CS 70 Discrete Mathematics and Probability Theory
Spring 2022 Satish Rao and Koushik Sen **HW 14**

1 Darts with Friends

Michelle and Alex are playing darts. Being the better player, Michelle's aim follows a uniform distribution over a disk of radius 1 around the center. Alex's aim follows a uniform distribution over a disk of radius 2 around the center.

(a) Let the distance of Michelle's throw from the center be denoted by the random variable X and let the distance of Alex's throw from the center be denoted by the random variable Y.

- What's the cumulative distribution function of X?
- What's the cumulative distribution function of Y?
- What's the probability density function of X?
- What's the probability density function of Y?

(b) What's the probability that Michelle's throw is closer to the center than Alex's throw? What's the probability that Alex's throw is closer to the center?

(c) What's the cumulative distribution function of $U = \max\{X,Y\}$?

(d) What's the cumulative distribution function of $V = \min\{X,Y\}$?

(e) What is the expectation of the absolute difference between Michelle's and Alex's distances from the center, that is, what is $\mathbb{E}[|X-Y|]$? [*Hint*: Use parts (c) and (d), together with the continuous version of the tail sum formula, which states that $\mathbb{E}[Z] = \int_0^\infty \mathbb{P}[Z \geq z]dz$.]

Solution:

(a)
- To get the cumulative distribution function of X, we'll consider the ratio of the area where the distance to the center is less than x, compared to the entire available area. This gives us the following expression:

$$\mathbb{P}[X \leq x] = \frac{\pi x^2}{\pi} = x^2, \qquad x \in [0,1].$$

- Using the same approach as the previous part:

$$\mathbb{P}[Y \leq y] = \frac{\pi y^2}{\pi \cdot 4} = \frac{y^2}{4}, \qquad y \in [0,2].$$

- We'll take the derivative of the CDF to get the following:

$$f_X(x) = \frac{\mathrm{d}\mathbb{P}[X \le x]}{\mathrm{d}x} = 2x, \qquad x \in [0,1].$$

- Using the same approach as the previous part:

$$f_Y(y) = \frac{\mathrm{d}\mathbb{P}[Y \le y]}{\mathrm{d}y} = \frac{y}{2}, \qquad y \in [0,2].$$

(b) We'll condition on Alex's outcome and then integrate over all the possibilities to get the marginal $\mathbb{P}(X \le Y)$ as following:

$$\mathbb{P}[X \le Y] = \int_0^2 \mathbb{P}[X \le Y \mid Y = y] f_Y(y)\,\mathrm{d}y = \int_0^1 y^2 \times \frac{y}{2}\,\mathrm{d}y + \int_1^2 1 \times \frac{y}{2}\,\mathrm{d}y$$
$$= \frac{1}{8} + \frac{3}{4} = \frac{7}{8}.$$

Note the range within which $\mathbb{P}[X \le Y] = 1$. This allowed us to separate the integral to simplify our solution. Using this, we can get $\mathbb{P}[Y \le X]$ by the following:

$$\mathbb{P}[Y \le X] = 1 - \mathbb{P}[X \le Y] = \frac{1}{8}$$

A similar approach to the integral above could be used to verify this result:

$$\mathbb{P}[Y \le X] = \int_0^1 \mathbb{P}[Y \le X \mid X = x] f_X(x)\,\mathrm{d}x = \int_0^1 \frac{x^2}{4} 2x\,\mathrm{d}x = \frac{1}{2}\int_0^1 x^3\,\mathrm{d}x = \frac{1}{8}.$$

(c) Getting the CDF of U relies on the insight that for the maximum of two random variables to be smaller than a value, they both need to be smaller than that value. Using this we can get the following result for $u \in [0,1]$:

$$\mathbb{P}[U \le u] = \mathbb{P}[X \le u]\mathbb{P}[Y \le u] = \left(u^2\right)\left(\frac{u^2}{4}\right) = \frac{u^4}{4}.$$

For $u \in [1,2]$ we have $\mathbb{P}[X \le u] = 1$; this makes

$$\mathbb{P}[U \le u] = \mathbb{P}[Y \le u] = \frac{u^2}{4}.$$

For $u > 2$ we have $\mathbb{P}[U \le u] = 1$ since CDFs of both X and Y are 1 in this range.

