CSE 21 Fall 2018
Homework 8
Due: **Sunday, December 9**, 2018 at 10:00pm

INSTRUCTIONS

Homework should be done in groups of **one to three** people. You are free to change group members at any time throughout the quarter. Problems should be solved together, not divided up between partners. Homework must be submitted through **Gradescope** by **a single representative**. Submissions must be received by **10:00pm** on the due date, and there are no exceptions to this rule.

You will be able to look at your scanned work before submitting it. Please ensure that your submission is **legible** (neatly written and not too faint) or your homework may not be graded.

Students should consult their textbook, class notes, lecture slides, instructors, TAs, and tutors when they need help with homework. Students should not look for answers to homework problems in other texts or sources, including the Internet. You may ask questions about the homework in office hours, but **not on Piazza**.

Your assignments in this class will be evaluated not only on the correctness of your answers, but on your ability to present your ideas clearly and logically. You should **always explain** how you arrived at your conclusions and **justify your answers** with mathematically sound reasoning. Whether you use formal proof techniques or write a more informal argument for why something is true, your answers should always be well-supported. Your goal should be to **convince the reader** that your results and methods are sound.

For questions that require pseudocode, you can follow the same format as the textbook, or you can write pseudocode in your own style, as long as you specify what your notation means. For example, are you using "=" to mean assignment or to check equality? You are welcome to use any algorithm from class as a subroutine in your pseudocode. For example, if you want to sort list A using InsertionSort, you can call InsertionSort(A) instead of writing out the pseudocode for InsertionSort.

REQUIRED READING Rosen 7.1 through 7.4

KEY CONCEPTS: basic probability principles, independence, conditional probability, expected value, conditional expectation, linearity of expectation, variance and concentration, Bayes' Theorem

---

1. (a.) (4 points) Assume that every time you go on a job interview, your chance of getting a job offer is 25%. How many job interviews must you go on so that the probability of your getting a job offer is greater than 95%?

   Solution: Suppose that A is the event that you get the job offer. Then A^c is the event that you do not get the job offer. $P(A) = 1/4$ and $P(A^c) = 1 - 1/4 = 3/4$. The probability that after n attempts you did not get offered the job is $P(A^c)^n = 3^n/4^n$. So the probability that after n attempts, you did get offered the job is $1 - P(A^c)^n = 1 - 3^n/4^n$. So the solution of this problem is to solve for n such that $1 - 3^n/4^n > 95/100$.

   $$1 - 3^n/4^n > 95/100 \implies 4^n/3^n > 100/5 \implies n > \log_{4/3}(100/5) \approx 10.8$$

   So, it would take 11 attempts or more to ensure that you have a 95% chance of getting a job offer.

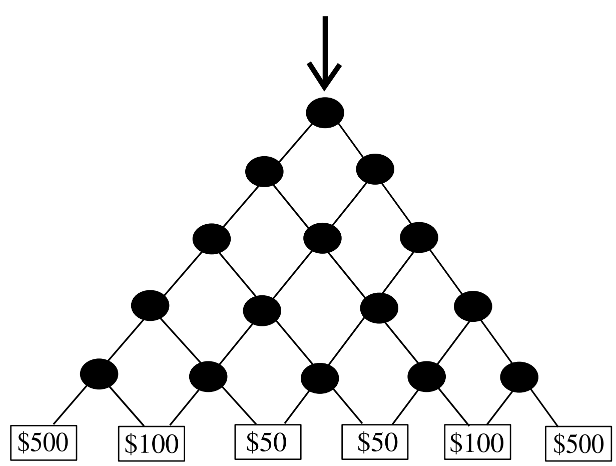
   (b.) (4 points) Suppose that you roll two dice and don't get to look at the outcome. Your friend looks at the outcome and tells you honestly that at least one of the dice came up 6. What is the probability that the sum of your two dice is 8?

   Solution: Let A be the event that at least one of your rolls is a 6 and let B be the event that the sum of your two dice is 8. The problem is to compute the probability that B occurs given that A has occurred, which is $P(B|A) = \dfrac{P(B \cap A)}{P(A)}$.

