MATH 241 Fall 2017 Homework #4

Professor Adam Kapelner

Due 6:30PM KY604 or KY258, Thursday, October 26, 2017

(this document last updated Wednesday 18$^{\text{th}}$ October, 2017 at 12:01am)

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 section about foundations of probability in Chapter 2 and the section about conditional probability and in/dependence in Chapter 3 in Ross. Chapter references are from the 7th edition.

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 15 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, upload `hwxx.tex` and `preamble.tex`, 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.

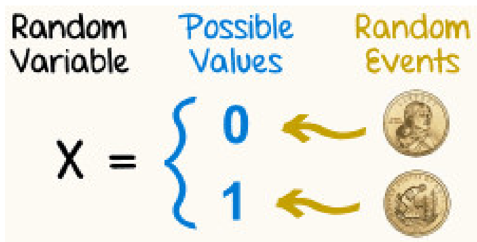
The document is available with spaces for you to write your answers. If not using LaTeX, print this document and write in your answers. I do not accept homeworks not on this printout. Keep this first page printed for your records. Write your name and section below (A, B or C).

NAME: ______________________________ SECTION (A, B or C): ______

Random Variables We now begin question about the second unit of this class: r.v.'s. Anything past this point is NOT covered on Midterm 1.

Problem 1

In class we spoke about how random variables map outcomes from the sample space to a number *i.e.* $X : \Omega \to \mathbb{R}$. That is they are set functions, just like the probability function which is $\mathbb{P} : 2^{\Omega} \to [0, 1]$. We will be investigating this concept here.



(a) [easy] Here is a way to produce $X \sim \text{Bernoulli}\left(\frac{1}{2}\right)$ using the Ω from a roll of a die. Map outcomes 1,2,3 to 0 and outcomes 4,5,6 to 1. This works because

$$\mathbb{P}(X = 0) = \mathbb{P}(\{\omega : X(\omega) = 0\}) = \mathbb{P}(\{1, 2, 3\}) = 1/2 \text{ and}$$
$$\mathbb{P}(X = 1) = \mathbb{P}(\{\omega : X(\omega) = 1\}) = \mathbb{P}(\{4, 5, 6\}) = 1/2.$$

Describe three other scenarios or devices that produce their own Ω's that also result in $X \sim \text{Bernoulli}\left(\frac{1}{2}\right)$. Be creative (i.e. not boring).

(b) [harder] We talked about in class how the sample space no longer needs to be considered once the random variable is described. Why? Use your answer to (a) to inspire this answer. Write it *in English* below.

(c) [difficult] Back to philosophy... Let's say X models the price difference that IBM stock moves in one day of trading. For instance, if the stock closed yesterday at \$56.24 and today it closed at \$57.24, the random variable would be \$1 for today. According to our definition of a random variable, there is a sample space with outcomes being drawn ($\omega \in \Omega$) that is "controlling" the value of X. Describe it the best you can *in English*. There are no right or wrong answers here, but your answer must be coherent and demonstrate you understand the question.

Problem 2

We will now study probability mass functions (PMF's) denoted as $p(x)$ and cumulative distribution functions (CDF's) denoted as $F(X)$ and review the r.v.'s we did in class.

(a) [easy] Draw the PMF for $X \sim \text{Bernoulli}(p)$.

(b) [easy] Draw the CDF for $X \sim \text{Uniform}(\{1, 3, 4, 9\})$.

(c) [harder] Using the r.v. from the previous question, what is $\mathbb{P}(X \in (3, 9))$? I am trying to trick you here.

(d) [difficult] In class we defined the Bernoulli r.v. as:

$$X \sim \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } 1-p \end{cases}$$

and put its PMF on the board. Write $p(x)$ for $X \sim \text{Bernoulli}\,(p)$ that is only valid for not only all values in the $\text{Supp}\,[X]$ but all values in $\mathbb{R}$. Hint: use the indicator function.

(e) [easy] What is the parameter space of X where $X \sim \text{Bernoulli}\,(p)$ and why?

(f) [difficult] Consider the PMF we discussed for $X \sim \text{Bernoulli}\left(\frac{1}{2}\right)$. Does $\int p(x)\,\mathrm{d}x = F(x) + C$ where the constant $C \in \mathbb{R}$? Explain. Think carefully about what integration really means.

(g) [difficult] How about the opposite? Consider the CDF we discussed for $X \sim \text{Bernoulli}\left(\frac{1}{2}\right)$. Does $\mathrm{d}/\mathrm{d}x[F(x)] = p(x)$? Explain. Think carefully about what differentiation really means.

Hypergeometric Distribution This is a very interesting random variable and we will cover it thoroughly between this homework and the next one.

Problem 3

The hypergeometric is sampling "without replacement." Imagine you have this bag of marbles with 37 marbles and 17 of them are black. We will define a "success" as drawing a black marble.

