CS 70 Discrete Mathematics and Probability Theory
Spring 2017 Rao

HW 13

1 Sundry

Before you start your homework, write down your team. Who else did you work with on this homework? List names and email addresses. (In case of homework party, you can also just describe the group.) How did you work on this homework? Working in groups of 3-5 will earn credit for your "Sundry" grade.

Please copy the following statement and sign next to it:

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up.

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up. (Signature here)

2 Function of a Markov Chain

Let X_n be a Markov chain on the state space $\mathscr{X}$ and let $f : \mathscr{X} \to \mathscr{X}$ be a function on the state space. Give an example to show that $f(X_n)$ may not be a Markov chain. [*Hint*: A sequence Y_n is not a Markov chain if it does not satisfy the Markov property

$$\Pr(Y_{n+1} = j \mid Y_n = i, Y_{n-1} = i_{n-1}, \ldots, Y_0 = i_0) = \Pr(Y_{n+1} = j \mid Y_n = i)$$

for all $i_0, \ldots, i_{n-1}, i, j \in \mathscr{X}$ and $n \in \mathbb{N}$.]

Solution:

Here is a simple example. Consider a Markov chain X_n on $\{0,1,2\}$ with

$$P(0,1) = P(1,2) = P(2,0) = 1.$$

Define $g(0) = g(1) = 0$ and $g(2) = 1$. If $\pi_0 = \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix}$, then $Y_n = g(X_n)$ is not a Markov chain. To see this, note that

$$\Pr[Y_2 = 0 \mid Y_1 = 0, Y_0 = 0] = 0 \neq \Pr[Y_2 = 0 \mid Y_1 = 0] = 0.5.$$

The intuition is that $g(X_n)$ contains less information than X_n if $g(\cdot)$ is many-to-one. Thus, although X_n contains all the information needed to predict its future, the same is not true of $g(X_n)$.

3 Reflecting Random Walk

Alice starts at vertex 0 and wishes to get to vertex n. When she is at vertex 0 she has a probability of 1 of transitioning to vertex 1. For any other vertex i, there is a probability of $1/2$ of transitioning to $i+1$ and a probability of $1/2$ of transitioning to $i-1$.

(a) What is the expected number of steps Alice takes to reach vertex n? Write down the hitting-time equations, but do not solve them yet.

(b) Solve the hitting-time equations. [*Hint*: Let R_i denote the expected number of steps to reach vertex n starting from vertex i. As a suggestion, try writing R_0 in terms of R_1; then, use this to express R_1 in terms of R_2; and then use this to express R_2 in terms of R_3, and so on. See if you can notice a pattern.]

Solution:

Formulate hitting time equations; the hard part is solving them. R_i represents the expected number of steps to get to vertex n starting from vertex i. In particular, $R_n = 0$ and we are interested in calculating R_0. We have the equations:

$$\begin{aligned}
R_0 &= 1 + R_1, \\
R_1 &= 1 + \frac{1}{2}R_0 + \frac{1}{2}R_2, \\
&\vdots \\
R_i &= 1 + \frac{1}{2}R_{i-1} + \frac{1}{2}R_{i+1}, \\
&\vdots \\
R_{n-1} &= 1 + \frac{1}{2}R_{n-2} + \frac{1}{2}R_n.
\end{aligned}$$

Plug in $R_0 = 1 + R_1$ to the second equation: $R_1 = 1 + 1/2 + (1/2)R_1 + (1/2)R_2$ which then implies $R_1 = 3 + R_2$. In fact, if $R_i = k + R_{i+1}$, then

$$R_{i+1} = 1 + \frac{1}{2}R_i + \frac{1}{2}R_{i+2} = 1 + \frac{k}{2} + \frac{1}{2}R_{i+1} + \frac{1}{2}R_{i+2},$$

which, after moving $(1/2)R_{i+1}$ to the left and multiplying by two, implies $R_{i+1} = k + 2 + R_{i+2}$.

Therefore, $R_0 = 1 + R_1 = 1 + 3 + R_2 = 1 + 3 + 5 + R_3 = \cdots = 1 + 3 + \cdots + 2n - 1 + R_n$ and since $R_n = 0$, we have $R_0 = n^2$.

4 Boba in a Straw

Imagine that Wan Fung is drinking milk tea and he has a very short straw: it has enough room to fit two boba (see Figure 1).

