich integral in a different way. For example, Talbot's method is based on trying to more carefully choose the contour of integration. But we do not even need to start with the Bromwich integral. We illustrate by briefly describing an entirely different inversion formula, involving differentation instead of integration.

The **Post-Widder formula** is

$$f(t) = \lim_{n \to \infty} \frac{(-1)^n}{n!} ((n+1)/t)^{n+1} \hat{f}^{(n)}((n+1)/t) \ ,$$

as given in Theorem 2 of the book chapter and (20) of the 1996 paper.

This formula can be derived and understood as

$$\frac{(-1)^n}{n!} ((n+1)/t)^{n+1} \hat{f}^{(n)}((n+1)/t) = E[f(X_{n,t})] \ ,$$

where $X_{n,t}$ is a random variable with a gamma distribution having mean t and variance $t^2/(n+1)$. Since the variance goes to 0 as $n \to \infty$, we have

$$X_{n,t} \Rightarrow t \quad \text{as} \quad n \to \infty \ .$$

Indeed, that can be proved from Chebyshev's inequality:

$$P(|X_{n,t} - t| > \epsilon) \le \frac{Var(X_{n,t})}{\epsilon^2} \ .$$

Since $Var(X_{n,t}) \to 0$ as $n \to \infty$, for any $\epsilon > 0$,

$$P(|X_{n,t} - t| > \epsilon) \to 0 \quad \text{as} \quad n \to \infty \ .$$

Under extra regularity conditions (e.g., uniform integrability),

$$E[f(X_{n,t})] \to t \quad \text{as} \quad n \to \infty \ .$$

To see that indeed

$$\frac{(-1)^n}{n!}((n+1)/t)^{n+1}\hat{f}^{(n)}((n+1)/t) = E[f(X_{n,t})] \ ,$$

start by writing

$$(-1)^n \hat{f}^{(n)}(\lambda) = \int_0^\infty e^{-\lambda x} x^n f(x)\, dx \ .$$

Next let $\lambda = (n+1)/t$ to get

$$(-1)^n \hat{f}^{(n)}((n+1)/t) = \int_0^\infty e^{-[(n+1)/t]x} x^n f(x)\, dx \ .$$

Then multiply both sides by $((n+1)/t)^{n+1}$ and divide both sides by $n!$ to get the desired equation.