   A^c is the event that neither of your rolls are 6 and $P(A^c) = (5/6)^2$. Therefore $P(A) = 1 - (5/6)^2 = 11/36$. $B \cap A$ is the event that at least one of your dice is 6 and the sum is 8. It is clear that $B \cap A = \{(6,2),(2,6)\}$ and the sample space has 36 elements therefore $P(B \cap A) = 2/36$.

   So the answer is $P(B|A) = \dfrac{P(B \cap A)}{P(A)} = \dfrac{2/36}{11/36} = 2/11$.

2. In a variant of the game Plinko from the gameshow The Price is Right, the player has one disk which they insert into the top of the board shown below. Each time the disk hits a peg (shown as a black circle), it has a 50% chance of falling to the left, and a 50% chance of falling to the right. Eventually, the disk lands in one of the bins at the bottom of the board. Each bin is marked with a dollar amount, and the player wins the amount of money shown on the bin in which the disk lands.



(a) (3 points) Suppose that bin 1 is the leftmost bin, and bins are numbered from 1 to 6 reading from left to right. In terms of i, find a formula for the probability that the disk falls into bin i.

Solution: Since every path to a bin requires 5 left or right moves, the bin number in which the disk lands depends upon how many of those moves are left moves and how many are right moves. To land in bin i, the disk must make $i-1$ right moves, and the rest left moves. The number of ways to select these $i-1$ right moves out of the 5 total moves is $\binom{5}{i-1}$, so the number of paths to bin i is $\binom{5}{i-1}$. The number of possible paths to any bin is given by $2^5 = 32$ since there are 2 choices at each of 5 pegs. Since each of these 32 paths is equally likely, the probability that the disk falls into bin i is given by the binomial distribution

$$P(\text{disk falls into bin } i) = \frac{\binom{5}{i-1}}{32}.$$

(b) (3 points) How much money does the player expect to win at this game?

Solution: Using the formula for expected value with the probabilites from part (a) gives

$$\frac{1}{32} * 500 + \frac{5}{32} * 100 + \frac{10}{32} * 50 + \frac{10}{32} * 50 + \frac{5}{32} * 100 + \frac{1}{32} * 500 = \$93.75.$$

(c) (3 points) How much money does the player expect to win if you know that the disk falls to the right the first time it hits a peg?

Solution: Now the conditional probability of the disk falling into bin i is $\frac{\binom{4}{i-2}}{16}$ if $i > 1$ and 0 if $i = 1$. The expected value formula gives

$$\frac{1}{16} * 100 + \frac{4}{16} * 50 + \frac{6}{16} * 50 + \frac{4}{16} * 100 + \frac{1}{16} * 500 = \$93.75.$$

Note that this is the same as part (a). Since the values on the bins are symmetric from left to right, knowing that the disk goes right at the first peg does not tell us any additional information about how much money is expected to be won.

(d) (3 points) How much money does the player expect to win if you know that the disk falls to the right the first and second time it hits a peg?

Solution: Now the conditional probability of the disk falling into bin i is $\frac{\binom{3}{i-3}}{8}$ if $i > 2$ and 0 if $i = 1$ or $i = 2$. The expected value formula gives

$$\frac{1}{8} * 50 + \frac{3}{8} * 50 + \frac{3}{8} * 100 + \frac{1}{8} * 500 = \$125.00$$

Note that here the expected value has gone up, since knowing that the disk goes right at the first two opportunities makes it even more likely to fall in the rightmost bin, which is worth a lot.

3. Gummi bears come in twelve flavors, and you have one of each flavor. Suppose you split your gummi bears among three people (Adam, Beth, Charlie) by randomly selecting a person to receive each gummi bear. Each person is equally likely to be chosen for each gummi bear.

