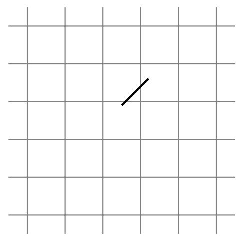
CS 70 Fall 2018 — Discrete Mathematics and Probability Theory Course Notes — HW 13

1 Buffon's Needle on a Grid

In this problem, we will consider Buffon's Needle, but with a catch. We now drop a needle at random onto a large grid, and example of which is shown below.



The length of the needle is 1, and the space between the grid lines is 1 as well.

Recall from class that a random throw means that the position of the center of the needle and its orientation are independent and uniformly distributed.

(a) Given that the angle between the needle and the horizontal lines is θ, what is the probability that the needle does not intersect any grid lines? Justify your answer.

(b) Continue part (a) to find the probability that the needle, when dropped onto the grid at random (with the angle chosen uniformly at random), intersects a grid line. Justify your answer.

Hint: You may use the fact that $\sin\theta\cos\theta = \frac{1}{2}\sin(2\theta)$ without proof.

(c) Let X be the number of times the needle intersects a gridline (so, in the example shown above, $X = 2$). Find $\mathbb{E}[X]$. Justify your answer.

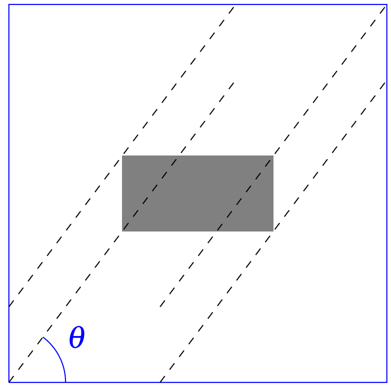
Hint: Can you do this without using your answer from part (b)?

(d) Combine the previous parts to figure out the probability that $X = 1$, i.e. the needle will only intersects one gridline. Justify your answer.

(e) We will fold the needle into an equilateral triangle, where each side is length $\frac{1}{3}$. What is the expected number of intersections that this triangle will have with the gridlines, when dropped onto the grid? Justify your answer.

Solution:

(a) Since the grid is large, we can think of the random process as picking the center of the needle uniformly at random from the unit square. We therefore see that the successful region is a rectangle with sides $1-\sin\theta$ and $1-\cos\theta$. The diagram below illustrates this. The dashed lines represent the farthest that the needle can go without touching any of the grid lines. The shaded region is where the center of the needle can go without touching any grid lines.



Therefore, the successful area is $(1-\sin\theta)(1-\cos\theta)$, which gives us the answer of

$$\boxed{(1-\sin\theta)(1-\cos\theta)}$$

(b) θ is chosen uniformly at random from 0 to $\frac{\pi}{2}$, so the pdf of θ is therefore $\frac{2}{\pi}$. Using the Law of Total Probability, we see that the probability that we do not intersect a grid line is

$$\int_0^{\frac{\pi}{2}} \frac{2(1-\sin\theta)(1-\cos\theta)}{\pi}\, d\theta.$$

Expanding out the integral yields

$$\begin{aligned}
\int_0^{\frac{\pi}{2}} \frac{2(1-\sin\theta)(1-\cos\theta)}{\pi}\, d\theta &= \frac{2}{\pi}\int_0^{\frac{\pi}{2}} 1-\sin\theta-\cos\theta+\sin\theta\cos\theta\, d\theta \\
&= \frac{2}{\pi}\int_0^{\frac{\pi}{2}} 1-\sin\theta-\cos\theta+\frac{1}{2}\sin 2\theta\, d\theta \\
&= \frac{2}{\pi}\left(\frac{\pi}{2}-1-1+\frac{1}{2}\right) \\
&= \frac{2}{\pi}\left(\frac{\pi}{2}-\frac{3}{2}\right) \\
&= 1-\frac{3}{\pi}
\end{aligned}$$

Therefore, the probability that the needle *does* intersect a gridline is

$$\boxed{\frac{3}{\pi}}$$

(c) Recall from the notes that in the original Buffon's needle problem, the probability that the needle will intersect a line is $\frac{2}{\pi}$. Let X_1 be equal to the number of times the needle will intersect a vertical grid line, and let X_2 equal the number of times the needle will intersect a horizontal grid line. Thus, $X = X_1 + X_2$.