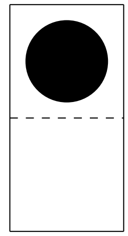


Figure 1: A straw with one boba currently inside. The straw only has enough room to fit two boba.

Here is a formal description of the drinking process: We model the straw as having two "components" (the top component and the bottom component). At any given time, a component can contain nothing, or one boba. As Wan Fung drinks from the straw, the following happens every second:

1. The contents of the top component enter Wan Fung's mouth.
2. The contents of the bottom component move to the top component.
3. With probability p, a new boba enters the bottom component; otherwise the bottom component is now empty.

Help Wan Fung evaluate the consequences of his incessant drinking!

(a) At the very start, the straw starts off completely empty. What is the expected number of seconds that elapse before the straw is completely filled with boba for the first time? [Write down the equations; you do not have to solve them.]

(b) Consider a slight variant of the previous part: now the straw is narrower at the bottom than at the top. This affects the drinking speed: if either (i) a new boba is about to enter the bottom component or (ii) a boba from the bottom component is about to move to the top component, then the action takes two seconds. If both (i) and (ii) are about to happen, then the action takes three seconds. Otherwise, the action takes one second. Under these conditions, answer the previous part again. [Write down the equations; you do not have to solve them.]

(c) Wan Fung was annoyed by the straw so he bought a fresh new straw (the straw is no longer narrow at the bottom). What is the long-run average rate of Wan Fung's calorie consumption? (Each boba is roughly 10 calories.)

(d) What is the long-run average number of boba which can be found inside the straw? [Maybe you should first think about the long-run distribution of the number of boba.]

Solution:

(a) We model the straw as a four-state Markov chain. The states are $\{(0,0),(0,1),(1,0),(1,1)\}$, where the first component of a state represents whether the top component is empty (0) or full

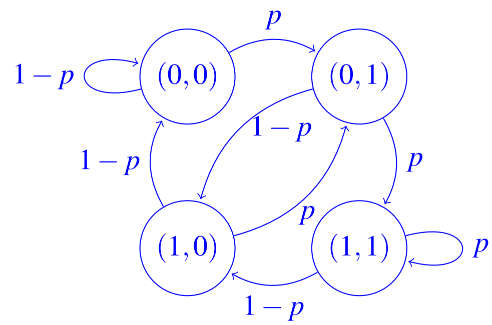


Figure 2: Transition diagram for the Markov chain.

(1); similarly, the second component represents whether the bottom component is empty or full. See Figure 2.

Now, we set up the hitting time equations. Let T denote the time it takes to reach state $(1,1)$, i.e. $T = \min\{n > 0 : X_n = (1,1)\}$. Let $\mathbf{E}_i[\cdot] = \mathbf{E}[\cdot \mid X_0 = i]$ denote the expectation starting from state i (for convenience of notation). The hitting-time equations are

$$\begin{aligned}
\mathbf{E}_{(0,0)}[T] &= 1 + (1-p)\mathbf{E}_{(0,0)}[T] + p\mathbf{E}_{(0,1)}[T],\\
\mathbf{E}_{(0,1)}[T] &= 1 + (1-p)\mathbf{E}_{(1,0)}[T] + p\mathbf{E}_{(1,1)}[T],\\
\mathbf{E}_{(1,0)}[T] &= 1 + (1-p)\mathbf{E}_{(0,0)}[T] + p\mathbf{E}_{(0,1)}[T],\\
\mathbf{E}_{(1,1)}[T] &= 0.
\end{aligned}$$

The question did not ask you to solve the equations. If you solved the equations anyway and would like to check your work, the hitting time is $\mathbf{E}_{(0,0)}[T] = (1+p)/p^2$.

(b) The new hitting-time equations are

$$\begin{aligned}
\mathbf{E}_{(0,0)}[T] &= (1-p)(1+\mathbf{E}_{(0,0)}[T]) + p(2+\mathbf{E}_{(0,1)}[T]),\\
\mathbf{E}_{(0,1)}[T] &= (1-p)(2+\mathbf{E}_{(1,0)}[T]) + p(3+\mathbf{E}_{(1,1)}[T]),\\
\mathbf{E}_{(1,0)}[T] &= (1-p)(1+\mathbf{E}_{(0,0)}[T]) + p(2+\mathbf{E}_{(0,1)}[T]),\\
\mathbf{E}_{(1,1)}[T] &= 0.
\end{aligned}$$

You did not have to solve the equations, but to get a sense for what the solution is like, solving the equations and plugging in $p = 1/2$ yields (after some tedious algebra) $\mathbf{E}_{(0,0)}[T] = 11$.

