CS 70 Spring 2021 — Discrete Mathematics and Probability Theory Course Notes — HW 10

# 1 Random Variables Warm-Up

Let $X$ and $Y$ be random variables, each taking values in the set $\{0,1,2\}$, with joint distribution

$$\begin{array}{lll} \mathbb{P}[X=0,Y=0]=1/3 & \mathbb{P}[X=0,Y=1]=0 & \mathbb{P}[X=0,Y=2]=1/3 \\ \mathbb{P}[X=1,Y=0]=0 & \mathbb{P}[X=1,Y=1]=1/9 & \mathbb{P}[X=1,Y=2]=0 \\ \mathbb{P}[X=2,Y=0]=1/9 & \mathbb{P}[X=2,Y=1]=1/9 & \mathbb{P}[X=2,Y=2]=0. \end{array}$$

(a) What are the marginal distributions of $X$ and $Y$?

(b) What are $\mathbb{E}[X]$ and $\mathbb{E}[Y]$?

(c) (optional) What are $\mathrm{Var}(X)$ and $\mathrm{Var}(Y)$?

(d) Let $I$ be the indicator that $X=1$, and $J$ be the indicator that $Y=1$. What are $\mathbb{E}[I]$, $\mathbb{E}[J]$ and $\mathbb{E}[IJ]$?

(e) In general, let $I_A$ and $I_B$ be the indicators for events $A$ and $B$ in a probability space $(\Omega,\mathbb{P})$. What is $\mathbb{E}[I_A I_B]$, in terms of the probability of some event?

**Solution:**

(a) By the law of total probabilty

$$\mathbb{P}[X=0]=\mathbb{P}[X=0,Y=0]+\mathbb{P}[X=0,Y=1]+\mathbb{P}[X=0,Y=2]=1/3+0+1/3=2/3$$

and similarly

$$\begin{aligned} \mathbb{P}[X=1] &= 0+1/9+0=1/9 \\ \mathbb{P}[X=2] &= 1/9+1/9+0=2/9. \end{aligned}$$

As a sanity check, these three numbers are all positive and they add up to $2/3+1/9+2/9=1$ as they should. The same kind of calculation gives

$$\begin{aligned} \mathbb{P}[Y=0] &= 1/3+0+1/9=4/9 \\ \mathbb{P}[Y=1] &= 0+1/9+1/9=2/9 \\ \mathbb{P}[Y=2] &= 1/3. \end{aligned}$$

(b) From the above marginal distributions, we can compute

$$\mathbb{E}[X] = 0\mathbb{P}[X=0] + 1\mathbb{P}[X=1] + 2\mathbb{P}[X=2] = 5/9$$
$$\mathbb{E}[Y] = 0\mathbb{P}[Y=0] + 1\mathbb{P}[Y=1] + 2\mathbb{P}[Y=2] = 8/9$$

(c) Again using our marginal distributions,

$$\mathbb{E}[X^2] = 0\mathbb{P}[X=0] + 1\mathbb{P}[X=1] + 4\mathbb{P}[X=2] = 1$$
$$\mathbb{E}[Y^2] = 0\mathbb{P}[Y=0] + 1\mathbb{P}[Y=1] + 4\mathbb{P}[Y=2] = 14/9$$

and thus

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 56/81$$
$$\mathrm{Var}(Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = 62/81.$$

We didn't ask you to do compute the covariance on the homework, but it is

$$\begin{aligned}
\mathrm{cov}(X,Y) &= \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] && (1)\\
&= 0\mathbb{P}[XY=0] + 1\mathbb{P}[XY=1] + 2\mathbb{P}[XY=2] + 4\mathbb{P}[XY=4] - 40/81 && (2)\\
&= 0\cdot 1/3 + 1\cdot 1/9 + 2\cdot 1/9 + 4\cdot 0 - 40/81 && (3)\\
&= -13/81. && (4)
\end{aligned}$$

(d) We know that taking the expectation of an indicator for some event gives the probability of that event, so

$$\mathbb{E}[I] = \mathbb{P}[X=1] = 1/9$$
$$\mathbb{E}[J] = \mathbb{P}[Y=1] = 2/9.$$

The random variable $IJ$ is equal to one if $I=1$ and $J=1$, and is zero otherwise. In other words, it is the indicator for the event that $I=1$ *and* $J=1$:

$$\mathbb{E}[IJ] = \mathbb{P}[I=1, J=1] = 1/9.$$

(e) By what we said in the previous part of the solution, $I_A I_B$ is the indicator for the event $A\cap B$, so

$$\mathbb{E}[I_A I_B] = \mathbb{P}[A\cap B].$$

# 2 Optimal Gambling

Jonathan has a coin that may be biased, but he doesn't think so. You disagree with him though, and he challenges you to a bet. You start off with $X_0$ dollars. You and Jonathan then play multiple rounds, and each round, you bet an amount of money of your choosing, and then coin is tossed. Jonathan will match your bet, no matter what, and if the coin comes up heads, you win and you take both yours and Jonathan's bet, and if it comes up tails, then you lose your bet.