By the linearity of expectation, we know that $\mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2]$. From the first paragraph, we know that $\mathbb{E}[X_1] = \mathbb{E}[X_2] = \frac{2}{\pi}$. Therefore, the answer is $\boxed{\dfrac{4}{\pi}}$.

(d) We know that the only possible values for X are 0, 1, and 2. Therefore, we know that

$$\mathbb{P}(X = 1) + \mathbb{P}(X = 2) = \frac{3}{\pi}.$$

We also know that the expected value is equal to $\frac{4}{\pi}$, so

$$\mathbb{P}(X = 1) + 2\mathbb{P}(X = 2) = \frac{4}{\pi}.$$

This gives us two equations and two unknowns. Solving for $\mathbb{P}(X = 1)$ yields

$$\boxed{\frac{2}{\pi}}.$$

(e) Let Y_i be the number of times the i-th side of the triangle intersects a grid line, for $i = 1, 2, 3$. Thus, the total number of times this triangle intersects a grid line is $Y_1 + Y_2 + Y_3$.

Let us revisit part (c), however. Let Z_1 be the number of times the first $\frac{1}{3}$ of the needle intersects a grid line, and let Z_2 be the number of times the second $\frac{1}{3}$ of the needle intersects the grid line, and similarly for Z_3. We know that

$$\mathbb{E}[Z_1 + Z_2 + Z_3] = \frac{4}{\pi}.$$

But, we know that $\mathbb{E}[Z_1] = \mathbb{E}[Y_1]$; they are both line segments of length $1/3$ and arise from the same random process. Therefore, we see that

$$\mathbb{E}[Z_1 + Z_2 + Z_3] = \mathbb{E}[Z_1] + \mathbb{E}[Z_2] + \mathbb{E}[Z_3] = \mathbb{E}[Y_1] + \mathbb{E}[Y_2] + \mathbb{E}[Y_3]$$

so the final answer is $\frac{4}{\pi}$.

2 Variance of the Minimum of Uniform Random Variables

Let n be a positive integer and let $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \text{Uniform}[0, 1]$. Find $\operatorname{var} Y$, where

$$Y := \min\{X_1, \ldots, X_n\}.$$

Hint: You may need to perform integration by parts.

Solution:

We know that the density of Y is $f(y) = n(1-y)^{n-1}$, for $y \in [0,1]$, and $\mathbb{E}[Y] = (n+1)^{-1}$. It remains to compute (via integration by parts)

$$\begin{aligned}
\mathbb{E}[Y^2] &= \int_0^1 y^2 \cdot n(1-y)^{n-1}\,\mathrm{d}y \\
&= n\int_0^1 y^2(1-y)^{n-1}\,\mathrm{d}y \\
&= -y^2(1-y)^n\Big|_0^1 + 2\int_0^1 y(1-y)^n\,\mathrm{d}y \\
&= \frac{2}{n+1}\int_0^1 y(n+1)(1-y)^n\,\mathrm{d}y.
\end{aligned}$$

Since $g(y) := (n+1)(1-y)^n$ is the density of the minimum of $n+1$ i.i.d. Uniform$[0,1]$ random variables, we recognize the last integral as the expectation of this minimum, which is $1/(n+2)$. Thus,

$$\mathbb{E}[Y^2] = \frac{2}{(n+1)(n+2)}$$

and so

$$\begin{aligned}
\operatorname{var} Y &= \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \\
&= \frac{2}{(n+1)(n+2)} - \frac{1}{(n+1)^2} \\
&= \frac{2(n+1)-(n+2)}{(n+1)^2(n+2)} \\
&= \frac{n}{(n+1)^2(n+2)}.
\end{aligned}$$

Fun Fact: For a non-negative random variable X with density f_X, one can extend the tail sum formula to give

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_0^\infty x^2 f_X(x)\,\mathrm{d}x \\
&= \int_0^\infty \Big(\int_0^x 2s\,\mathrm{d}s\Big) f_X(x)\,\mathrm{d}x \\
&= \int_0^\infty 2s\int_s^\infty f_X(x)\,\mathrm{d}x\,\mathrm{d}s \\
&= \int_0^\infty 2s\,\mathbb{P}(X \ge s)\,\mathrm{d}s
\end{aligned}$$

and this gives another way to compute $\mathbb{E}[Y^2]$ in this problem. You can derive a similar formula to compute any moment $\mathbb{E}[X^k]$ for $k \in \mathbb{N}$.

