Statistics 100A
Homework 3 Solutions

Ryan Rosario

Chapter 4

1. *Two balls are chosen randomly from an urn containing 8 white, 4 black, and 2 orange balls. Suppose that we win \$2 for each black ball selected and we lose \$1 for each white ball selected. Let X denote our winnings. What are the possible values of X, and what are the probabilities associated with each value?*

   First construct the sample space of all possible draws. We must draw 2 balls at random from this url:

$$S = \{(W,W), (W,B), (W,O), (B,W), (B,O), (B,B), (O,W), (O,B), (O,O)\}$$

   Also, we need to assign values of X for each $s \in S$.

s	X
(W,W)	-2
(W,B)	$+1$
(W,O)	-1
(B,W)	$+1$
(B,O)	$+2$
(B,B)	$+4$
(O,W)	-1
(O,B)	$+2$
(O,O)	0

   Now we must compute the probabilities for each value of X.

   $X = +1$ when we draw a white ball and black ball, or when we draw a black ball and a white ball. We assume sampling without replacement:

$$P(X=1) = \frac{8}{14} \cdot \frac{4}{13} + \frac{4}{14} \cdot \frac{8}{13} = \boxed{0.3516}$$

   $X = -1$ when we draw a white ball and orange ball, or an orange ball and white ball.

$$P(X=-1) = \frac{8}{14} \cdot \frac{2}{13} + \frac{2}{14} \cdot \frac{8}{13} = \boxed{0.176}$$

   $X = +2$ when we draw a black ball followed by an orange ball, or an orange ball followed by a black ball.

$$P(X=+2) = \frac{4}{14} \cdot \frac{2}{13} + \frac{2}{14} \cdot \frac{4}{13} = \boxed{0.088}$$

   $X = -2$ when we draw two white balls.

$$P(X=-2) = \frac{8}{14} \cdot \frac{7}{13} = \boxed{0.308}$$

$X = +4$ when two black balls are drawn.

$$P(X = +4) = \frac{4}{14} \cdot \frac{3}{13} = \boxed{0.066}$$

$X = 0$ when two orange balls are drawn.

$$P(X = 0) = \frac{2}{14} \cdot \frac{1}{13} = \boxed{0.011}$$

Then the probability distribution is given by

x	P(X = x)
-2	0.308
-1	0.176
0	0.011
$+1$	0.3516
$+2$	0.088
$+4$	0.066

4. *Five men a 5 women are ranked according to their scores on an examination. Assume that no two scores are alike and all 10! possible rankings are equally likely. Let X denote the highest ranking achieved by a woman. Find $P(X = i), i = 1, 2, 3, \ldots, 8, 9, 10..$*

   First off, note that the lowest possible position is 6, which means that all five women score worse than the five men. So $\boxed{P(X = i) = 0 \text{ for } i = 7, 8, 9, 10}$.

   Now let's find how many different ways there are to arrange the 10 people such that the women all scored lower than the men. There are 5! ways to arrange the women and 5! ways to arrange the men, so

$$P(X = 6) = \frac{5!5!}{10!} = \boxed{0.004}$$

   Now consider the top woman scoring 5th on the exam.

				W					

   There are 5 possible positions for the lower scoring women, and we have 4 women that must be assigned to these ranks, which can be accomplished in $\binom{5}{4}$ different ways. Additionally, these 5 women can be arranged in 5! ways, and the men can be arranged in 5! ways. Thus,

$$P(X = 5) = \frac{\binom{5}{4} 5!5!}{10!} = \boxed{0.0198}$$

   Similarly for $P(X = 4)$, there are 6 positions for the 4 remaining women. The women can be arranged in 5! ways and the men can be arranged in 5! ways.

$$P(X = 4) = \frac{\binom{6}{4} 5!5!}{10!} = \boxed{0.0595}$$

and so on...

$$P(X=3) = \frac{\binom{7}{4} 5!5!}{10!} \approx \boxed{0.139}$$

$$P(X=2) = \frac{\binom{8}{4} 5!5!}{10!} \approx \boxed{0.278}$$

$$P(X=1) = \frac{\binom{9}{4} 5!5!}{10!} = \boxed{0.5}$$

8. *If the die in problem 7 is assumed fair, calculate the probabilities associated with the random variables in parts (a) through (d).*