(d) Getting the CDF of V relies on a similar insight that for the minimum of two random variables to be greater than a value, they both need to be greater than that value. Taking the complement of this will give us the CDF of V. This allows us to get the following result. For $v \in [0,1]$:

$$\mathbb{P}[V \le v] = 1 - \mathbb{P}[V \ge v] = 1 - \mathbb{P}[X \ge v]\mathbb{P}[Y \ge v] = 1 - \left(1 - \mathbb{P}[X \le v]\right)\left(1 - \mathbb{P}[Y \le v]\right)$$
$$= 1 - \left(1 - v^2\right)\left(1 - \frac{v^2}{4}\right) = \frac{5v^2}{4} - \frac{v^4}{4}.$$

For $v > 1$, we get $\mathbb{P}[X > v] = 0$, making $\mathbb{P}[V \le v] = 1$.

(e) We can subtract V from U to get this difference. Using the tail-sum formula to calculate the expectation, we can get the following result:

$$\begin{aligned}\mathbb{E}[|X-Y|] = \mathbb{E}[U-V] = \mathbb{E}[U]-\mathbb{E}[V] &= \int_0^2 \mathbb{P}[U \geq u]\,\mathrm{d}u - \int_0^1 \mathbb{P}[V \geq v]\,\mathrm{d}v \\ &= \int_0^1 \left(1-\frac{u^4}{4}\right)\mathrm{d}u + \int_1^2 \left(1-\frac{u^2}{4}\right)\mathrm{d}u - \int_0^1 \left(1-\frac{5v^2}{4}+\frac{v^4}{4}\right)\mathrm{d}v \\ &= \frac{19}{20}+\frac{5}{12}-\frac{19}{30} = \frac{11}{15}.\end{aligned}$$

Alternatively, you could derive the density of U and V and use those to calculate the expectation. For $u \in [0,1]$:

$$f_U(u) = \frac{\mathrm{d}\mathbb{P}[U \leq u]}{\mathrm{d}u} = u^3.$$

For $u \in [1,2]$:

$$f_U(u) = \frac{\mathrm{d}\mathbb{P}[U \leq u]}{\mathrm{d}u} = \frac{u}{2}.$$

Using this we can calculate $\mathbb{E}[U]$ as:

$$\mathbb{E}[U] = \int_0^2 u f_U(u)\,\mathrm{d}u = \int_0^1 u^4\,\mathrm{d}u + \frac{1}{2}\int_1^2 u^2\,\mathrm{d}u = \frac{1}{5}+\frac{7}{6} = \frac{41}{30}.$$

To calculate $\mathbb{E}[V]$ we will use the following PDF for $v \in [0,1]$:

$$f_V(v) = \frac{\mathrm{d}\mathbb{P}[V \leq v]}{\mathrm{d}v} = \frac{5v}{2} - v^3.$$

We can get the $\mathbb{E}[V]$ by the following:

$$\mathbb{E}[V] = \int_0^1 v f_V(v)\,\mathrm{d}v = \int_0^1 \left(\frac{5v^2}{2} - v^4\right)\mathrm{d}v = \frac{5}{6}-\frac{1}{5} = \frac{19}{30}.$$

Combining the two results gives us the same result as above:

$$\mathbb{E}[|X-Y|] = \mathbb{E}[U-V] = \mathbb{E}[U]-\mathbb{E}[V] = \frac{41}{30}-\frac{19}{30} = \frac{11}{15}.$$

2 Continuous LLSE

Suppose that X and Y are uniformly distributed on the shaded region in the figure below.

