MATH 241 Fall 2021 Homework #4

Professor Adam Kapelner

Due by email 11:59PM November 14, 2021

(this document last updated Wednesday 3rd November, 2021 at 6:54pm)

Instructions and Philosophy

The path to success in this class is to do many problems. Unlike other courses, exclusively doing reading(s) will not help. Coming to lecture is akin to watching workout videos; thinking about and solving problems on your own is the actual "working out". Feel free to "work out" with others; **I want you to work on this in groups.**

Reading is still *required*. For this homework set, read the sections in Ross about about discrete random variables especially the Bernoulli, the Binomial, the Hypergeometric, the geometric, the uniform. Then read about the concepts of expectation and variance and the formulas associated with these two concepts.

The problems below are color coded: green problems are considered *easy* and marked "[easy]"; yellow problems are considered *intermediate* and marked "[harder]", red problems are considered *difficult* and marked "[difficult]" and purple problems are extra credit. The *easy* problems are intended to be "giveaways" if you went to class. Do as much as you can of the others; I expect you to at least attempt the *difficult* problems.

This homework is worth 100 points but the point distribution will not be determined until after the due date. See syllabus for the policy on late homework.

Up to [see syllabus] points are given as a bonus if the homework is typed using LaTeX. Links to instaling LaTeX and program for compiling LaTeX is found on the syllabus. You are encouraged to use `overleaf.com`. If you are handing in homework this way, read the comments in the code; there are two lines to comment out and you should replace my name with yours and write your section. If you are asked to make drawings, you can take a picture of your handwritten drawing and insert them as figures or leave space using the "\vspace" command and draw them in after printing or attach them stapled.

The document is available with spaces for you to write your answers. If not using LaTeX, print this document and write in your answers. I do not accept homeworks not on this printout. Keep this first page printed for your records. Write your name below.

NAME: ____

Problem 1

We will continue our look at the bernoulli-binomial-hypergeometric rv's.

(a) [difficult] Let $X \sim \text{Hypergeometric}\,(n,\, K,\, N)$ and let $p = K/N$. Show that the PMF of the rv when $N \to \infty$ is $\text{Binomial}\,(n,\, p)$.

(b) [harder] Let $X \sim \text{Hypergeometric}\,(n,\, K,\, N)$. You can think of this rv as the sum of $X_1, \ldots, X_n$ where each X_i is either 1 or 0 indicating whether or not the ith draw is one of the K special items or not. How are each of these X_i distributed?

(c) [harder] Are these X_i identically distributed? Yes/No

(d) [harder] Are these X_i independent? Yes/No and prove your answer.

(e) [harder] The human mouth has 32 teeth. If the probability of a cavity at some point in a lifetime is 5%, is it possible to calculate the probability of 7 cavities during a lifetime using a binomial r.v. model $X \sim \text{Binomial}(32, 5\%)$ and computing $\mathbb{P}(X = 7)$? Why or why not?

Problem 2

We will be investigating r.v.'s by imagining a trip the grocery store to buy ingredients for guacamole.

(a) [easy] You buy *one* avocado at the grocery store which is good (c.f. a bad avocado which may have brown inside because it's partially rotten). Call this probability of good p. Model the number of *good* avocados you have using a random variable. All you need to write is $X \sim$ something. You do not need to write the PMF, draw the PMF, draw the CDF, etc.

(b) [difficult] You take one avocado, cut it open and see if it's rotten. You keep doing this until you see a *rotten* avocado. Model the number of avocados you cut open using a r.v. Call this r.v. X. Be careful between "rotten" and "good".

(c) [easy] Write the PMF for the r.v. you created in (b).

(d) [easy] Write the support for the r.v. you created in (b).

(e) [easy] What is the probability you stop when looking at the third avocado?

(f) [easy] Use the sigma notation for summing (e.g. $\sum_{i=1}^{5}$) to calculate the probability that you stop between 4 and 37 avocados (including 4 and including 37). Since you don't know p you cannot actually compute a numerical value for this probability. Leave it in sigma notation.

(g) [harder] Let's say at some point in your avocado shopping that you learned how to detect rotten avocadoes by using the "squeeze test" and you used this learning to select new avocados. What assumption(s) would be violated?

(h) [harder] Grocery stores usually put the old avocados on the top of the avocado basked and thus the new avocados on the bottom. If your strategy was just to pick the "top" avocado each time, what assumption(s) would be violated?