   Let $X = f(X_1, X_2)$ where X_1 and X_2 are the faces showing on the first and second roll and f is given by

   (a) $f(X_1, X_2) = \max\{X_1, X_2\}$

   The maximum value of the two rolls is 1 and the maximum value of the two rolls is 6. Therefore $X = \{1, 2, 3, 4, 5, 6\}$.

x	Possibilities	$P(X=x)$
1	{⚀ and ⚀}	$\frac{1}{36}$
2	{⚁ and (⚀⚁)} or {(⚀) and ⚁}	$\frac{3}{36}$
3	{⚂ and (⚀⚁)} or {(⚀⚁) and ⚂}	$\frac{5}{36}$
4	{⚃ and (⚀⚁⚂⚃)} or {(⚀⚁⚂) and ⚃}	$\frac{7}{36}$
5	{⚄ and (⚀⚁⚂⚃⚄)} or {(⚀⚁⚂⚃) and ⚄}	$\frac{9}{36}$
6	{⚅ and (⚀⚁⚂⚃⚄⚅)} or {(⚀⚁⚂⚃⚄) and ⚅}	$\frac{11}{36}$

   (d) $f(X_1, X_2) = X_1 - X_2$

   In the extreme case, the first roll shows a 1 and the second shows a 6, or vice verse. Then, $X = \{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$.

x	Possibilities	$P(X=x)$
-5	{(⚀,⚅)}	$\frac{1}{36}$
-4	{(⚀,⚄), (⚁,⚅)}	$\frac{2}{36}$
-3	{(⚀,⚃), (⚁,⚄), (⚂,⚅)}	$\frac{3}{36}$
-2	{(⚀,⚂), (⚁,⚃), (⚂,⚄), (⚃,⚅)}	$\frac{4}{36}$
-1	{(⚀,⚁), (⚁,⚂), (⚂,⚃), (⚃,⚄), (⚄,⚅)}	$\frac{5}{36}$
0	{(⚀,⚀), (⚁,⚁), (⚂,⚂), (⚃,⚃), (⚄,⚄), (⚅,⚅)}	$\frac{6}{36}$

   For brevity, note that $P(X=i) = P(X=-i)$ for $i = 1, 2, 3, 4, 5, 6$.

14. *Five distinct numbers are randomly distributed to players numbered 1 through 5. Whenever two players compare their numbers, the one with the highest one is declared the winner. Initially, players 1 and 2 compare their numbers; the winner then compares with player 3, and so on. Let X denote the number of times player 1 is a winner. Find $P(X = i), i = 0, 1, 2, 3, 5$. Find the probability that player 1 wins an even number of times.*

This problem ended up being much easier than the solution lead me to believe!

We let X be the number of times player 1 is a winner.

Let's start with the case that player 1 never wins. That is, player 1 loses to player 2. If we construct the sample space, we would see that $P(X = 0) = \frac{1}{2}$ because exactly one-half of the 5! permutations have the first number (player 1) greater than the second number (player 2).

$$P(X = 0) = \frac{\frac{1}{2}5!}{5!} = \boxed{\frac{1}{2}}$$

Player 1 wins exactly 1 game if player 3 has a larger number than player 1, but player 1 has a larger number than player 2. The number of ways this can happen is the same as the number of ways that player 2 loses to player 1 and player 3 (hypothetically), etc. Per the solution I used, $P(X = 1) = P(Y_2 < Y_1 < Y_3)$ where Y_i denotes the number given to player i. When $i \neq j \neq$, there are 3! ways to arrange the inequality $Y_i < Y_j < Y_k$. Exactly one of these arrangements gives us the inequality $Y_2 < Y_1 < Y_3$. Therefore,