That is, X and Y have the joint distribution:

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

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

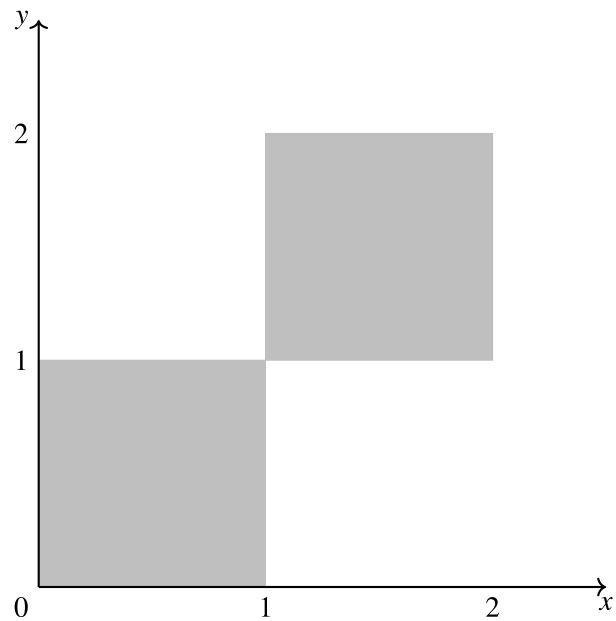

Figure 1: The joint density of (X,Y) is uniform over the shaded region.

(b) Compute the marginal distribution of X.

(c) Compute $L[Y \mid X]$, the best linear estimator of Y given X.

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

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:

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

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

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

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, $\mathbb{E}[Y \mid X]$ 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 $\mathbb{E}[Y \mid X = x] = 1/2$ in this region. Conditioned on $X = x$ for $1 < x < 2$, $Y \sim U[1,2]$, so $\mathbb{E}[Y \mid X = x] = 3/2$ in this region. See Figure **??**.

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

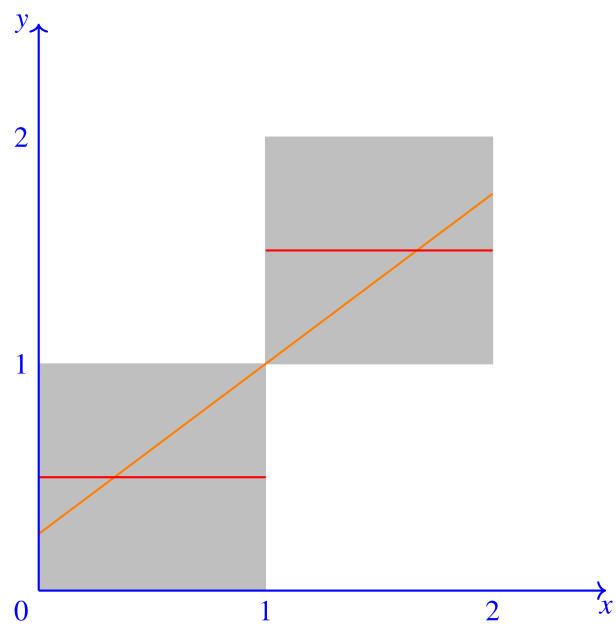

Figure 2: $L[Y \mid X]$ is the orange line. $\mathbb{E}[Y \mid X]$ is the red function.

3 Chebyshev's Inequality vs. Central Limit Theorem

Let n be a positive integer. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with the following distribution:

$$\mathbb{P}[X_i = -1] = \frac{1}{12}; \qquad \mathbb{P}[X_i = 1] = \frac{9}{12}; \qquad \mathbb{P}[X_i = 2] = \frac{2}{12}.$$

(a) Calculate the expectations and variances of X_1, $\sum_{i=1}^n X_i$, $\sum_{i=1}^n (X_i - \mathbb{E}[X_i])$, and

$$Z_n = \frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{\sqrt{n/2}}.$$

(b) Use Chebyshev's Inequality to find an upper bound b for $\mathbb{P}[|Z_n| \geq 2]$.

