Solutions

Please simplify your answers to the extent reasonable without a calculator, show your work, and explain your answers, concisely. If you set up an integral or a sum that you cannot evaluate, leave it as it is; and if the result is needed for the next part, say how you would use the result if you had it.

1. Let \( Y_1, Y_2 \in \{0, 1\} \) be Bernoulli random variables. Suppose

\[
\begin{align*}
\Pr(Y_2 = 0 \mid Y_1 = 0) &= r \\
\Pr(Y_2 = 1 \mid Y_1 = 1) &= s
\end{align*}
\]

and \( \Pr(Y_1 = 1) = p \).

a. [4 points] What is the joint probability function for \( Y_1 \) and \( Y_2 \)?

\[
\begin{align*}
\Pr(Y_1 = 0 \text{ and } Y_2 = 0) &= \Pr(Y_2 = 0 \mid Y_1 = 0) \Pr(Y_1 = 0) = r(1 - p) \\
\Pr(Y_1 = 0 \text{ and } Y_2 = 1) &= \Pr(Y_2 = 1 \mid Y_1 = 0) \Pr(Y_1 = 0) = (1 - r)(1 - p) \\
\Pr(Y_1 = 1 \text{ and } Y_2 = 0) &= \Pr(Y_2 = 0 \mid Y_1 = 1) \Pr(Y_1 = 1) = (1 - s)p \\
\Pr(Y_1 = 1 \text{ and } Y_2 = 1) &= \Pr(Y_2 = 1 \mid Y_1 = 1) \Pr(Y_1 = 1) = s(1 - p)
\end{align*}
\]

With the values of \( Y_1 \) and \( Y_2 \) labeling the rows and columns of a matrix, respectively, this joint probability function is:

\[
\begin{pmatrix}
0 & 1 \\
0 & \begin{pmatrix} r(1 - p) & (1 - r)(1 - p) \\
(1 - s)p & ps \end{pmatrix}
\end{pmatrix}
\]

b. [4 points] For what values of \( r \), \( s \) and \( p \) are \( Y_1 \) and \( Y_2 \) independent?

To be independent, the joint probability matrix above must be singular, \( i.e. \), have nonzero determinant because one row is a multiple of the other since the distribution is a product distribution. Its determinant is \( rsp(1 - p) - (1 - r)(1 - s)p(1 - p) = (r + s - 1)p(1 - p) \) which vanishes if \( p \in \{0, 1\} \) or if \( r + s = 1 \).

c. [4 points] What is \( \Pr(Y_2 = 1) \)?

\[
\Pr(Y_2 = 1) = (1 - r)(1 - p) + sp.
\]

d. [4 points] For what values of \( r \), \( s \) and \( p \) is \( \Pr(Y_2 = 1) = \Pr(Y_1 = 1) \)?

Solving \( p = (1 - r)(1 - p) + sp \) for \( p \) gives

\[
p = \frac{1 - r}{2 - r - s},
\]

as long as \( r + s \neq 2 \).

2. Consider a sequence of customers entering a store. Let \( Y_i \in \{0, 1\} \) denote the number of items the \( i^{th} \) customer buys, for \( 0 < i \in \mathbb{N} \). Suppose every customer, after the first, sees what the previous customer does; if the previous customer bought nothing, the
current customer also buys nothing, with probability $r$, and if the previous customer bought an item, the current customer does too, with probability $s$.

a. [4 points] Suppose $0 < r = s < 1$. Without doing any calculation, explain what is the fraction of customers who buy an item in the infinite number of customers limit.

In this case buying and not-buying are completely symmetrical, so in the infinite number of customer limit, half of the customers buy an item.

b. [4 points] For general $0 < r, s < 1$, what is $\lim_{n \to \infty} \Pr(Y_n = 1)$?

The transition probability matrix is

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r & 1 - r \\ 1 - s & s \end{pmatrix},$$

and we want to find $\pi = (1 - \pi_1 \quad \pi_1)$ that solves $\pi P = \pi$. Multiplying out gives us the equation

$$r(1 - \pi_1) + (1 - s)\pi_1 = 1 - \pi_1,$$

which we solve to get

$$\lim_{n \to \infty} \Pr(Y_n = 1) = \pi_1 = \frac{1 - r}{2 - r - s}.$$

(Notice that this is the same as the answer to problem 1.d, i.e., that the probability of each customer buying an item is the same.)

c. [8 points] Still for general $0 < r, s < 1$, what is $\lim_{n \to \infty} \Pr(Y_n = 1 \text{ and } Y_{n+1} = 1)$?