$$P(X = 1) = \frac{\frac{1}{3!}5!}{5!} = \frac{1}{3!} = \boxed{\frac{1}{6}}$$

Now consider the case that player 1 wins exactly 2 games. This means $Y_1 > Y_2$ and $Y_1 > Y_3$ but $Y_1 < Y_4$. The easiest way to do this is to consider all of the possible ways that this can happen,

$$\begin{aligned}
S_{\text{Player 1 wins 2 games.}} &= \{(Y_1, Y_2, Y_3, Y_4)\} \\
&= \{(3,1,2,4), (3,1,2,5), (3,2,1,4), (3,2,1,5), (4,1,2,5), (4,2,1,5) \\
&\qquad (4,1,3,5), (4,3,1,5), (4,2,3,5), (4,3,2,5)\}
\end{aligned}$$

Then,

$$P(X = 2) = \frac{10}{5!} = \boxed{\frac{1}{12}}$$

To win 3 games, $Y_1 > Y_2$, and $Y_1 > Y_3$ and $Y_1 > Y_4$ but $Y_1 < Y_5$. One way to think of this is that we must restrict $Y_1 = 4$ and $Y_5 = 5$ and vary the others, giving 3! arrangements. Or, we can just list out the different combinations.

$$\begin{aligned}
S_{\text{Player 1 wins 3 games.}} &= \{(Y_1, Y_2, Y_3, Y_4, Y_5)\} \\
&= \{(4,1,2,3,5), (4,2,3,1,5), (4,3,2,1,5), \\
&\qquad (4,1,3,2,5), (4,2,1,3,5), (4,3,1,2,5)\}
\end{aligned}$$

$$P(X = 3) = \frac{3!}{5!} = \frac{1}{20}$$

To win all 4 games, $Y_1 > Y_2 > Y_3 > Y_4 > Y_5$.

$$\begin{aligned} S_{\text{Player 1 wins 4 games.}} &= \{(Y_1, Y_2, Y_3, Y_4, Y_5)\} \\ &= \{(5, *, *, *, *)\} \end{aligned}$$

where $*$ denotes a wildcard meaning these players have some other number besides 5 here. For brevity, I do not list all of the arrangements, since we have enough information to solve the problem once we restrict $Y_1 = 5$. Thus,

$$P(X = 4) = \frac{4!}{5!} = \frac{1}{5}$$

Finally,

$$P(X = \text{even}) = P(X = 2) + P(X = 4) = \frac{1}{12} + \frac{1}{5} = \frac{17}{60}$$

19. *If the distribution of X is given by*

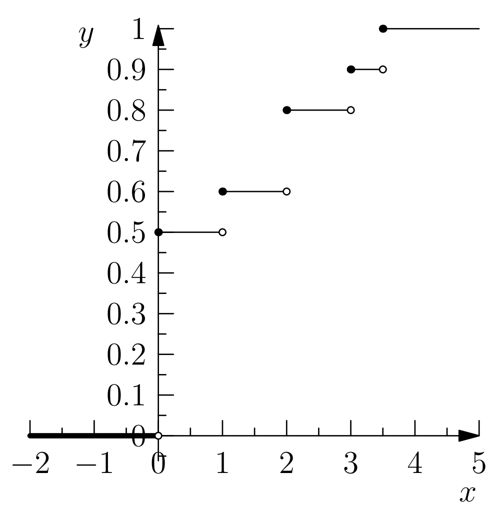
$$F(b) = \begin{cases} 0 & b < 0 \\ \frac{1}{2} & 0 \le b < 1 \\ \frac{3}{5} & 1 \le b < 2 \\ \frac{4}{5} & 2 \le b < 3 \\ \frac{9}{10} & 3 \le b < 3.5 \\ 1 & b \ge 3.5 \end{cases}$$

calculate the probability mass function of X.