   (a) (3 point) What is the probability that Adam gets exactly three gummi bears?
   Solution: There are $\binom{12}{3}$ ways to select which three gummi bears Adam gets. The probability that all three of these particular gummi bears go to Adam is $\left(\frac{1}{3}\right)^3$ and the probability that the other nine gummi bears each go to Beth or Charlie is $\left(\frac{2}{3}\right)^9$. Therefore, the probability that Adam gets exactly three gummi bears is

$$\binom{12}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^9 \approx 0.212$$

   (b) (3 points) If you know that each person recieved at least one gummi bear, what is the probability that Adam gets exactly three gummi bears?
   Solution: This is a question of conditional probability, so using the formula for conditional probability gives

$$P(\text{Adam gets 3} \mid \text{everyone gets at least 1}) = \frac{P(\text{Adam gets 3 AND everyone gets at least 1})}{P(\text{everyone gets at least 1})}$$

   For the numerator, the probability that Adam gets three and everyone gets at least one is the number of ways in which both of these things happen, divided by the number of ways to distribute the gummi bears with no restrictions. To count the ways in which Adam gets three and everyone gets at least one, we can first pick which three gummi bears Adam will get, and then consider the number of ways to distribute the remaining 9 gummi bears among Beth and Charlie, so that they each get at least one. There are $\binom{12}{3}$ ways to select which gummi bears Adam gets. There are 2^9 ways to distribute the remaining gummi bears among Beth and Charlie, but two of these ways give all the gummi bears to the same person. Therefore, the number of ways in which Adam gets three and everyone gets at least one is $\binom{12}{3}(2^9-2)$ The number of ways to distribute the gummi bears with no restrictions is 3^{12} because each of the 12 gummi bears can go to any of the 3 people. So the numerator is $\frac{\binom{12}{3}(2^9-2)}{3^{12}}$.
   For the denominator, the probability that everyone gets at least one gummi bear is the number of ways in which everyone gets at least one gummi bear, divided by the number of ways to distribute the gummi bears with no restricitons. Using the solution to Group HW5 problem 3c to cont the number of ways in which everyone gets at least one gummi bear, we get $\frac{3^{12}-3*2^{12}+3}{3^{12}}$.
   Plugging this into the formula for conditional probability gives

$$P(\text{Adam gets 3} \mid \text{everyone gets at least 1}) = \frac{\frac{\binom{12}{3}(2^9-2)}{3^{12}}}{\frac{3^{12}-3*2^{12}+3}{3^{12}}} \approx 0.216$$

   (c) (3 points) What is the expected number of gummi bears that Adam gets?

Solution: If X is the number of gummi bears that Adam gets, we can write $X = \sum_{i=1}^{12} X_i$ where $X_i = 1$ if Adam gets gummi bear i, and 0 otherwise. Using linearity of expectation, $E(X) = \sum_{i=1}^{12} E(X_i)$, and $E(X_i) = \frac{1}{3} * 1 + \frac{2}{3} * 0$ for all i, since the probability of Adam getting any particular gummi bear is $\frac{1}{3}$. Therefore,

$$E(X) = \sum_{i=1}^{12} \frac{1}{3} = 4.$$

Another Solution: We can compute the expected value using the definition of expected value. Using the same argument as in part a, but for a general k, the probability that Adam gets k gummi bears is $\binom{12}{k}\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{12-k}$. Therefore the definition of expected value gives

$$\sum_{k=0}^{12}\binom{12}{k}\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{12-k} * k = 4.$$

Yet Another Solution: We do not expect Adam to have any more or less gummi bears than Beth or Charlie, since they are each equally likely to get each gummi bear, so there is no bias towards any one person. Since we know that the total number of gummi bears is 12, this means we expect Adam, Beth, and Charlie to each get 4.

(d) (3 points) If you know that each person recieved at least one gummi bear, what is the expected number of gummi bears that Adam gets?

Solution: We know Adam gets at least one gummi bear, and we also know that there are two gummi bears that go to Beth and Charlie that Adam does not get. Of the remaning 9 gummi bears, the expected number of those that Adam gets is $\sum_{i=1}^{9} \frac{1}{3} = 3$ by linearity of expectation (see part c). Therefore the expected total number of gummi bears that Adam gets is $1 + 3 = 4$.