The probability transition matrix for the states of consecutive customers is

$$P = \begin{pmatrix} 00 & 01 & 10 & 11 \\ 00 & r & 1 - r & 0 & 0 \\ 01 & 0 & 0 & 1 - s & s \\ 10 & r & 1 - r & 0 & 0 \\ 11 & 0 & 0 & 1 - s & s \end{pmatrix}.$$

Solving $\pi = (\pi_0 \quad \pi_1 \quad \pi_2 \quad \pi_3) = \pi P$ requires solving:

$$r\pi_0 + r\pi_2 = \pi_0$$

$$(1 - r)\pi_0 + (1 - r)\pi_2 = \pi_1$$

$$(1 - s)\pi_1 + (1 - s)\pi_3 = \pi_2$$

$$s\pi_1 + s\pi_3 = \pi_2$$

The first two of these equations imply $\pi_0 = (r/(1 - r))\pi_1$ and the last two imply $\pi_3 = (s/(1 - s))\pi_2$. Plugging these back in to the first and last equations gives $\pi_1 = \pi_2$. Requiring that the sum of the $\pi_i$ be 1 gives

$$\pi_1 = \frac{(1 - r)(1 - s)}{2 - r - s} = \pi_2.$$
which implies that
\[
\lim_{n \to \infty} \Pr(Y_n = 1 \text{ and } Y_{n+1} = 1) = \pi_3 = \frac{s(1 - r)}{2 - r - s}.
\]

An alternate, cleverer, solution is to notice that
\[
\lim_{n \to \infty} \Pr(Y_n = 1 \text{ and } Y_{n+1} = 1) = \lim_{n \to \infty} \Pr(Y_{n+1} = 1 \mid Y_n = 1) \Pr(Y_n = 1)
\]
\[
= s \lim_{n \to \infty} \Pr(Y_n = 1)
\]
\[
= s \cdot \frac{1 - r}{2 - r - s},
\]
where we have used the answer to 2.b for the last step.

3. [16 points] Now suppose people enter a store according to a Poisson process with rate \(\lambda\). The store owner wants to know if someone buying something makes the next person more likely to buy something (as it did in problem 2 for \(s > 1/2\)). To help answer this question, suppose each person who enters the store buys something with probability \(p\), independently of what anyone else does. Let \(C(t)\) be the number of people who buy something and are followed by the next customer also buying something, both before time \(t\). What is \(E[C(t)]\)?

\[
E[C(t)] = \sum_{n=2}^{\infty} E[C(t) \mid X(t) = n] \Pr(X(t) = n)
\]
\[
= \sum_{n=2}^{\infty} \sum_{i=1}^{n-1} E[Y_i = 1 \text{ and } Y_{i+1} = 1] \Pr(X(t) = n)
\]
\[
= \sum_{n=2}^{\infty} (n - 1) p^2 \frac{\lambda^n}{n!} e^{-\lambda t}
\]
\[
= p^2 e^{-\lambda t} \left( \lambda t \sum_{n=2}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} - \sum_{n=2}^{\infty} \frac{\lambda^n}{n!} \right)
\]
\[
= p^2 e^{-\lambda t} \left( \lambda t (e^{\lambda t} - 1) - (e^{\lambda t} - 1 - \lambda t) \right)
\]
\[
= p^2 (e^{-\lambda t} - 1 + \lambda t).
\]

Extra Credit. [5 points] Suppose instead of the people’s purchasing decisions being independent, they are dependent as in problem 2. Now what is \(E[C(t)]\)?

Since we don’t know what the first customer does, or even what the probability distribution for the first customer’s action is, we have to make some assumption. If we assume that customer probabilities are stationary, the only change in the calculation above is that the \(p^2\) is replaced by
\[
\lim_{n \to \infty} \Pr(Y_n = 1 \text{ and } Y_{n+1} = 1) = \frac{s(1 - r)}{2 - r - s}
\]
4. Let \( \{X(t) \mid t \in \mathbb{R}_{\geq 0}\} \) be a Poisson point process with rate \( \lambda \) on \( \mathbb{R}_{\geq 0} \). For each point \( 0 < i \in \mathbb{N} \) of the process, let \( D_i \) be the distance to its nearest neighbor.

a. [3 points] Write \( D_i \) in terms of the sojourn times \( \{S_j \mid j \in \mathbb{N}\} \).

\[
D_1 = S_1 \quad \text{and for } i > 1, \quad D_i = \min\{S_{i-1}, S_i\}.
\]

b. [3 points] Are the \( \{D_i \mid 0 < i \in \mathbb{N}\} \) independent?