(a) [easy] Let's say you draw one marble. Call this r.v. X. Is it hypergeometric?

(b) [easy] The hypergeometric distribution has three parameters. What are the parameters for X?

(c) [easy] Write, but do not draw, the PDF, $p(x)$ for the r.v. X where x is the number of successes.

(d) [easy] What is the support of this r.v.?

(e) [harder] There is another variable we learned about in class with this same support. Show that X is distributed as this type of r.v. and find its parameter(s).

(f) [easy] Now imagine you draw 4 marbles without replacement. Call this r.v. X (and forget about the previous r.v. X from this question, parts a-e). How is X distributed? Use the notation in class and find its parameters.

(g) [easy] What is the support of X?

(h) [easy] Write, but do not draw, the PMF of X.

(i) [easy] Draw the PMF of X.

(j) [easy] Draw the CDF of X.

(k) [easy] What is the probability of getting 4 successes in a row? Use the PMF.

(l) [easy] Now imagine you draw 27 marbles without replacement. Call this r.v. X (and forget about the previous r.v. X). How is X distributed? Use the notation in class and find its parameters.

(m) [easy] What is the support of X?

(n) [easy] Why is $0 \notin \text{Supp}[X]$?

(o) [easy] Write, but do not draw, the PMF of X.

(p) [difficult] Find the mode of this distribution. "Mode" is defined as the most likely outcome result.

Problem 4

Generally, the hypergeometric has three parameters. We will solve for its support here under several disjoint conditions and then in class we will generalize it. Call X a hypergeometric r.v. with all its parameters free - meaning they can take on any value, so please use the notation n, K, N in your answers as we did in class. This problem is just copying from your class notes.

(a) [easy] Using the usual parameterization of the hypergeometric, describe the parameter space. You need to say what sets each of the parameters "lives" in.

(b) [easy] Write, but do not draw, the PMF of X.

(c) [easy] Let's say $n < K$ and $n < N - K$. What is the support of X in this situation?

(d) [easy] Let's say $n < K$ and $n \geq N - K$. What is the support of X in this situation?

(e) [easy] Let's say $n \geq K$ and $n < N - K$. What is the support of X in this situation?

(f) [easy] Let's say $n \geq K$ and $n \geq N - K$. What is the support of X in this situation?

(g) [harder] Find a formula using a sum for the CDF of the general hypergeometric r.v.

(h) [E.C.] Demonstrate that the sum of the PMF over the support is 1 (on a separate piece of paper).

Problem 5

We will now look at the binomial in general.

(a) [harder] Show using the definition of equals in distribution that $X_1 \stackrel{d}{=} X_2$ if $X_1 \sim \text{Bernoulli}\,(p)$ and $X_2 \sim \text{Binomial}\,(1,\, p)$.

(b) [harder] Imagine an infinite bag where 47% of the balls are "successes". If I draw 87 balls, what is the probability I get 29 success balls?

(c) [harder] Imagine I have a bag with 300 balls where 141 of the balls are "successes". If I draw 87 balls *with replacement*, what is the probability I get 29 success balls?

(d) [harder] Why is your answers to (c) and (d) the same?

(e) [easy] Let $X_1, \ldots, X_n \stackrel{iid}{\sim} \text{Bernoulli}\,(p)$. Give a real-life example of this situation

(f) [easy] Let $T_n = X_1 + \ldots + X_n$ where $X_1, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}\,(p)$. How is T_n distributed?

(g) [difficult] Let $X_1, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}\,(p)$ and $T_n = \sum_{i=1}^n X_i$. Derive the distribution of T_n from first principles just like we did in the notes.

Problem 6

Imagine two Bernoulli r.v.'s X_1 and X_2 which model two fair coin flips where Heads is mapped to 1 and tails is mapped to 0. The probability of heads is $1/2$.

(a) [easy] Given no other information, explain using the definition of r.v. independence why these two r.v.'s are independent.

(b) [easy] Given no other information, explain using the definition of equality in distribution why $X_1 \overset{d}{=} X_2$.

(c) [easy] Are $X_1, X_2 \stackrel{iid}{\sim} \text{Bernoulli}\,(p)$?

(d) [harder] Now imagine these two coins were linked using some sort of sorcery. They are flipped at the same time but are guaranteed to flip the same way. That is, if the first coin goes heads, the second coin must go heads (and if the first coin goes tails, the second coin must go tails).

Explain using the definition of r.v. independence why these two r.v.'s are *dependent*.

(e) [difficult] Using the same two sorcery-controlled coins, explain using the definition of equality in distribution why or why not $X_1 \stackrel{d}{=} X_2$.

(f) [easy] Let $T_2 = X_1 + X_2$. Is $T_2 \sim \text{Binomial}\left(2, \frac{1}{2}\right)$? Why or why not?