(i) [harder] Is the process of finding a bad avocado "memoryless"? Yes / no. Define memoryless and explain your answer.

Problem 3

This is the fun part of the homework. You're going to do some data-generating / rv-realizing experiments. Let's make some magic!

(a) [easy] Grab a cup and 8 pennies (or nickels, or dimes, etc). Use a magic marker to mark four of them (front and back). If you shake the cup and pull out three coins, let X be the r.v. for how many marked coins you pull out? How is X distributed? Write "$X \sim$ something" below.

(b) [easy] Using as fact that $\mathbb{E}[X] = n\frac{K}{N}$ when $X \sim \text{Hypergeometric}(n, K, N)$ which we proved in class, calculate $\mathbb{E}[X]$ for the r.v. you constructed in part (a).

(c) [easy] Shake the cup and take out 3 coins. How many were marked? Repeat this five times. Record your data below. That is, just write down the five numbers separated by commas.

(d) [easy] Find $\bar{x}$ from the data you recorded in part (c).

(e) [easy] Is $\bar{x} \approx \mathbb{E}[X]$? If not, what could you change in the experiment to make $\bar{x}$ closer to $\mathbb{E}[X]$?

(f) [easy] Now forget that the coins are marked. If you shake the cup and flip all 8 coins, let X be the r.v. for how many heads are flipped. How is X distributed? Write "$X \sim$ something" below.

(g) [easy] Using the fact we proved in class that $\mathbb{E}[X] = np$ when $X \sim \text{Binomial}(n, p)$, calculate $\mathbb{E}[X]$ for the r.v. you constructed in part (f).

(h) [easy] Shake the cup and count the number of heads. Repeat this five times. Record your data below.

(i) [easy] Find $\bar{x}$ from the data you recorded in part (h).

(j) [easy] Is $\bar{x} \approx \mathbb{E}[X]$? If not, what could you change in the experiment to make $\bar{x}$ closer to $\mathbb{E}[X]$?

(k) [easy] Now imagine one coin in the cup and success is defined as getting a head. Further imagine that you don't stop flipping this coin until you get a head. Let X be the r.v. for how many flips you make. How is X distributed? Write "$X \sim$ something" below.

(l) [easy] Using the fact we proved in class that $\mathbb{E}[X] = 1/p$ when $X \sim \text{Geometric}(p)$, calculate $\mathbb{E}[X]$ for the r.v. you constructed in part (k).

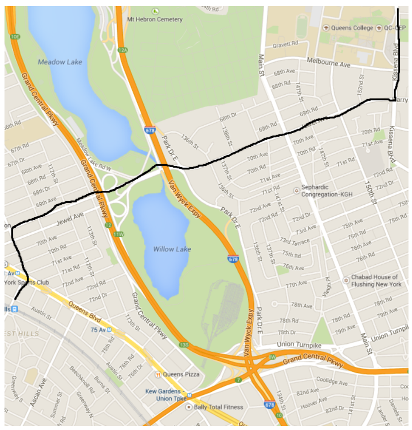
(m) [easy] Flip until you get a head. Repeat this five times. Record your data below.

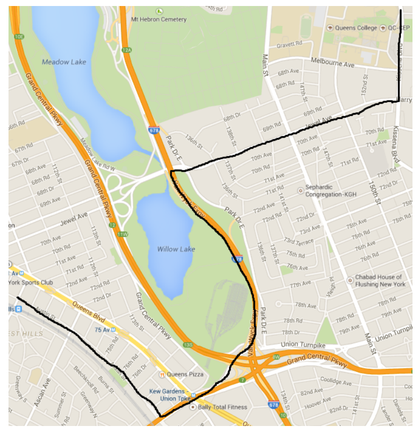
(n) [easy] Find $\bar{x}$ from the data you recorded in part (m).

(o) [harder] Is $\bar{x} \approx \mathbb{E}[X]$? If not, what could you change in the experiment to make $\bar{x}$ closer to $\mathbb{E}[X]$?