(c) This part is actually more straightforward than it might initially seem: the average rate at which Wan Fung consumes boba must equal the average rate at which boba enters the straw, which is p per second. Hence, his long-run average calorie consumption rate is $10p$ per second.

(d) We compute the stationary distribution. The balance equations are

$$\begin{aligned}
\pi(0,0) &= (1-p)\pi(0,0) + (1-p)\pi(1,0),\\
\pi(0,1) &= p\pi(0,0) + p\pi(1,0),\\
\pi(1,0) &= (1-p)\pi(0,1) + (1-p)\pi(1,1),\\
\pi(1,1) &= p\pi(0,1) + p\pi(1,1).
\end{aligned}$$

Expressing everything in terms of $\pi(0,0)$, we find

$$\pi(0,1) = \pi(1,0) = \frac{p}{1-p}\pi(0,0),$$
$$\pi(1,1) = \frac{p^2}{(1-p)^2}\pi(0,0).$$

From the normalization condition we have

$$\pi(0,0)\left(1+\frac{2p}{1-p}+\frac{p^2}{(1-p)^2}\right) = 1,$$

so $\pi(0,0) = (1-p)^2$. Hence, the stationary distribution is

$$\pi(0,0) = (1-p)^2,$$
$$\pi(0,1) = \pi(1,0) = p(1-p),$$
$$\pi(1,1) = p^2.$$

In states $(0,1)$ and $(1,0)$, there is one boba in the straw; in state $(1,1)$, there are two boba in the straw. Therefore, the long-run average number of boba in the straw is

$$\pi(0,1)+\pi(1,0)+2\pi(1,1) = 2p(1-p)+2p^2 = 2p.$$

Alternate Solution: The goal of the question was to have you work through the balance equations, but there is a simple solution. Observe that at any given time after at least two seconds have passed, each component has probability p of being filled with boba. Therefore, the number of boba in the straw is like a binomial distribution with 2 independent trials and success probability p, so the average number of boba in the straw is $2p$.

5 Online Product Sampling

Suppose you are manufacturing bags of coffee beans. You want to guarantee that your product tastes amazing. In fact, the manufacturing process is not perfect, and a fraction p of your bags do not taste amazing. In the industry, it is crucial for you to perform inspections.

One strategy for performing inspections is the following:

1. **High Sampling**: Inspect every bag of coffee until m consecutive bags taste amazing; then switch to low sampling.
2. **Low Sampling**: Sample one bag out of every k bags uniformly at random (in other words, for each bag, sample it with probability $1/k$); if the sample does not taste amazing, then switch to high sampling.

[As a side-note: why do we not simply inspect every bag? The answer is that an inspected bag of coffee can no longer be sold.]

We will use a Markov chain model on $\mathscr{X} = \{0,\dots,m\}$. The state $i \in \{0,\dots,m-1\}$ represents the **High Sampling** state where we have tasted i consecutive amazing bags. The state m represents the **Low Sampling** state.

(a) Write down the balance equations, but do not solve them.

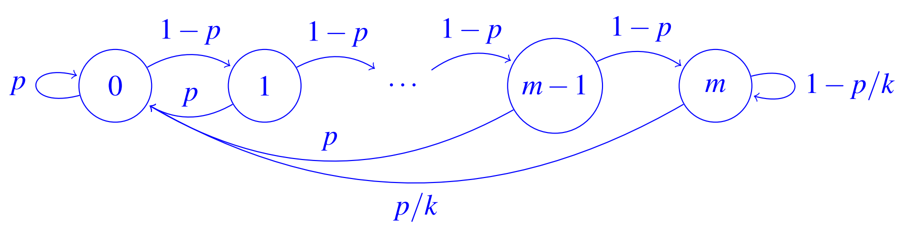
(b) For the rest of this problem, you may assume that $m = 2$ (so we have a three-state Markov chain). Compute the stationary distribution π.

(c) Under this inspection strategy, what is the long-run fraction of bags which are inspected?