No, since \( D_1 \) and \( D_2 \), for example, both depend upon \( S_1 \). For example, let \( 0 \leq s < t \);

\[
\Pr(D_1 > s, D_2 > t) = \Pr(S_1 > s, S_1 > t, S_2 > t) = \Pr(S_1 > t) \Pr(S_2 > t),
\]

while

\[
\Pr(D_1 > s) \Pr(D_2 > t) = \Pr(S_1 > s) \Pr(S_1 > t, S_2 > t) = \Pr(S_1 > s) \Pr(S_1 > t) \Pr(S_2 > t).
\]

c. [10 points] What is the probability density function for each \( D_i \)?

\[
f_{D_1}(d) = f_{S_1}(d) = \lambda e^{-\lambda d}. \quad \text{For } i > 1,
\]

\[
\Pr(d < D_i \leq d + \Delta d) = \Pr(d < \text{distance to nearest point} \leq d + \Delta d)
\]

\[
= \Pr(d < \text{distance to point on left} \leq d + \Delta d \text{ and } d + \Delta d < \text{distance to point on right})
\]

\[
+ \Pr(d < \text{distance to point on right} \leq d + \Delta d \text{ and } d + \Delta d < \text{distance to point on left})
\]

\[
= 2e^{-\lambda d} \cdot \lambda \Delta d e^{-\lambda \Delta d} \cdot e^{-\lambda d},
\]

which implies \( f_{D_i}(d) = 2\lambda e^{-2\lambda d} \), i.e., just the probability distribution for a sojourn time with double the rate. Alternatively,

\[
\Pr(D_i > d) = \Pr(S_{i-1} > d, S_i > d) = \Pr(S_{i-1} > d) \Pr(S_i > d) = e^{-\lambda d} e^{-\lambda d} = e^{-2\lambda d}.
\]

From this we can compute the probability density function:

\[
f_{D_i}(d) = \frac{d}{dd} \Pr(D_i > d) = -\frac{d}{dd} e^{-2\lambda d} = 2\lambda e^{-2\lambda d},
\]

for \( d \geq 0 \).

5. [20 points] An ant starts at one vertex of a cube and at each time step walks along an edge to an adjacent vertex, choosing each possible edge with equal probability \( 1/3 \).
What is the probability the ant returns to its original vertex before reaching the opposite vertex?

Let the state of the ant be described by how many edges away from the corner at which it starts it is, \textit{e.g.}, the opposite corner is state 3. Making states 0 and 3 absorbing states, the probability transition matrix is

\[
P = \begin{pmatrix}
0 & 1 & 2 & 3 \\
0 & 1 & 0 & 0 \\
1 & 0 & 2/3 & 0 \\
2 & 0 & 0 & 1/3 \\
3 & 0 & 0 & 0 & 1 \\
\end{pmatrix}.
\]

Since after the first step the ant is at state 1, we must compute \( \nu_i = \Pr(\text{absorbed at state 0} \mid \text{from state } i) \) for \( i = 1 \). Considering the next step the ant takes, we get

\[
\nu_1 = \frac{1}{3} + \frac{2}{3} \nu_2,
\]

\[
\nu_2 = \frac{2}{3} \nu_1.
\]

Solving these equations gives \( \nu_1 = 3/5 \).

6. Let \( X, Y \) be random variables with a bivariate normal distribution such that \( \mathbb{E}[Y] = 0 \) and \( \text{Var}[X] = 1 = \text{Var}[Y] \). Suppose \( \mathbb{E}[X \mid Y = y] = 2 - y/3 \).

a. [6 points] What is \( \mathbb{E}[X] \)?

\[
\mathbb{E}[X] = \int_{-\infty}^{\infty} \mathbb{E}[X \mid Y = y] f_Y(y) \, dy = \int_{-\infty}^{\infty} (2 - y/3) f_Y(y) \, dy = \mathbb{E}[2 - Y/3] = 2 - \mathbb{E}[Y]/3 = 2.
\]

b. [10 points] What is \( \text{Cov}[X, Y] \)?

\[
\text{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X] \mathbb{E}[Y] = \mathbb{E}[XY]
\]

\[
= \int_{-\infty}^{\infty} \mathbb{E}[XY \mid Y = y] f_Y(y) \, dy
\]

\[
= \int_{-\infty}^{\infty} y \mathbb{E}[X \mid Y = y] f_Y(y) \, dy
\]

\[
= \int_{-\infty}^{\infty} (2y - y^2/3) f_Y(y) \, dy
\]

\[
= \mathbb{E}[2Y - Y^2/3] = 2\mathbb{E}[Y] - \mathbb{E}[Y^2]/3 = -1/3,
\]

because \( \mathbb{E}[Y] = 0 \) and thus \( \mathbb{E}[Y^2] = \text{Var}[Y] = 1 \). Notice that we did not need to use the fact that the joint distribution of \( X \) and \( Y \) is bivariate normal.