(a) Now suppose you actually secretly know that the bias of the coin is $\frac{1}{2} < p < 1$! You use the following strategy: on each round, you will bet a fraction $q$ of the money you have at the start of the round. Let's say you play $n$ rounds. What is the probability that you win exactly $k$ of the rounds? What is the amount of money you would have if you win exactly $k$ rounds? [*Hint*: Does the order in which you win the games affect your profit?]

(b) Let $X_n$ denote the amount of money you have on round $n$. $X_0$ represents your initial assets and is a constant value. Show that $\mathbb{E}[X_n] = ((1-p)(1-q)+p(1+q))^n X_0$.

You may use the binomial theorem in your answer:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{(n-k)}$$

[*Hint*: Try computing a sum over the number of rounds you win out of the $n$ rounds you play - use your answers from the previous part.]

(c) What value of $q$ will maximize $\mathbb{E}[X_n]$? For this value of $q$, what is the distribution of $X_n$? Can you predict what will happen as $n \to \infty$? [*Hint*: Under this betting strategy, what happens if you ever lose a round?]

(d) The problem with the previous approach is that we were too concerned about expected value, so our gambling strategy was too extreme. Let's start over: again we will use a gambling strategy in which we bet a fraction $q$ of our money at each round. Express $X_n$ in terms of $n$, $q$, $X_0$, and $W_n$, where $W_n$ is the number of rounds you have won up until round $n$. [*Hint*: Does the order in which you win the games affect your profit?]

**Solution:**

(a) The number of rounds we win out of $n$ rounds played is described by a binomial distribution with $n$ trials and probability of success $p$. So the probability we win exactly $k$ rounds is $\binom{n}{k} p^k (1-p)^{n-k}$.

You win $k$ times and each time you win, your fortune is multiplied by $1+q$; you lose $n-k$ times, and each time you lose, your fortune is multiplied by $1-q$. Therefore, the amount of money you will have is $X_0(1+q)^k(1-q)^{n-k}$.

(b) As seen in the previous part, $X_n$ can take on $n$ different values, described by $X_0(1+q)^k(1-q)^{n-k}$ for $1 \le k \le n$. We also know the probability that $X_n$ takes each of those values, and so can compute expectation directly from the definition:

$$\begin{aligned}
\mathbb{E}[X_n] &= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \cdot X_0(1+q)^k(1-q)^{n-k} \\
&= X_0 \sum_{k=0}^{n} \binom{n}{k} (p(1+q))^k((1-p)(1-q))^{n-k} \\
&= ((1-p)(1-q)+p(1+q))^n X_0
\end{aligned}$$

(c) We want $(1-p)(1-q)+p(1+q)$ to be as large as possible. Note that this is linear in $q$, and the coefficient for $q$ is $p-(1-p)>0$. Hence, we should take $q$ to be as large as possible, which is 1 (you cannot bet more money than you actually have).

For this value of $q$, note that on each round you either double your money or go broke. Hence, the distribution is:

$$X_n = \begin{cases} 2^n X_0, & \text{with probability } p^n \\ 0, & \text{with probability } 1-p^n \end{cases}$$

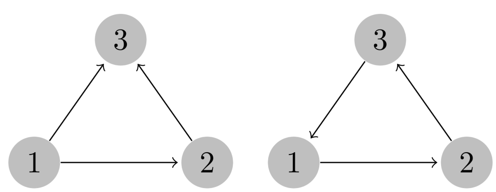
Uh-oh. As $n \to \infty$, the probability that you are broke approaches 1. The issue here is that your expected fortune grows large, but the probability that you are rich grows vanishingly small. In general, $X_n \to 0$ as $n \to \infty$ does not necessarily imply that $\mathbb{E}[X_n] \to 0$, which is what we see here.

(d) As in part (a): you win $W_n$ times and each time you win, your fortune is multiplied by $1+q$; you lose $n-W_n$ times, and each time you lose, your fortune is multiplied by $1-q$. Therefore,

$$X_n = X_0(1-q)^{n-W_n}(1+q)^{W_n}.$$

# 3 Random Tournaments

A *tournament* is a directed graph in which every pair of vertices has exactly one directed edge between them—for example, here are two tournaments on the vertices $\{1,2,3\}$:




In the first tournament above, $(1,2,3)$ is a *Hamiltonian path*, since it visits all the vertices exactly once, without repeating any edges, but $(1,2,3,1)$ is not a valid *Hamiltonian cycle*, because the tournament contains the directed edge $1 \to 3$ and not $3 \to 1$. In the second tournament, $(1,2,3,1)$ is a *Hamiltonian cycle*, as are $(2,3,1,2)$ and $(3,1,2,3)$; for this problem we'll say that these are all different Hamiltonian cycles, since their start/end points are different.