3 Erasures, Bounds, and Probabilities

Alice is sending 1000 bits to Bob. The probability that a bit gets erased is p, and the erasure of each bit is independent of the others.

Alice is using a scheme that can tolerate up to one-fifth of the bits being erased. That is, as long as Bob receives at least 801 of the 1000 bits correctly, he can decode Alice's message.

In other words, Bob becomes unable to decode Alice's message only if 200 or more bits are erased. We call this a "communication breakdown", and we want the probability of a communication breakdown to be at most 10^{-6}.

1. Use Markov's inequality to upper bound p such that the probability of a communications breakdown is at most 10^{-6}.
2. Use Chebyshev's inequality to upper bound p such that the probability of a communications breakdown is at most 10^{-6}.
3. As the CLT would suggest, approximate the fraction of erasures by a Gaussian random variable (with suitable mean and variance). Use this to find an approximate bound for p such that the probability of a communications breakdown is at most 10^{-6}.

Solution:

1. Let the indicator random variable for the i^{th} bit's erasure be X_i. That is, $X_i = 1$ if the i^{th} bit is erased, and 0 otherwise. $\Pr(X_i = 1) = p$, and the X_is are all independent of one another.

   Let X be the total number of erasures. Then $X = X_1 + X_2 + \ldots + X_{1000}$. A communications breakdown happens if and only if $X \geq 200$. So we would like to have $\Pr(X \geq 200) \leq 10^{-6}$.

   Further, let μ_X denote the mean of X and σ_X^2 denote the variance of X. From previous discussion sections, $\mu_X = 1000p$ and $\sigma_X^2 = 1000p(1-p)$.

   Applying Markov's inequality to X (which is non-negative), we have:

$$\Pr(X \geq 200) \leq \frac{E[X]}{200} = \frac{1000p}{200} = 5p$$

   So, if p is such that $5p \leq 10^{-6}$, then our objective of $\Pr(X \geq 200) \leq 10^{-6}$ is met. So the upper bound for p is $10^{-6}/5$, or 2×10^{-7}.

2. Chebyshev's inequality states the following:

$$\Pr(\|X - \mu_X\| \geq k\sigma_X) \leq \frac{1}{k^2}$$

   So we need to choose a k given by:

$$k = \frac{200 - 1000p}{\sqrt{1000p(1-p)}}$$

   Note: The above is valid only for $200 - 1000p > 0$, or $p < 0.2$, since k has to be positive. But as we will see below, our upper bound for p will be below 0.2, so there is no problem.

Proceeding with the above value of k, and substituting for μ_X and σ_X, we obtain:

$$\Pr(\|X - 1000p\| \geq 200 - 1000p) \leq \frac{1}{\left(\frac{(200-1000p)^2}{1000p(1-p)}\right)}$$

Simplifying, we get:

$$\Pr(\|X - 1000p\| \geq 200 - 1000p) \leq \frac{p(1-p)}{40(1-5p)^2}$$

Now we know the following:

$$\begin{aligned}\Pr(X \geq 200) &= \Pr(X - 1000p \geq 200 - 1000p) \\ &\leq \Pr(\|X - 1000p\| \geq 200 - 1000p) \\ &\leq \frac{p(1-p)}{40(1-5p)^2}\end{aligned}$$

As before, to meet our objective, we just have to ensure that

$$\frac{p(1-p)}{40(1-5p)^2} \leq 10^{-6},$$

which yields an upper bound of about 3.998×10^{-5} for p.