(c) Can you use b to bound $\mathbb{P}[Z_n \geq 2]$ and $\mathbb{P}[Z_n \leq -2]$?

(d) As $n \to \infty$, what is the distribution of Z_n?

(e) We know that if $Z \sim \mathcal{N}(0,1)$, then $\mathbb{P}[|Z| \leq 2] = \Phi(2) - \Phi(-2) \approx 0.9545$. As $n \to \infty$, can you provide approximations for $\mathbb{P}[Z_n \geq 2]$ and $\mathbb{P}[Z_n \leq -2]$?

Solution:

(a) $\mathbb{E}[X_1] = -1/12 + 9/12 + 4/12 = 1$, and

$$\mathrm{Var}(X_1) = \frac{1}{12} \cdot 2^2 + \frac{9}{12} \cdot 0^2 + \frac{2}{12} \cdot 1^2 = \frac{1}{2}.$$

Using linearity of expectation and variance (since $X_1, \ldots, X_n$ are independent), we find that $\mathbb{E}[\sum_{i=1}^n X_i] = n$ and $\mathrm{Var}(\sum_{i=1}^n X_i) = n/2$.

Again, by linearity of expectation, $\mathbb{E}[\sum_{i=1}^n X_i - n] = n - n = 0$. Subtracting a constant does not change the variance, so $\mathrm{Var}(\sum_{i=1}^n X_i - n) = n/2$, as before.

Using the scaling properties of the expectation and variance, $\mathbb{E}[Z_n] = 0/\sqrt{n/2} = 0$ and $\mathrm{Var}(Z_n) = (n/2)/(n/2) = 1$.

(b)

$$\mathbb{P}[|Z_n| \geq 2] \leq \frac{\mathrm{Var}(Z_n)}{2^2} = \frac{1}{4}$$

(c) $1/4$ for both, since $\mathbb{P}[Z_n \geq 2] \leq \mathbb{P}[|Z_n| \geq 2]$ and $\mathbb{P}[Z_n \leq -2] \leq \mathbb{P}[|Z_n| \geq 2]$.

(d) By the Central Limit Theorem, we know that $Z_n \to \mathcal{N}(0,1)$, the standard normal distribution.

(e) Since $Z_n \to \mathcal{N}(0,1)$, we can approximate $\mathbb{P}[|Z_n| \geq 2] \approx 1 - 0.9545 = 0.0455$. By the symmetry of the normal distribution, $\mathbb{P}[Z_n \geq 2] = \mathbb{P}[Z_n \leq -2] \approx 0.0455/2 = 0.02275$.

It is interesting to note that the CLT provides a much smaller answer than Chebyshev. This is due to the fact that the CLT is applied to a particular kind of random variable, namely the (scaled) sum of a bunch of random variables. Chebyshev's inequality, however, holds for any random variable, and is therefore weaker.

4 Playing Blackjack

You are playing a game of Blackjack where you start with \$100. You are a particularly risk-loving player who does not believe in leaving the table until you either make \$400, or lose all your money. At each turn you either win \$100 with probability p, or you lose \$100 with probability $1 - p$.

(a) Formulate this problem as a Markov chain; i.e. define your state space, transition probabilities, and determine your starting state.

(b) Compute the probability that you end the game with \$400.

Solution:

(a) Since it is only possible for us to either win or lose \$100, we define the following state space $\chi = \{0, 100, 200, 300, 400\}$. The following are the transition probabilities:

$$\mathbb{P}[X_j = 0 | X_{j-1} = 0] = \mathbb{P}[X_j = 400 | X_{j-1} = 400] = 1$$
$$\mathbb{P}[X_j = i + 100 | X_{j-1} = i] = p \text{ for } i \in \{100, 200, 300\}$$
$$\mathbb{P}[X_j = i - 100 | X_{j-1} = i] = 1 - p \text{ for } i \in \{100, 200, 300\}$$

(b) We want to find the probability that we are "absorbed" by state 400 before we are absorbed by state 0. We can calculate this probability by leveraging the memoryless property of Markov Chains. Define a_i as the probability of reaching state 400 before 0 starting at state i.