Consider the following way of choosing a random tournament $T$ on $n$ vertices: independently for each (unordered) pair of vertices $\{i,j\} \subset \{1,\ldots,n\}$, flip a coin and include the edge $i \to j$ in the graph if the outcome is heads, and the edge $j \to i$ if tails. What is the expected number of Hamiltonian paths in $T$? What is the expected number of Hamiltonian cycles?

**Solution:**

Each possible Hamiltonian path in the graph corresponds to a permutation $\sigma$ of the numbers $1,\ldots,n$, where $\sigma(1)$ is the starting vertex, $\sigma(2)$ is the second vertex visited, etc. If we write $I_\sigma$ for the indicator random variable that $\sigma$ corresponds to an actual Hamiltonian cycle in $T$, then

$$\mathbb{E}[\#\text{ Hamiltonian Paths}] = \mathbb{E}\left[\sum_{\sigma} I_{\sigma}\right] = \sum_{\sigma} \mathbb{P}[\sigma \text{ is a Hamiltonian path in } T]$$

In order for each $\sigma$ to correspond to an actual Hamiltonian path in $T$, the edges $\sigma(i) \to \sigma(i+1)$, for $i = 1, \ldots, n-1$ must all be included in the graph. Since the orientations of the edges in $T$ are independent, with $\sigma(i) \to \sigma(i+1)$ occurring with probability $1/2$, the probability that they are all included is $2^{-(n-1)}$. There are $n!$ possible permutations, so we have

$$\mathbb{E}[\#\text{ Hamiltonian Paths}] = \frac{n!}{2^{n-1}}.$$

The situation for Hamiltonian cycles is similar. Each possible Hamiltonian cycle each possible cycle corresponds to a permutation $\sigma$, but this time in order for $\sigma$ to be a valid Hamiltonian cycle, $T$ must include the edges $\sigma(i) \to \sigma(i+1)$ for all $i = 1, \ldots, n-1$, as well as the edge $\sigma(n) \to \sigma(1)$. As above, these $n$ edges are oriented independently of one another, so

$$\mathbb{E}[\#\text{ Hamiltonian Cycles}] = \frac{n!}{2^{n}}.$$

# 4 Class Enrollment

Lydia has just started her CalCentral enrollment appointment. She needs to register for a marine science class and CS 70. There are no waitlists, and she can attempt to enroll once per day in either class or both. The CalCentral enrollment system is strange and picky, so the probability of enrolling successfully in the marine science class on each attempt is $\mu$ and the probability of enrolling successfully in CS 70 on each attempt is $\lambda$. Also, these events are independent.

(a) Suppose Lydia begins by attempting to enroll in the marine science class everyday and gets enrolled in it on day $M$. What is the distribution of $M$?

(b) Suppose she is not enrolled in the marine science class after attempting each day for the first 5 days. What is the conditional distribution of $M$ given $M > 5$?

(c) Once she is enrolled in the marine science class, she starts attempting to enroll in CS 70 from day $M+1$ and gets enrolled in it on day $C$. Find the expected number of days it takes Lydia to enroll in both the classes, i.e. $\mathbb{E}[C]$.

(d) Suppose instead of attempting one by one, Lydia decides to attempt enrolling in both the classes from day 1. Let $M$ be the number of days it takes to enroll in the marine science class, and $C$ be the number of days it takes to enroll in CS 70. What is the distribution of $M$ and $C$ now? Are they independent?

(e) Let $X$ denote the day she gets enrolled in her first class and let $Y$ denote the day she gets enrolled in both the classes. What is the distribution of $X$?

(f) What is the expected number of days it takes Lydia to enroll in both classes now, i.e. $\mathbb{E}[Y]$.

(g) What is the expected number of classes she will be enrolled in by the end of 14 days?

**Solution:**

(a) $M \sim \text{Geometric}(\mu)$.

(b) Given that $M > 5$, the random variable $M$ takes values in $\{6,7,\dots\}$. For $i = 6,7,\dots,$

$$\mathbb{P}[M = i | M > 5] = \frac{\mathbb{P}[M = i \wedge M > 5]}{\mathbb{P}[M > 5]} = \frac{\mathbb{P}[M = i]}{\mathbb{P}[M > 5]} = \frac{\mu(1-\mu)^{i-1}}{(1-\mu)^5} = \mu(1-\mu)^{i-6}.$$

If $K$ denotes the additional number of days it takes to get enrolled in the marine science class after day 5, i.e. $K = M - 5$, then conditioned on $M > 5$, the random variable $K$ has the geometric distribution with parameter $\mu$. Note that this is the same as the distribution of $M$. This is known as the memoryless property of geometric distribution.