First we must find $P(X = i), i = 0, 1, 2, 3, 3.5$. To do that, we note that $F(b) = P(X \le b)$ and that

$$P(X = i) = P(X \le i) - P(X < i) = F(i) - P(X < i)$$

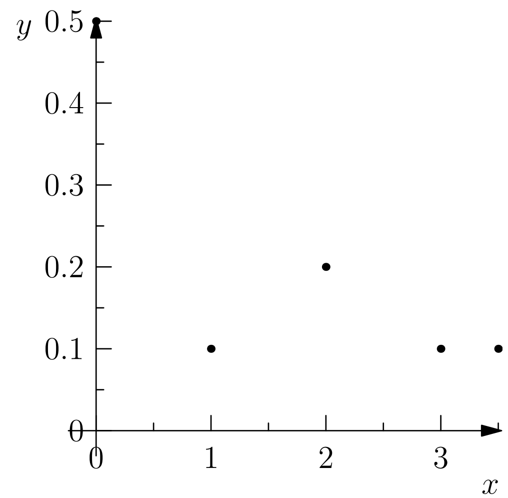
This is really just a jump function between values of b. F is also called the *cumulative distribution function (CDF)*.



$P(X=i)$	$F(i)-P(X<i)$	Calculation	$p(i)$
$P(X=0)$	$F(0)-P(X<0)$	$\frac{1}{2}-0$	$\frac{1}{2}$
$P(X=1)$	$F(1)-P(X<1)$	$\frac{3}{5}-\frac{1}{2}$	$\frac{1}{10}$
$P(X=2)$	$F(2)-P(X<2)$	$\frac{4}{5}-\frac{3}{5}$	$\frac{1}{5}$
$P(X=3)$	$F(3)-P(X<3)$	$\frac{9}{10}-\frac{4}{5}$	$\frac{1}{10}$
$P(X=3.5)$	$F(3.5)-P(X<3.5)$	$1-\frac{9}{10}$	$\frac{1}{10}$

Which is represented as,

$$\boxed{p(x)=\begin{cases}\frac{1}{2} & \text{if } x=0\\ \frac{1}{10} & \text{if } x=1\\ \frac{1}{5} & \text{if } x=2\\ \frac{1}{10} & \text{if } x=3\\ \frac{1}{10} & \text{if } x=3.5\\ 0 & \text{otherwise}\end{cases}}$$



20. *A gambling book recommends the following "winning strategy" for the game of roulette. It recommends that a gambler bet \$1 on red. If red appears (which has a probability of $\frac{18}{38}$), then the gambler should take her \$1 profit and quit, If the gambler loses the bet (which has probability $\frac{20}{38}$ of occurring), she should make additional \$1 bets on red on each of the next two spins of the roulette wheel and then quit. Let X denote the gambler's winnings when she quits.*

(a) *Find $P(X > 0)$.*

First construct the sample space. Let R be the event that the ball lands on a red space and let R' be the complement of R. The sequence of spins will be of length 1 (if she wins on the first spin), or of length 3 (if she does not win on the first spin).

$$S = \{R, R'RR, R'R'R, R'RR', R'R'R'\}$$

If the wheel lands on a red space, then she wins \$1 *profit*, otherwise she wins nothing and loses the \$1 she bet. We can compute the values for X.

s	X
R	1
$R'RR$	1
$R'R'R$	-1
$R'RR'$	-1
$R'R'R'$	-3

Then $P(X > 0) = P(X = 1)$ which only happens one of two ways; she either wins on the first bet, or she loses on the first bet but wins on the next two spins. We assume that each spin of the roulette wheel is independent. Then,

$$\begin{aligned} P(X > 0) &= P(X = 1) \\ &= P(R) + P(R'RR) \\ &= \frac{18}{38} + \frac{20}{38}\frac{18}{38}\frac{18}{38} \\ &= \boxed{0.5918} \end{aligned}$$

The probability that she wins any profit is 0.5918.

(b) *Are you convinced that the strategy is indeed a "winning" strategy? Explain your answer!*

Typically this question would be answered after calculating $E(X)$. Based on the probability in the previous part, it *may* seem like a winning strategy since the probability of profiting from the game is 0.5918 which is slightly better than flipping a coin. Actually, this is $\boxed{\text{not}}$ a winning strategy, as we will see in the next part.

(c) *Find $E(X)$.*

By definition,

$$E(X) = \sum_{i=1}^{n} xP(X = x)$$

.

So, we need to construct the probability distribution.