(d) Under this inspection strategy, what is the long-run fraction of bags of coffee which are sold (i.e. not inspected) and don't taste amazing?

Solution:

(a) The Markov chain diagram is:



Explanation: from a state $i \in \{0,\dots,m-1\}$, there is a $1-p$ chance that we inspect a bag of coffee which tastes amazing, moving to state $i+1$; otherwise we inspect a bag of coffee that does not taste amazing, so we go back to 0. In state m, there is a $1/k$ chance that that we sample the bag of coffee, and a further p chance that the bag of coffee does not taste amazing; hence, with probability p/k we move to state 0. Otherwise, we remain at state m.

The balance equations are:

$$\begin{aligned}
\pi(0) &= p\sum_{i=0}^{m-1}\pi(i) + \frac{p}{k}\pi(m),\\
\pi(i) &= (1-p)\pi(i-1), \qquad i = 1,\dots,m-1,\\
\pi(m) &= (1-p)\pi(m-1) + \left(1-\frac{p}{k}\right)\pi(m).
\end{aligned}$$

Along with the balance equations, we include the normalization condition $\sum_{i=0}^{m}\pi(i) = 1$.

(b) We know that $\pi(1) = (1-p)\pi(0)$. From the last equation, $(p/k)\pi(2) = (1-p)\pi(1)$, or

$$\pi(2) = \frac{k}{p}(1-p)\pi(1) = \frac{k}{p}(1-p)^2\pi(0).$$

We have expressed everything in terms of $\pi(0)$, so we don't need the first balance equation (remember that one of the balance equations is always redundant). Instead, we will use the normalization condition $\pi(0)+\pi(1)+\pi(2)=1$, or

$$\pi(0)\left(1+(1-p)+\frac{k}{p}(1-p)^2\right)=1.$$

Therefore,

$$\pi(0)=\left(2-p+\frac{k}{p}(1-p)^2\right)^{-1}=\frac{p}{1+(k-1)(1-p)^2}.$$

Hence, the stationary distribution is

$$\begin{aligned}\pi(0)&=\frac{p}{1+(k-1)(1-p)^2},\\ \pi(1)&=\frac{p(1-p)}{1+(k-1)(1-p)^2},\\ \pi(2)&=\frac{k(1-p)^2}{1+(k-1)(1-p)^2}.\end{aligned}$$

(c) When we are in states 0 or 1, we inspect each bag. In state 2, we only inspect a fraction $1/k$ of the bags. Hence, the long-run fraction of inspected bags f is

$$\begin{aligned}f&=\pi(0)+\pi(1)+\frac{1}{k}\pi(2)=1-\pi(2)+\frac{1}{k}\pi(2)=1-\frac{(k-1)(1-p)^2}{1+(k-1)(1-p)^2}\\ &=\frac{1}{1+(k-1)(1-p)^2}.\end{aligned}$$

(d) First we look at $1-f$, the fraction of bags which are not inspected. Out of the bags which are not inspected, the fraction of bags which don't taste amazing is p. Hence,

$$p(1-f)=\frac{(k-1)p(1-p)^2}{1+(k-1)(1-p)^2}.$$

The above quantity is a useful measurement of the quality of goods sold. Think of k and m as fixed numbers (which depend on the amount of resources available to the manufacturer). Then, maximizing the above quantity over all $p\in[0,1]$ yields the worst-possible case, i.e. the largest possible fraction of bags of coffee which do not taste amazing, and hence we can provide a customer guarantee without ever knowing the true value of p.

Bonus: Although you were not asked to solve the problem for general m, here is the full solution in case you are curious about how it is solved. First, we find the stationary distribution.

From $\pi(i)=(1-p)\pi(i-1)$, $i=1,\dots,m-1$, we see that $\pi(i)=(1-p)^i\pi(0)$. From the last equation, we see that $(p/k)\pi(m)=(1-p)\pi(m-1)$, or

$$\pi(m)=\frac{k}{p}(1-p)\pi(m-1)=\frac{k}{p}(1-p)^m\pi(0).$$

We have expressed everything in terms of $\pi(0)$, so we don't need the first balance equation (remember that one of the balance equations is always redundant). Instead, we will use the normalization condition $\sum_{i=0}^{m}\pi(i) = 1$, or

$$\pi(0)\left(\sum_{i=0}^{m-1}(1-p)^i + \frac{k}{p}(1-p)^m\right) = 1.$$

We recall the formula for a finite geometric series to solve for $\pi(0)$:

$$\pi(0) = \left(\frac{1-(1-p)^m}{p} + \frac{k}{p}(1-p)^m\right)^{-1} = \frac{p}{1+(k-1)(1-p)^m}.$$