(c) We have $C - M \sim \text{Geometric}(\lambda)$. Thus $\mathbb{E}[M] = 1/\mu$ and $\mathbb{E}[C-M] = 1/\lambda$. And hence $\mathbb{E}[C] = \mathbb{E}[M] + \mathbb{E}[C-M] = 1/\mu + 1/\lambda$.

(d) $M \sim \text{Geometric}(\mu)$, $C \sim \text{Geometric}(\lambda)$. Yes they are independent.

(e) We have $X = \min\{M,C\}$ and $Y = \max\{M,C\}$. We also use the following definition of the minimum:

$$\min(m,c) = \begin{cases} m & \text{if } m \le c; \\ c & \text{if } m > c. \end{cases}$$

Now, for all $k \in \{1,2,\dots\}$, $\min(M,C) = k$ is equivalent to $(M = k) \cap (C \ge k)$ or $(C = k) \cap (M > k)$. Hence,

$$\begin{aligned} \mathbb{P}[X = k] &= \mathbb{P}[\min(M,C) = k] \\ &= \mathbb{P}[(M = k) \cap (C \ge k)] + \mathbb{P}[(C = k) \cap (M > k)] \\ &= \mathbb{P}[M = k] \cdot \mathbb{P}[C \ge k] + \mathbb{P}[C = k] \cdot \mathbb{P}[M > k] \end{aligned}$$

(since $M$ and $C$ are independent)

$$= [(1-\mu)^{k-1}\mu](1-\lambda)^{k-1} + [(1-\lambda)^{k-1}\lambda](1-\mu)^k$$

(since $M$ and $C$ are geometric)

$$\begin{aligned} &= ((1-\mu)(1-\lambda))^{k-1}(\mu + \lambda(1-\mu)) \\ &= (1 - \mu - \lambda + \lambda\mu)^{k-1}(\mu + \lambda - \mu\lambda). \end{aligned}$$

But this final expression is precisely the probability that a geometric r.v. with parameter $\mu + \lambda - \mu\lambda$ takes the value $k$. Hence $X \sim \text{Geom}(\mu + \lambda - \mu\lambda)$.

An alternative, slightly cleaner approach is to work with the *tail probabilities* of the geometric distribution, rather than with the usual point probabilities as above. In other words, we can work with $\mathbb{P}[X \geq k]$ rather than with $\mathbb{P}[X = k]$; clearly the values $\mathbb{P}[X \geq k]$ specify the values $\mathbb{P}[X = k]$ since $\mathbb{P}[X = k] = \mathbb{P}[X \geq k] - \mathbb{P}[X \geq (k+1)]$, so it suffices to calculate them instead. We then get the following argument:

$$\begin{aligned}
\mathbb{P}[X \geq k] &= \mathbb{P}[\min(M,C) \geq k] \\
&= \mathbb{P}[(M \geq k) \cap (C \geq k)] \\
&= \mathbb{P}[M \geq k] \cdot \mathbb{P}[C \geq k] && \text{since } M, C \text{ are independent} \\
&= (1-\mu)^{k-1}(1-\lambda)^{k-1} && \text{since } M, C \text{ are geometric} \\
&= ((1-\mu)(1-\lambda))^{k-1} \\
&= (1-\mu-\lambda+\mu\lambda)^{k-1}.
\end{aligned}$$

This is the tail probability of a geometric distribution with parameter $\mu + \lambda - \mu\lambda$, so we are done.

(f) From part (e) we get $\mathbb{E}[X] = 1/(\mu + \lambda - \mu\lambda)$. From part (d) we have $\mathbb{E}[M] = 1/\mu$ and $\mathbb{E}[C] = 1/\lambda$. We now observe that $\min\{m,c\} + \max\{m,c\} = m + c$. Using linearity of expectation we get $\mathbb{E}[X] + \mathbb{E}[Y] = \mathbb{E}[M] + \mathbb{E}[C]$. Thus $\mathbb{E}[Y] = 1/\mu + 1/\lambda - 1/(\mu + \lambda - \mu\lambda)$.

(g) Let $I_M$ and $I_C$ be the indicator random variables of the events "$M \leq 14$" and "$C \leq 14$" respectively. Then $I_M + I_C$ is the number of classes she will be enrolled in within 14 days. Hence the answer is $\mathbb{E}[I_M] + \mathbb{E}[I_C] = \mathbb{P}[M \leq 14] + \mathbb{P}[C \leq 14] = 1 - (1-\mu)^{14} + 1 - (1-\lambda)^{14}$