Problem 4

Consider a taxi ride from Forest Hills to QC. For the purposes of this exercise, assume there are only two routes. This is close to realistic. There is the "Van Wyck" (outlined in black on the right below) and "Jewel Ave" which is the Q64 bus route (outlined in black on the left below). The only determinant of route selection is whether or not there is traffic on the Van Wyck. If there is traffic, I take Jewel Ave route; if not, I take the Van Wyck route. The probability of traffic on the Van Wyck is 30%. The Jewel Ave route is 2.3 miles and takes 13 min and the Van Wyck route is 8 min and is 3.6 miles.





(a) [easy] Let W be the r.v. which models the time I travel in the taxi. What is its distribution? Use the notation we used in class.

(b) [easy] What is $\text{Supp}\,[W]$?

(c) [easy] Compute $\mathbb{E}\,[W]$ from the definition of expectation.

(d) [easy] Write a sentence that synthesizes what part (c) means.

(e) [easy] Let D be the r.v. which models the distance I travel in the taxi. What is its distribution? Use the notation we used in class.

(f) [easy] Compute $\mathbb{E}[D]$.

(g) [difficult] Are the r.v.'s W and D dependent? Justify your answer *in English.*

(h) [easy] Write a sentence that synthesizes what part (f) means.

(i) [easy] The taxi charges \$0.35\min. Let M be the r.v. which is what I pay for time on my trip home. Find the distribution of M.

(j) [easy] Write M as a function of W.

(k) [easy] Calculate $\mathbb{E}[M]$ based on the formula we learned in class about expectations of r.v.'s scaled by a constant.

(l) [easy] The taxi charges \$1.75\mi of distance covered. Let L be the r.v. which is what I pay for mileage on my trip home. Find the distribution of L.

(m) [easy] Write L as a function of D.

(n) [easy] Calculate $\mathbb{E}[L]$ based on the formula we learned in class about expectations of r.v.'s scaled by a constant.

(o) [easy] The taxi also includes a base fare of \$2.55. Let B be the r.v. which models the total bill for my taxi ride. Write B as a function of W and D.

(p) [harder] We didn't really cover this in class, but you should be able to do it. W and D are one-to-one so the scaled W and scaled D sum is really one r.v. Find $\mathbb{E}[B]$ based also on the formula we learned in class about the expectation of a r.v. with a constant added.

Problem 5

Imagine rolling two fair dice (no sorcery). Let X_1 be the r.v. corresponding to the first die and let X_2 be the r.v. corresponding to the second die. Let the outcome results be \$1 if you roll a 1, \$2 if you roll a 2, ..., and \$6 if you roll a six.

(a) [easy] What brand name r.v. is X_1 distributed as? Write $X \sim$ something and make sure the parameters are correct.

(b) [easy] Does $X_1 \overset{d}{=} X_2$? Yes or no is fine.

(c) [easy] Are X_1 and X_2 independent? Yes or no is fine.

(d) [easy] Compute $\mathbb{E}[X_1]$ from first principles.

(e) [easy] Compute $\mathbb{V}\text{ar}[X_2]$ from first principles.

(f) [easy] The standard deviation is also called "standard error" and it sometimes denoted "SE." Use your answer in (e) to find $\mathbb{SE}[X_i]$ for $i \in \{1, 2\}$. Please just use the square root and do not rederive the variance again from scratch.

(g) [easy] Draw the PMF for X_i for $i \in \{1, 2\}$ and mark $\mathbb{E}[X_i]$ and $\mathbb{SE}[X_i]$ on the graph similar to how we did in class.

(h) [easy] Imagine the game where you just double the winnings of a single roll. This would be equivalent to just multiplying the r.v. by a scale factor of 2. Calculate $\mathbb{E}[2X_i]$, $\mathbb{V}\text{ar}[2X_i]$ and $\mathbb{S}\text{E}[2X_i]$ from the formulas we learned in class.

(i) [easy] Draw the PMF for $2X_i$ for $i \in \{1, 2\}$ and mark $\mathbb{E}[2X_i]$ and $\mathbb{S}\text{E}[2X_i]$ that you calculated in (h) on the graph.

(j) [difficult] Draw the PMF for $X_1 + X_2$. This involves taking a convolution. Since convolution won't be on the midterm or final, I'm going to give a hint. There is 1 way to get 2 or 12, 2 ways to get 3 or 11, 3 ways to get 4 or 10, 4 ways to get 5 or 9, 5 ways to get 6 or 8 and 6 ways to get 7.