3. Let Y be equal to the fraction of erasures, i.e. $\frac{X}{1000}$. Using properties of expectation and variance, we can see that

$$\begin{aligned}\mathbb{E}[Y] &= p \\ \operatorname{Var}(Y) &= \operatorname{Var}(X) \cdot \frac{1}{1000^2} = \frac{p(1-p)}{1000}\end{aligned}$$

Therefore, by Central Limit Theorem, we can say that Y is roughly a normal distribution with that mean and variance. Since we are interested in the event that $Y \geq 0.2$, let's figure out how many standard deviations above the mean 0.2 is:

$$\frac{0.2 - p}{\sqrt{\frac{p(1-p)}{1000}}} = \frac{(0.2 - p)\sqrt{1000}}{\sqrt{p(1-p)}}.$$

Therefore, the probability that we get a failure should be approximately (by CLT),

$$1 - \Phi\left(\frac{(0.2 - p)\sqrt{1000}}{\sqrt{p(1-p)}}\right)$$

where Φ is the CDF of a standard normal variable. Setting this to be at most 10^{-6} gives us

$$\Phi\left(\frac{(0.2-p)\sqrt{1000}}{\sqrt{p(1-p)}}\right) \geq 1-10^{-6}$$

And, since $\Phi^{-1}(1-10^{-6}) \approx 4.753$, we solve the inequality

$$\frac{(0.2-p)\sqrt{1000}}{\sqrt{p(1-p)}} \geq 4.753$$

This yields that we need $p \leq 0.1468$.

Note that this gives quite a different value from the previous parts. This is because the Central Limit Theorem gives a much tighter approximation for tail events than Markov's and Chebyshev's. Therefore, we do not need p to be so low to achieve a communication breakdown probability of 10^{-6}. The other bounds required us to need a probability of on the order of 10^{-5}, but here we realize that we only need it to be less than 0.1468. Quite drastic!

4 Sampling a Gaussian With Uniform

In this question, we will see one way to generate a normal random variable if we have access to a random number generator that outputs numbers between 0 and 1 uniformly at random.

As a general comment, remember that showing two random variables have the same CDF or PDF is sufficient for showing that they have the same distribution.

(a) First, let us see how to generate an exponential random variable with a uniform random variable. Let $U_1 \sim Uniform(0,1)$. Prove that $-\ln U_1 \sim Expo(1)$.

(b) Let $N_1, N_2 \sim \mathcal{N}(0,1)$, where N_1 and N_2 are independent. Prove that $N_1^2 + N_2^2 \sim Expo(1/2)$.

Hint: You may use the fact that over a region R, if we convert to polar coordinates $(x,y) \to (r,\theta)$, then the double integral over the region R will be

$$\iint_R f(x,y)\,dx\,dy = \iint_R f(r\cos\theta, r\sin\theta)\cdot r\,dr\,d\theta.$$

(c) Suppose we have two uniform random variables, U_1 and U_2. How would you transform these two random variables into a normal random variable with mean 0 and variance 1?

Hint: What part (b) tells us is that the point (N_1, N_2) will have a distance from the origin that is distributed as the square root of an exponential distribution. Try to use U_1 to sample the radius, and then use U_2 to sample the angle.

Solution:

(a) The CDF of an exponential $Expo(\lambda)$ distribution is $1-e^{-\lambda t}$. Let us prove that the $-\ln(U_1)$ also has the same CDF.

We see that

$$\begin{aligned}\mathbb{P}(-\ln(U_1) \le t) &= \mathbb{P}(\ln(U_1) \ge -t) \\ &= \mathbb{P}(U_1 \ge e^{-t}) \\ &= 1-e^{-t}\end{aligned}$$

This shows that $-\ln(U_1)$ has an exponential distribution with $\lambda = 1$.