We also know that for $i \in \{100, 200, 300\}$, we have the following relation:

$$a_i = (1-p)a_{i-100} + pa_{i+100} \text{ for } i \in \{100, 200, 300\}$$

We also know that $a_0 = 0$, since if you are at state 0, then there is no chance that you end up at state 400. We also have $a_{400} = 1$ since if we are at state 400, then we have already succeeded in our goal to reach 400.

We have three unknowns $(a_{100}, a_{200}, a_{300})$ and three equations, and we can now solve this system of equations for a_{100}.

$$a_0 = 0, a_{400} = 1$$
$$\implies a_i = (1-p)a_{i-100} + pa_{i+100} \text{ for } i \in \{100, 200, 300\}$$
$$a_{100} = pa_{200}$$
$$a_{200} = (1-p)a_{100} + pa_{300} \implies a_{200}[1 - p(1-p)] = pa_{300}$$
$$\implies a_{200} = \frac{pa_{300}}{1 - p(1-p)}$$
$$a_{300} = (1-p)a_{200} + p \implies a_{300} = \frac{(1-p)pa_{300}}{1 - p(1-p)} + p$$
$$\implies a_{300} = \frac{p(1 - p(1-p))}{1 - 2p(1-p)}$$
$$\implies a_{200} = \frac{p^2}{1 - 2p(1-p)}$$
$$\implies a_{100} = \frac{p^3}{1 - 2p(1-p)}$$

This problem is called Gambler's Ruin, where it is used to show that even if p is decently large, after playing a large number of games without stopping, you will end up at 0 dollars with high probability.

5 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:

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

We can write this in terms of the differences $D_i := R_{i+1} - R_i$. If we take the recurrence relation $R_i = 1 + \frac{1}{2}R_{i-1} + \frac{1}{2}R_{i+1}$, we can rearrange the equation:

$$\begin{aligned}
R_i &= 1 + \frac{1}{2}R_{i-1} + \frac{1}{2}R_{i+1} \\
2R_i &= 2 + R_{i-1} + R_{i+1} \\
R_i - R_{i-1} - 2 &= R_{i+1} - R_i \\
D_{i-1} - 2 &= D_i
\end{aligned}$$

Furthermore, we know that $D_0 := R_1 - R_0 = -1$ from the very first hitting time equation. Since we have shown that D_i decreases by 2 every time, we know that $D_i = -2i - 1$. How do we get back R_i from knowing D_i? Well, we see that

$$R_i = (R_i - R_{i-1}) + (R_{i-1} - R_{i-2}) + \cdots + (R_1 - R_0) + R_0 = D_{i-1} + D_{i-2} + \cdots + D_0 + R_0$$

Therefore, we have $R_i = -1 - 3 - 5 - \cdots - (2i-1) + R_0$. What is the sum of the first i odd integers? Here is how you would derive it. Let $S = 1 + 3 + 5 + \cdots + (2i-1)$. Then, we can also write S backwards, as $S = (2i-1) + (2i-3) + \cdots + 5 + 3 + 1$. Lining up the terms, we see:

$$\begin{aligned} S &= 1 \qquad\quad + 3 \qquad\quad + \cdots + (2i-3) + (2i-1) \\ S &= (2i-1) + (2i-3) + \cdots + 3 \qquad\quad + 1 \end{aligned}$$

Adding these together gives us $2S = 2i + 2i + 2i + \cdots + 2i = 2i^2$. Solving for S yields $S = i^2$.