Hence, the stationary distribution is

$$\begin{aligned}\pi(i) &= \frac{p(1-p)^i}{1+(k-1)(1-p)^m}, \qquad i = 0,\dots,m-1,\\ \pi(m) &= \frac{k(1-p)^m}{1+(k-1)(1-p)^m}.\end{aligned}$$

When we are in states $0,\dots,m-1$, we inspect each bag. In state m, we only inspect a fraction $1/k$ of the bags. Hence, the long-run fraction of inspected bags f is

$$\begin{aligned}f &= \sum_{i=0}^{m-1}\pi(i) + \frac{1}{k}\pi(m) = 1-\pi(m) + \frac{1}{k}\pi(m) = 1 - \frac{(k-1)(1-p)^m}{1+(k-1)(1-p)^m}\\ &= \frac{1}{1+(k-1)(1-p)^m}.\end{aligned}$$

We look at $1-f$, the fraction of bags which are not inspected. Out of the bags which are not inspected, the fraction of bags which don't taste amazing is p. Hence, the long-run fraction of bags of coffee which are sold and don't taste amazing is

$$p(1-f) = \frac{(k-1)p(1-p)^m}{1+(k-1)(1-p)^m}.$$

6 Continuous LLSE

Suppose that X and Y are uniformly distributed on the shaded region in Figure 3.

That is, X and Y have the joint distribution:

$$f_{X,Y}(x,y) = \begin{cases} 1/2, & 0 \le x \le 1, 0 \le y \le 1 \\ 1/2, & 1 \le x \le 2, 1 \le y \le 2 \end{cases}$$

(a) Do you expect X and Y to be positively correlated, negatively correlated, or neither?

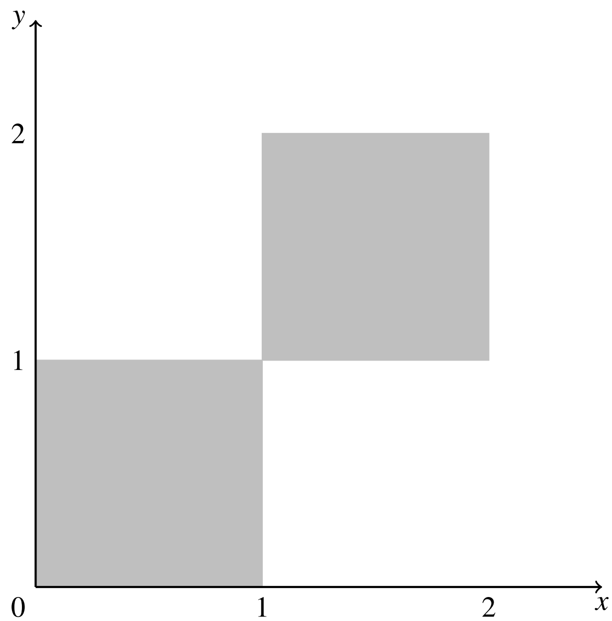


Figure 3: The joint density of (X,Y) is uniform over the shaded region.

(b) Compute the marginal distribution of X.

(c) Compute $L[Y \mid X]$.

(d) What is $\mathbf{E}[Y \mid X]$?

Solution:

(a) Positively correlated, because high values of Y correspond to high values of X.

(b) Intuitively, if we slice the joint distribution at any $x \in [0,2]$, then the probability is the same, so we should expect X to be uniformly distributed on $[0,2]$. We verify this by explicit computation:

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y = 1\{0 \le x \le 1\} \int_0^1 \frac{1}{2}\,\mathrm{d}y + 1\{1 \le x \le 2\} \int_1^2 \frac{1}{2}\,\mathrm{d}y \\
&= \frac{1}{2} 1\{0 \le x \le 2\}
\end{aligned}
$$