(k) [easy] Calculate $\mathbb{E}[X_1 + X_2]$, $\mathbb{V}\text{ar}[X_1 + X_2]$ and $\mathbb{SE}[X_1 + X_2]$ from the formulas we learned in class. Do not use the PMF from the last question; use the formulas from class.

(l) [easy] Imagine the general case of $X_1, \ldots, X_n \stackrel{iid}{\sim}$ with mean μ and variance σ^2. Define $\bar{X}$ as we did in class. Redo the derivation of $\mathbb{E}\left[\bar{X}\right] = \mu$.

(m) [easy] Imagine the general case of $X_1, \ldots, X_n \stackrel{iid}{\sim}$ with mean μ and variance σ^2. Define $\bar{X}$ as we did in class. Redo the derivation of $\mathbb{SE}\left[\bar{X}\right] = \sigma/\sqrt{n}$.

(n) [difficult] Now you have the choice between game A — where you roll n times and average the winnings (*i.e.* you collect $\bar{X}_n$ dollars at the end) or game B — where you roll one die and collect the amount you make on just one roll. Use your answers to the relevant previous questions (I won't tell you which ones explicitly) to explain why you would choose game A over B or vice versa. I want multiple sentences *in English*. You must convince me you understand the tradeoff that game A and B are making.

(o) [harder] Let Z be the standardized r.v. for $\bar{X}_n$. Standardization of a r.v. X is defined as subtracting its mean and dividing by its standard error. For $\bar{X}$ this would be:

$$Z := \frac{\bar{X} - \mathbb{E}\left[\bar{X}\right]}{\mathbb{SE}\left[\bar{X}\right]} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Prove from the formulas in class that $\mathbb{E}[Z] = 0$ and $\mathbb{V}\text{ar}[Z] = \mathbb{SE}[Z] = 1$. Hint: use those two rules about $\mathbb{V}\text{ar}[aX]$ and $\mathbb{V}\text{ar}[X + c]$ you just rederived

(p) [easy] Why is do you think "standardization" is an appropriate term for what we did above?

(q) [harder] Returning to X, the dice game in the beginning of the problem (the outcome results being \$1 if you roll a 1, \$2 if you roll a 2, ..., and \$6 if you roll a six), you calculated variance using the definition $\mathbb{V}\text{ar}[X] := \mathbb{E}\left[(X-\mu)^2\right]$ assuming the classic squared error loss: $e(x, \mu) := (x-\mu)^2$. Imagine we defined a new variance metric using absolute loss, $\widetilde{e(x, \mu)} := |x - \mu|$. We'll denote this "new variance" with a big squiggly symbol, $\widetilde{\mathbb{V}\text{ar}}[X] := \mathbb{E}[|X - \mu|]$ just to make sure you don't confuse it with the standard definition of $\mathbb{V}\text{ar}[X]$. Calculate $\widetilde{\mathbb{V}\text{ar}}[X]$ and include units.

Problem 6

More simple r.v. practice.

(a) [easy] You know that T_n is the sum of $n \stackrel{iid}{\sim}$ bernoulli r.v.s with parameter p. Show that $\mathbb{V}\text{ar}\left[T_n\right]$ can be easily derived using the variance-sum formula we learned in class.

(b) [difficult] Show for any two r.v.'s X and Y which are independent that $\mathbb{V}\text{ar}\left[X \times Y\right] = \mu_X^2 \sigma_Y^2 + \mu_Y^2 \sigma_X^2 + \sigma_X^2 \sigma_Y^2$. Remember, two r.v.'s multiplied together is a new r.v., $g(X, Y)$.

(c) [easy] Let $a_1, a_2, \ldots, a_n$ be a sequence of constants. Let $X_1, \ldots, X_n$ be a sequence of r.v.'s which thereby share the same mean μ. Create a simplified expression for $\mathbb{E}\left[a_1 X_1 + \ldots + a_n X_n\right]$ as the simplest combination of symbols $a_1, a_2, \ldots, a_n$ and μ.

(d) [harder] Let $a_1, a_2, \ldots, a_n$ be a sequence of constants. Assume $X_1, \ldots, X_n$ are a sequence of $\stackrel{iid}{\sim}$ r.v.'s which thereby share the same variance σ^2. Create a simplified expression for $\mathbb{SE}\left[a_1 X_1 + \ldots + a_n X_n\right]$ as the simplest combination of symbols $a_1, a_2, \ldots, a_n$ and μ.