Another Solution: We can compute the expected value using the definition of expected value. Using the same argument as in part b, but for a general k, the probability that Adam gets k gummi bears is $\frac{\binom{12}{k}(12^{12-k} - 2)}{3^{12} - 3 * 2^{12} + 3}$. Notice that for each person to get at least one gummi bear, Adam must get at least one gummi bear, but no more than ten gummi bears. Therefore the definition of expected value gives

$$\sum_{k=1}^{10} \frac{\binom{12}{k}(12^{12-k} - 2)}{3^{12} - 3 * 2^{12} + 3} * k = 4.$$

Yet Another Solution: We do not expect Adam to have any more or less gummi bears than Beth or Charlie, since they are each equally likely to get each gummi bear, so there is no bias towards any one person. This is true even when we know that they all get at least one gummi bear. Since we know that the total number of gummi bears is 12, this means we expect Adam, Beth, and Charlie to each get 4.

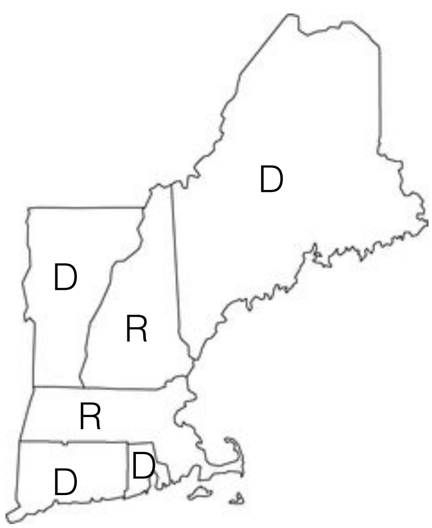
4. (a) (6 points) Suppose there are n people assigned to m different tasks. Assume that each person is randomly assigned a task and that for each person, all tasks are equally likely. Use linearity of expectation to find the expected number of people working on a task alone.

Solution: Let X be the random variable that represents the number of people that are working on a task alone. The solution is to compute the expected value $E(X)$. Suppose the people are numbered $1, 2, \ldots, n$. Let X_i be the random variable such that $X_i = 1$ if person i is alone and $X_i = 0$ otherwise. It follows that $X = \sum_{i=1}^{n} X_i$. Therefore $E(X) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$.
What is the probability that person i is alone? Person i is assigned a certain task and for each other person there is a $\frac{m-1}{m}$ chance that they are doing a different task than person i. So the chance that all $n-1$ other people are doing a different task than person i is $P(X_i = 1) = \left(\frac{m-1}{m}\right)^{n-1}$. So we have that the expected value of each X_i is $E(X_i) = \left(\frac{m-1}{m}\right)^{n-1}$.

$$E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \left(\frac{m-1}{m}\right)^{n-1} = n\left(\frac{m-1}{m}\right)^{n-1}.$$

(b) (6 points) Suppose each of the 50 states of the USA votes either Democrat or Republican in the next election. Assume that each state votes Democrat with a 50% probability, and Republican with a 50% probability. If there are 109 borders between states, use linearity of expectation to find the expected number of borders that separate a Democratic voting state from a Republican voting state.

For example, the following map shows 6 states with 7 borders between states. There are 5 borders that separate a Democratic voting state from a Republican voting state.



Solution: Let X be the random variable that represents the number of borders that separate a Democratic voting state from a Republican voting state. Suppose each border is numbered $1, 2, \ldots, 109$. Then let X_i be the random variable such that $X_i = 1$ if border i separates a Democratic state from a Republican State and $X_i = 0$ otherwise. For each border, the possibilities of the states that it separates is $\{(D, D), (D, R), (R, D), (R, R)\}$. And since each state votes Democratic or Republican with equal probability, each of these possibilities is equally likely. Therefore $P(X_i = 1) = 2/4 = 1/2$ and $P(X_i = 0) = 2/4 = 1/2$ for each border i. Therefore the expected value $E(X_i) = (1)(1/2) + (0)(1/2) = 1/2$ for each border i.