(b) We compute the CDF of $N_1^2+N_2^2$. We want the probability that $N_1^2+N_2^2 \le t$ for some t. This means that we are integrating the joint distribution over a circle of radius $\sqrt{t}$, centered at the origin. We therefore compute the following integral

$$\iint_{(x,y):x^2+y^2\le t} \frac{1}{2\pi} e^{-(x^2+y^2)/2}\,dx\,dy = \int_0^{2\pi}\int_0^{\sqrt{t}} \frac{1}{2\pi} r e^{-r^2/2}\,dr\,d\theta$$

Evaluating this integral yields

$$\int_0^{2\pi} -\frac{e^{-r^2/2}}{2\pi}\Big|_0^{\sqrt{t}}\,d\theta = \int_0^{2\pi} \frac{1-e^{-t/2}}{2\pi}\,d\theta = 1-e^{-t/2}.$$

This proves that $N_1^2+N_2^2 \sim Expo(1/2)$.

(c) We will sample the point (N_1,N_2) using uniform random variables U_1 and U_2. We first sample the radius, which we know is an exponential distribution. Therefore, we know that $-2\ln(U_1)$ is an exponential $1/2$ distribution, so $\sqrt{-2\ln(U_1)}$ can be our radius. Since the (N_1,N_2) joint distribution is rotationally symmetric, we know that we can pick our angle uniformly at random once the radius is determined. Therefore, we let $\theta = 2\pi U_2$.

We will actually arrive at two Gaussians, so we can just take N_1, which will be

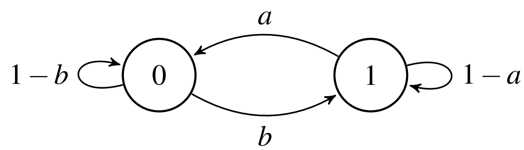
$$\boxed{\sqrt{-2\ln(U_1)}\cos(2\pi U_2)}$$

5 Markov Chain Terminology

In this question, we will walk you through terms related to Markov chains.

1. (Irreducibility) A Markov chain is irreducible if, starting from any state i, the chain can transition to any other state j, possibly in multiple steps.
2. (Periodicity) $d(i) := \gcd\{n>0 \mid P^n(i,i) = \mathbb{P}[X_n = i \mid X_0 = i] > 0\}$, $i \in \mathscr{X}$. If $d(i) = 1\ \forall i \in \mathscr{X}$, then the Markov chain is aperiodic; otherwise it is periodic.

3. (Matrix Representation) Define the transition probability matrix P by filling entry (i,j) with probability $P(i,j)$.
4. (Invariance) A distribution π is invariant for the transition probability matrix P if it satisfies the following balance equations: $\pi = \pi P$.



(a) For what values of a and b is the above Markov chain irreducible? Reducible?

(b) For $a = 1$, $b = 1$, prove that the above Markov chain is periodic.

(c) For $0 < a < 1$, $0 < b < 1$, prove that the above Markov chain is aperiodic.

(d) Construct a transition probability matrix using the above Markov chain.

(e) Write down the balance equations for this Markov chain and solve them. Assume that the Markov chain is irreducible.

Solution:

(a) The Markov chain is irreducible if both a and b are non-zero. It is reducible if at least one is 0.

(b) We compute $d(0)$ to find that:

$$d(0) = \gcd\{2,4,6,...\} = 2.$$

Thus, the chain is periodic.

(c) We compute $d(0)$ to find that:

$$d(0) = \gcd\{1,2,3,...\} = 1.$$

Thus, the chain is aperiodic.

(d)

$$\begin{bmatrix} 1-b & b \\ a & 1-a \end{bmatrix}$$

(e)

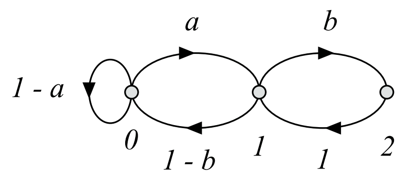
$$\pi(0) = (1-b)\pi(0) + a\pi(1),$$
$$\pi(1) = b\pi(0) + (1-a)\pi(1).$$

One of the equations is redundant. We throw out the second equation and replace it with $\pi(0) + \pi(1) = 1$. This gives the solution

$$\pi = \frac{1}{a+b}\begin{bmatrix} a & b \end{bmatrix}.$$

6 Analyze a Markov Chain

Consider the Markov chain $X(n)$ with the state diagram shown below where $a, b \in (0,1)$.