(c) $\mathbf{E}[X] = \mathbf{E}[Y] = 1$ by symmetry. Since X is uniform on $[0,2]$, $\operatorname{var}(X) = 4 \cdot 1/12 = 1/3$ (since the variance of a $U[0,1]$ random variable is $1/12$). We compute the covariance:

$$
\begin{aligned}
\mathbf{E}[XY] &= \int_{-\infty}^{-\infty} \int_{-\infty}^{\infty} xy f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \int_0^1 xy \cdot \frac{1}{2}\,\mathrm{d}x\,\mathrm{d}y + \int_1^2 \int_1^2 xy \cdot \frac{1}{2}\,\mathrm{d}x\,\mathrm{d}y \\
&= \frac{1}{2} \left(\int_0^1 x\,\mathrm{d}x \int_0^1 y\,\mathrm{d}y + \int_1^2 x\,\mathrm{d}y \int_1^2 y\,\mathrm{d}y \right) = \frac{1}{2} \left(\frac{1}{4} + \frac{9}{4} \right) = \frac{5}{4}
\end{aligned}
$$

So $\text{cov}(X,Y) = 5/4 - 1 \cdot 1 = 1/4$. The LLSE is

$$L[Y \mid X] - 1 = \frac{1/4}{1/3}(X-1)$$
$$L[Y \mid X] = \frac{3}{4}X + \frac{1}{4}$$

(d) The easiest way to solve this is to look at the picture of the joint density, and draw horizontal lines through middles of each of the two squares. Intuitively, $\mathbf{E}[Y \mid X]$ means "for each slice of $X = x$, what is the best guess of Y"? Slightly more formally, one can argue that conditioned on $X = x$ for $0 < x < 1$, $Y \sim U[0,1]$, so $\mathbf{E}[Y \mid X = x] = 1/2$ in this region. Conditioned on $X = x$ for $1 < x < 2$, $Y \sim U[1,2]$, so $\mathbf{E}[Y \mid X = x] = 3/2$ in this region. See Figure 4.

$$\mathbf{E}[Y \mid X = x] = \begin{cases} 1/2, & 0 \le x \le 1 \\ 3/2, & 1 \le x \le 2 \end{cases}$$

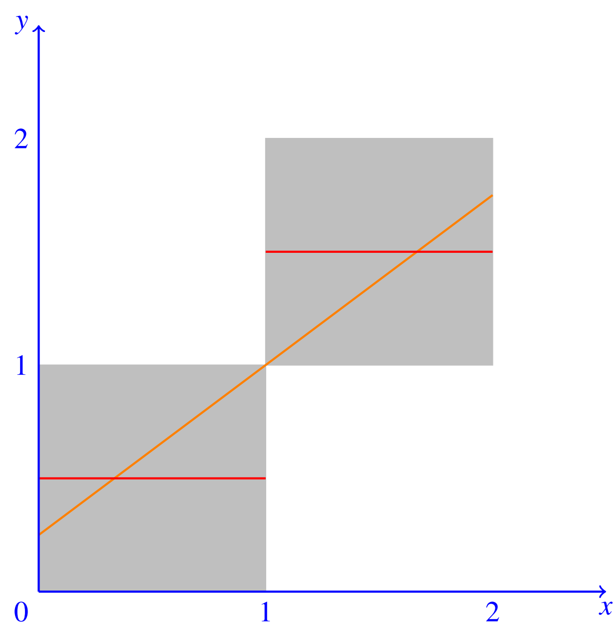


Figure 4: The LLSE is the orange line. The MMSE is the red function.

7 First Exponential to Die

Let X and Y be $\text{Expo}(\lambda_1)$ and $\text{Expo}(\lambda_2)$ respectively, independent. What is $\Pr(\min(X,Y) = X)$, the probability that the first of the two to die is X? [*Hint*: Consider the ideas in Midterm 2, Problem 5, Part 11. Note that in continuous probability, integrals are analagous to sums in discrete probabilty.]

Solution:

One has

$$\Pr(\min(X,Y)=X)=\Pr(Y>X)=\int_0^\infty \Pr(Y>X \mid X=x) f_X(x)\,\mathrm{d}x=\int_0^\infty e^{-\lambda_2 x}\cdot\lambda_1 e^{-\lambda_1 x}\,\mathrm{d}x$$
$$=-\frac{\lambda_1}{\lambda_1+\lambda_2}\,e^{-(\lambda_1+\lambda_2)x}\Big|_{x=0}^{\infty}=\frac{\lambda_1}{\lambda_1+\lambda_2}.$$