Now that we know this fact, we see that $R_i = R_0 - i^2$. Since we know that $R_n = 0$, we see that $R_0 - n^2 = 0$, and thus $R_0 = n^2$.

6 Boba in a Straw

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

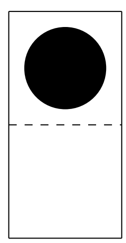

Figure 3: A straw with one boba currently inside. The straw only has enough room to fit two boba.

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 Jonathan drinks from the straw, the following happens every second:

1. The contents of the top component enter Jonathan'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 Jonathan evaluate the consequences of his incessant drinking!

(a) Draw the Markov chain that models this process, and show that it is both irreducible and aperiodic.

(b) 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.]

(c) 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.]

(d) Jonathan was annoyed by the straw so he bought a fresh new straw (same as the straw from Figure 1). What is the long-run average rate of Jonathan's calorie consumption? (Each boba is roughly 10 calories.)

(e) 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.]

(f) What is the long run probability that the amount of boba in the straw doesn't change from one second to the next?

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 (1); similarly, the second component represents whether the bottom component is empty or full. See Figure **??**. This chain is irreducible as we can get from any state to any other with the cycle $(0,0) \to (0,1) \to (1,0) \to (0,1) \to (1,1) \to (1,0) \to (0,0)$. Furthermore, this chain contains a self-loop, so it is aperiodic.

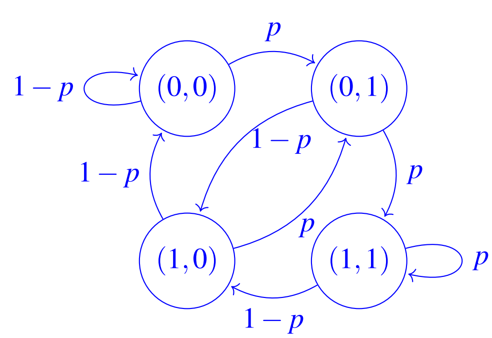

Figure 4: Transition diagram for the Markov chain.

(b) 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 $\mathbb{E}_i[\cdot] = \mathbb{E}[\cdot \mid X_0 = i]$ denote the expectation starting from state i (for convenience of notation). The hitting-time equations are

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

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 $\mathbb{E}_{(0,0)}[T] = (1+p)/p^2$.

(c) The new hitting-time equations are

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

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) $\mathbb{E}_{(0,0)}[T] = 11$.

(d) This part is actually more straightforward than it might initially seem: the average rate at which Jonathan 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.

(e) 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

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

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

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

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 $2p$.

(f) The long run probability that the amount of boba doesn't change is the probability that either (a) we are in state $(0,1)$ and transition to $(1,0)$, (b) we are in state $(1,0)$ and transition to $(0,1)$, (c) we are in state $(1,1)$ and transition to $(1,1)$, or (d) we are in state $(0,0)$ and transition to $(0,0)$. In the long run, the probability we are in a particular state is given by the stationary distribution, so we have

$$\begin{aligned}
\mathbb{P}[(0,0) \to (0,0)] &= \pi(0,0)\mathbb{P}[X_{n+1} = (0,0) \mid X_n = (0,0)] = (1-p)^3 \\
\mathbb{P}[(0,1) \to (1,0)] &= \pi(0,1)\mathbb{P}[X_{n+1} = (1,0) \mid X_n = (0,1)] = p(1-p)^2 \\
\mathbb{P}[(1,0) \to (0,1)] &= \pi(1,0)\mathbb{P}[X_{n+1} = (0,1) \mid X_n = (1,0)] = p^2(1-p) \\
\mathbb{P}[(1,1) \to (1,1)] &= \pi(1,1)\mathbb{P}[X_{n+1} = (1,1) \mid X_n = (1,1)] = p^3.
\end{aligned}$$

Thus, our overall probability is $\boxed{p^3 + p^2(1-p) + p(1-p)^2 + (1-p)^3}$.