(a) Show that this Markov chain is aperiodic;

(b) Calculate $\mathbb{P}[X(1) = 1, X(2) = 0, X(3) = 0, X(4) = 1 \mid X(0) = 0]$;

(c) Calculate the invariant distribution;

(d) Let $T_i = \min\{n \geq 0 \mid X(n) = i\}$, T_i is the number of steps until we transit to state i for the first time. Calculate $\mathbb{E}[T_2 \mid X(0) = 1]$.

Solution:

(a) The Markov chain is irreducible because $a, b \in (0,1)$. Also, $P(0,0) > 0$, so that

$$\gcd\{n > 0 \mid P^n(0,0) > 0\} = \gcd\{1,2,3,\ldots\} = 1,$$

which shows that the Markov chain is aperiodic.
We can also notice from the definition of aperiodic Markov chain, that if a Markov chain has a self loop with nonzero probability ($P(0,0) > 0$ in this example), which says the smallest number of steps from a state to itself if 1, it is aperiodic.

(b) As a result of the Markov property, we know our state at timestep n depends only on timestep $n-1$. We see that the probability is

$$P(0,1) \times P(1,0) \times P(0,0) \times P(0,1) = a(1-b)(1-a)a.$$

(c) The balance equations are

$$\pi(0) = (1-a)\pi(0) + (1-b)\pi(1)$$
$$\pi(1) = a\pi(0) + \pi(2).$$

After some simple manipulations, we see that they imply the following equations:

$$a\pi(0) = (1-b)\pi(1)$$
$$b\pi(1) = \pi(2).$$

These equations express the equality of the probability of a jump from i to $i+1$ and from $i+1$ to i, for $i = 0$ and $i = 1$, respectively. These relations are called the "detailed balance equations". From these equations we find successively that

$$\pi(1) = \frac{a}{1-b}\pi(0) \text{ and } \pi(2) = b\pi(1) = \frac{ab}{1-b}\pi(0).$$

The normalization equation is

$$1 = \pi(0) + \pi(1) + \pi(2) = \pi(0)\left[1 + \frac{a}{1-b} + \frac{ab}{1-b}\right]$$
$$= \pi(0)\frac{1-b+a+ab}{1-b},$$

so that

$$\pi(0) = \frac{1-b}{1-b+a+ab}.$$

Thus,

$$\pi = \frac{1}{1-b+a+ab}\begin{bmatrix}1-b & a & ab\end{bmatrix}.$$

(d) We define

$$\beta(i) = \mathbb{E}[T_2 \mid X(0) = i], i = 0,1,2.$$

The FSE are $\beta(2) = 0$ and

$$\beta(0) = 1 + (1-a)\beta(0) + a\beta(1)$$
$$\beta(1) = 1 + (1-b)\beta(0).$$

The first equation is equivalent to

$$\beta(0) = \frac{1}{a} + \beta(1).$$

Substituting this expression in the second equation, we get

$$\beta(1) = 1 + (1-b)\left(\frac{1}{a} + \beta(1)\right) = (1-b)\beta(1) + \frac{1+a-b}{a},$$

so that

$$\beta(1) = \frac{1+a-b}{ab}.$$

7 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).

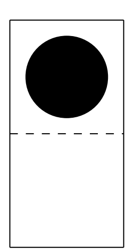


Figure 1: A straw with one boba currently inside. The straw only has enough room to fit two boba.

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

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

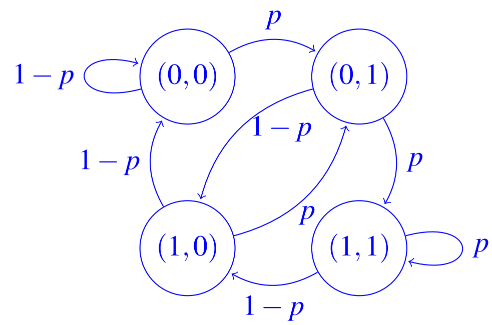


Figure 2: Transition diagram for the Markov chain.

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

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

(c) 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.

(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

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

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 $2p$.