It follows that $X = \sum_{i=1}^{109} X_i$ and that $E(X) = E\left(\sum_{i=1}^{109} X_i\right) = \sum_{i=1}^{109} E(X_i)$. Therefore, $E(X) = \sum_{i=1}^{109} E(X_i) = \sum_{i=1}^{109} 1/2 = 109/2 = 54.5$.

5. Suppose we flip a fair coin n times where $n > 2$ and consider the sample space of all possible coin toss sequences.

   - Let A be the event that the first flip is Tails.
   - Let B be the event that the first and second flips are the same.
   - Let X be the random variable that counts the number of Tails that appear.
   - Let Y be the random variable that counts the number of Heads that appear minus the number of Tails that appear.

   For each pair of quantities, decide which of them is greater, or if they are both the same. Justify your answer in words.

   (a) (2 points) $P(A|B)$ and $P(A)$

   Solution: $P(A) = P(A|B)$ because knowing that the first and second flips are the same does not influence the chance that the first flip is tails. Also, we can compute $P(A) = 1/2$ and $P(A|B) = P(A \cap B)/P(B) = \frac{1/4}{1/2} = 1/2$.

   (b) (2 points) $E(Y|A)$ and $E(Y)$

   Solution: $E(Y|A) < E(Y)$ because given that A has occurred, we know that at least one of the tosses is a tails so this will decrease the expected value of the number of heads minus the number of tails.

   (c) (2 points) $V(X + Y)$ and $V(X)$

   Solution: $V(X + Y) = V(X)$ because $X + Y$ counts the number of heads and Y counts the number of tails. Each toss is equally likely so the variances are the same.

   (d) (2 points) $V(2X + Y)$ and $V(Y)$

   Solution: $V(2X + Y) < V(Y)$ because $2X + Y$ will always be the number of tosses so $V(2X + Y) = 0$ and the value of Y changes so $V(Y) > 0$.

6. (a) (3 points) Use the definition of expected value to calculate the expected number of edges in a randomly selected simple undirected graph on n labeled vertices, where we assume each simple undirected graph is equally likely to be selected.

Solution: The number of simple undirected graphs on n labeled vertices is $2^{\binom{n}{2}}$ because for each of the $\binom{n}{2}$ pairs of vertices, they either are connected with an edge or not connected with an edge (2 options.)

The number of simple undirected graphs on n labeled vertices with exactly k edges is $\binom{\binom{n}{2}}{k}$ because there are $\binom{n}{2}$ possible edges, and a graph is determined by a choice of which k of them to include.

Therefore, since each graph is equally likely to be selected, the probability of a randomly selected graph having k edges is $\frac{\binom{\binom{n}{2}}{k}}{2^{\binom{n}{2}}}$. The definition of expected value says that the expected number of edges in a randomly selected simple undirected graph on n labeled vertices is

$$\sum_{k=0}^{\binom{n}{2}} \frac{\binom{\binom{n}{2}}{k}}{2^{\binom{n}{2}}} * k.$$

(b) (3 points) Use linearity of expecation to calculate the expected number of edges in a randomly selected simple undirected graph on n labeled vertices, where we assume each simple undirected graph is equally likely to be selected.

Solution: For any simple undirected graph, let X be the number of edges in the graph. For each of the $\binom{n}{2}$ possible edges, order them in some way, and let $X_i = 1$ if edge i is included in the graph, and 0 otherwise. Then $X = \sum_{i=1}^{\binom{n}{2}} X_i$, and applying linearity of expectation gives $E(X) = \sum_{i=1}^{\binom{n}{2}} E(X_i)$. Since every graph is equally likely to be selected, and half of the graphs include any particular edge, we know the probability that $X_i = 1$ is $\frac{1}{2}$ for any i. Therefore $E(X_i) = \frac{1}{2} * 1 + \frac{1}{2} * 0 = \frac{1}{2}$. Then

$$E(X) = \sum_{i=1}^{\binom{n}{2}} \frac{1}{2} = \frac{\binom{n}{2}}{2}.$$

(c) (3 points) Suppose we generate a simple undirected graph on n labeled vertices by randomly deciding, for each possible edge, whether or not it should appear in the graph. Suppose that each possible edge has a probability p of appearing. Use the definition of expected value to calculate the expected number of edges in a graph generated in this way.