$x =$	$P(X = x)$
-3	$P(R'R'R') = \left(\frac{20}{38}\right)^3 = 0.1458$
-1	$P(R'R'R) + P(R'RR') = 2 \cdot \frac{18}{38} \cdot \left(\frac{20}{38}\right)^2 = 0.2624$
$+1$	(part a) $= 0.5918$

Then,

$$E(X) = -3 \cdot 0.1458 + (-1) \cdot 0.2624 + 1 \cdot 0.5918 = -0.108 \approx \boxed{-\$0.11}$$

She should expect to **lose** 11 cents. Therefore, this is not a winning strategy.

21. *A total of 4 buses carrying 148 students from the same school arrives at a football stadium. The buses carry, respectively, 40, 33, 25, and 50 students. One of the students is randomly selected. Let X denote the number of students that were on the bus carrying this randomly selected student. One of the 4 bus drivers is also randomly selected selected. Let Y denote the number of students on her bus.*

(a) *Which of $E(X)$ or $E(Y)$ do you think is larger? Why?*

In the first sampling method, we draw one of 148 students at random. Naturally, the probability of choosing a student from the fullest bus is higher than the probability of drawing a student from the other buses. Thus, the fuller buses are weighted higher than the other buses, so we would expect $E(X)$ to be larger. In the second method, one of the four bus drivers is randomly selected. Each bus driver is **equally likely** to be chosen! Each bus is weighted the sum, therefore, $E(Y) \leq E(X)$. Equality holds when each bus contains the same number of students.

Analogy: Consider calculating your GPA two different ways. Suppose your grades this quarter are as follows: A, C, C. The class that you earned an A in is worth 5 units and the classes you earned the Cs in are worth 4 units. Computing your GPA using the standard method of weighting the grades by units would weight the higher grade heavier than each of the two Cs. This calculation is similar to the calculation of $E(X)$. The GPA in this case is 3.00.

Now suppose we ignore the course units and recompute the GPA such that each grade is weighted the same. Then the GPA is 2.67. This calculation is similar to $E(Y)$. In this particular case, we see that the weighted method is higher.

(b) *Compute $E(X)$ and $E(Y)$.*

First we need the probability distributions for X and Y.

X corresponds to the method of drawing a student at random. The probability that a student comes from bus $i, i = \{A, B, C, D\}$ is simply the number of students on that bus divided by the total number of students.

$x =$	$P(X = x)$
40	$\frac{40}{148}$
33	$\frac{33}{148}$
25	$\frac{25}{148}$
50	$\frac{50}{148}$

Then, $E(X) = 40 \cdot \frac{40}{148} + 33 \cdot \frac{33}{148} + 25 \cdot \frac{25}{148} + 50 \cdot \frac{50}{148} = \boxed{39.28}$.

Y corresponds to the method of drawing a bus driver at random. The probability that a bus driver comes from a particular bus is constant, $\frac{1}{4}$.

$x =$	$P(X = x)$
40	$\frac{1}{4}$
33	$\frac{1}{4}$
25	$\frac{1}{4}$
50	$\frac{1}{4}$

Then, $E(Y) = 40 \cdot \frac{1}{4} + 33 \cdot \frac{1}{4} + 25 \cdot \frac{1}{4} + 50 \cdot \frac{1}{4} = \boxed{37}$.

30. *A person tosses a fair coin until a tail appears for the first time. If the tail appears on the nth flip, the person wins 2^n dollars. Let X denote the player's winnings. This problem is known as the St. Petersburg paradox.*

This is an example of the geometric distribution. We use the geometric distribution when we want to calculate the probability that it takes n trials to get a success. In this case, a success is the event that a tail is showing.