Solution: The number of edges in a simple undirected graph can range from 0 to $\binom{n}{2}$. The number of graphs having k edges is $\binom{\binom{n}{2}}{k}$. The probability of each such graph is $p^k(1-p)^{\binom{n}{2}-k}$ since each of the k edges appears with probability p, and each of the other edges does not appear with probability $1-p$. Therefore, the probability of a graph having exactly k eges is

$$\binom{\binom{n}{2}}{k} p^k (1-p)^{\binom{n}{2}-k}.$$

The definition of expected value says the expected number of edges is

$$\sum_{k=0}^{\binom{n}{2}} \binom{\binom{n}{2}}{k} p^k (1-p)^{\binom{n}{2}-k} * k.$$

(d) (3 points) Use linearity of expectation to calculate the expected number of edges in a randomly generated simple undirected graph on n labeled vertices, where we assume each possible edge appears with probability p.

Solution: For any simple undirected graph, let X be the number of edges in the graph. For each of the $\binom{n}{2}$ possible edges, order them in some way, and let $X_i = 1$ if edge i is included in the graph, and 0 otherwise. Then $X = \sum_{i=1}^{\binom{n}{2}} X_i$, and applying linearity of expectation gives $E(X) = \sum_{i=1}^{\binom{n}{2}} E(X_i)$. Since each edge i appears with probability p, $E(X_i) = p * 1 + (1-p) * 0 = p$. Then

$$E(X) = \sum_{i=1}^{\binom{n}{2}} p = \binom{n}{2} * p.$$

This makes sense because if each edge appears with probability p, we would expect any graph to have proportion p of all possible edges.

(e) (3 points) Use linearity of expectation to calculate the expected number of triangles in a randomly generated simple undirected graph on n labeled vertices, where we assume each possible edge appears with probability p. A triangle in a graph is formed by three vertices, all of which are connected by an edge.

Solution: For any simple undirected graph, let X be the number of triangles in the graph. The maximum number of possible triangles is $\binom{n}{3}$ since a triangle is determined by a set of three vertices. Order these possible triangles in some way. For each of the $\binom{n}{3}$ possible triangles, let $X_i = 1$ if triangle i is included in the graph, and 0 otherwise. Then $X = \sum_{i=1}^{\binom{n}{3}} X_i$, and applying linearity of expectation gives $E(X) = \sum_{i=1}^{\binom{n}{3}} E(X_i)$. Since each edge i appears with probability p, the probability of any particular triangle appearing in the graph is p^3, since all three edges must be present in a triangle. Then $E(X_i) = p^3 * 1 + (1-p^3) * 0 = p^3$. Then

$$E(X) = \sum_{i=1}^{\binom{n}{3}} p^3 = \binom{n}{3} * p^3.$$

7. (8 points) Tesla is planning to introduce a new self-driving car. The company commissions a marketing report for each new car that predicts either the success or the failure of the car. Of the new cars introduced by the company so far, 65% have been successes. Furthermore, 70% of their successful cars were predicted to be successes, while 40% of failed cars were predicted to be successes. Find the probability that this new self-driving car will be successful if its success has been predicted.

   Solution: The problem is asking to find the probability that a car will be successful if its success has been predicted. Let E be the event that the car's success has been predicted and let F be the event that the car will be a success. Then we want to compute $P(F|E)$. Using Bayes' theorem, we have that

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|\bar{F})P(\bar{F})}.$$

   According to the problem, $P(F) = 0.65$, $P(\bar{F}) = 0.35$, $P(E|F) = 0.70$, $P(E|\bar{F}) = 0.40$. So

$$P(F|E) = \frac{(0.7)(0.65)}{(0.7)(0.65) + (0.4)(0.35)} \approx 0.76.$$