(a) *Show that $E(X) = +\infty$.*

Proof. We know that X denotes the player's winnings, so $X = 2^n$ for $n \geq 1$. The probability that a player wins on the nth game is $\left(\frac{1}{2}\right)^n$. Then,

$$E(X) = \sum_{i=1}^{\infty} xP(X = x) = \sum_{i=1}^{\infty} 2^n \cdot \left(\frac{1}{2}\right)^n = \sum_{i=1}^{\infty} 1 = +\infty$$

□

(b) *Would you be willing to play \$1 million to play this game once?*

Hell $\boxed{\text{no}}$! Suppose you pay \$1 million to play once and you win on the first game. You only win \$2. Suppose you lose on the first game (which is likely), then you lost your bet. $E(X) = \infty$ is the expected value in the long run, not on one game!

(c) *Would you be willing to play \$1 million for each game if you could play for as long as you liked and only had to settle up when you stopped playing?*

$\boxed{\text{Yes}}$, the expectation in the long run is "something large," so we could expect to profit. This is not a universally accepted answer though. From Wikipedia,

"According to the usual treatment of deciding when it is advantageous and therefore rational to play, you should therefore play the game at any price if offered the opportunity. Yet, in published descriptions of the paradox, e.g. (Martin, 2004), many people expressed disbelief in the result. Martin quotes Ian Hacking as saying "few of us would pay even \$25 to enter such a game" and says most commentators would agree."

I highly recommend the Wikipedia article on the St. Petersburg paradox!

37. *Find Var(X) and Var(Y) for X and Y as given in Problem 21.*

By definition,

$$\text{Var}(X) = E\left[(X - E(X))^2\right] = E(X^2) - E(X)^2$$

x	x^2	$P(X = x)$
40	1600	$\frac{40}{148}$
33	1089	$\frac{33}{148}$
25	625	$\frac{25}{148}$
50	2500	$\frac{50}{148}$

$$E(X^2) = \sum_{i=1}^{n} x^2 P(X = x) = 1600 \cdot \frac{40}{148} + 1089 \cdot \frac{33}{148} + 625 \cdot \frac{25}{148} + 2500 \cdot \frac{50}{148} = 1625.4$$

Thus,

$$\text{Var}(X) = E(X^2) - E(X)^2 = 1625.4 - 39.28^2 = \boxed{82.2}$$

Similarly,

y	y^2	$P(Y = y)$
40	1600	$\frac{1}{4}$
33	1089	$\frac{1}{4}$
25	625	$\frac{1}{4}$
50	2500	$\frac{1}{4}$

$$E(Y) = \frac{1}{4}(1600 + 1089 + 625 + 2500) = 1453.5$$

$$\text{Var}(Y) = E(Y^2) - E(Y)^2 = 1453.5 - 37^2 = \boxed{84.5}$$

38. *If* $E(X) = 1$ *and* $Var(X) = 5$, *find*

Recall the following properties of expectation and variance

$$E(c) = c, \qquad E(cX) = cE(X), \qquad E(X+Y) = E(X) + E(Y)$$

$$\text{Var}\,(c) = 0, \qquad \text{Var}\,(cX) = c^2\text{Var}\,(x)\,, \qquad \text{Var}\,(X+Y) = \text{Var}\,(X) + \text{Var}\,(Y) \text{ iff X, Y indpdt}$$

$$\text{Var}\,(x) = E(X^2) - E(X)^2$$

(a) $E\left[(2+X)^2\right]$

$$\begin{aligned}
E\left[(2+X)^2\right] &= E\left[(2+X)(2+X)\right] \\
&= E\left[4 + X^2 + 4X\right] \\
&= E(4) + E(X^2) + E(4X) \\
&= 4 + E(X^2) + 4E(X) \\
&\quad \text{Recall that Var}\,(x) = E(X^2) - E(X)^2 \\
&= 4 + \left(\text{Var}\,(X) + E(X)^2\right) + 4E(X) \\
&= 4 + \left(5 + 1^2\right) + 4 \cdot 1 \\
&= \boxed{14}
\end{aligned}$$

(b) $\text{Var}\,(4+3X)$

$$\begin{aligned}
\text{Var}\,(4+3X) &= \text{Var}\,(4) + \text{Var}\,(3X) \\
&= 0 + 3^2\text{Var}\,(X) \\
&= 9 \cdot \text{Var}\,(X) \\
&= 9 \cdot 5 \\
&= \boxed{45}
\end{aligned}$$