

MATH 241 Spring 2015 Homework #6

Professor Adam Kapelner

Due 5PM, Tuesday, Mar 24, 2015

(this document last updated Tuesday 17th March, 2015 at 11:10pm)

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*. But for this homework set, read the relevant parts of Chapter 4 on the bernoulli, hypergeometric, binomial, geometric and negative binomial r.v.’s.

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]”, **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. If the problem asks you for a computation, **round to two or three decimals (do not answer in an exact fraction).**

This homework is worth 100 points but the point distribution will not be determined until after the due date. Late homework will be penalized 10 points per day.

Between 1–15 points are arbitrarily given as a bonus (conditional on quality) if the homework is typed using L^AT_EX (15 points only if it is perfect L^AT_EX). Links to installing L^AT_EX and program for compiling L^AT_EX are found on the syllabus. You are encouraged to use overleaf.com (make sure you upload both the hwxx.tex and the preamble.tex file). 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 as a figure 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 L^AT_EX, print this document and write in your answers. **Handing it in without this printout is NOT ACCEPTABLE.** You will have to redo it.

Keep this page printed for your records. Write your name and section below where section A is if you’re registered for the 9:15AM–10:30AM lecture and section B is if you’re in the 12:15PM–1:30PM lecture.

NAME: _____ SECTION (A or B): _____

Independence and equality of distribution of r.v.'s Since we haven't covered much else, this majority of this assignment will be about this distribution.

Problem 1

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 $X_1 \stackrel{d}{=} X_2$.
- (c) [easy] Are $X_1, X_2 \stackrel{iid}{\sim}$ 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) [harder] 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] Are $X_1, X_2 \stackrel{iid}{\sim}$ Bernoulli(p) if they are modeled by these two sorcery-controlled coins?

The Binomial All binomial, all the time...

Problem 2

We will look at hypergeometric distributions with large N . If N is really large, sampling without replacement can be approximated by sampling with replacement. In the limit, it is sampling with replacement.

Parts (a-e) are from HW #5 so I've filled in the answers already for you. Other answers are filled in as well

- (a) [easy] We will now begin deriving the binomial in pieces. Parameterize a hypergeometric by setting $K = pN$. What is the parameter space for p ?

$$p \in \left\{ \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N} \right\}$$

- (b) [easy] Write the PMF $p(x)$ for this r.v. using the p parameterization using x as the free variable.

$$p(x) = \frac{\binom{pN}{x} \binom{(1-p)N}{n-x}}{\binom{N}{n}}$$

- (c) [easy] What limit do we take and why are we taking this limit?

We are taking the $\lim_{N \rightarrow \infty}$ because we are trying to derive the PMF of sampling n successes or failures with replacement.

- (d) [easy] Rewrite the PMF without choose notation using only factorials and simplify the fraction by moving the factorial terms from denominator, $\binom{N}{n}$, to the numerator.

$$p(x) = \lim_{N \rightarrow \infty} \frac{(pN)!}{x!(pN-x)!} \frac{((1-p)N)!}{(n-x)!((1-p)N-n+x)!} \frac{n!(N-n)!}{N!}$$

- (e) [easy] Which three terms can you factor out from the limit expression? Show that they are equivalent to $\binom{n}{x}$.

Since we are taking the $\lim_{N \rightarrow \infty}$, p and n and x are all constants. Thus, we can pull out all terms that are not functions of N . We can pull out the $x!$, $(n-x)!$ and $n!$ (see the terms in part d colored purple). If we pull them out, we get:

$$\frac{1}{x!} \times \frac{1}{(n-x)!} \times n! = \binom{n}{x}$$

- (f) [difficult] Within the limit, you now have three ratios. Write these ratios by canceling out the common terms. For instance $10!/6! = 10 \times 9 \times 8 \times 7$ and $6!/10! = 1/(10 \times 9 \times 8 \times 7)$. This is difficult because you have to get the indexing right.

- (g) [harder] How many terms are in the numerator? How many terms are in the denominator.

There are x terms in the first piece of the numerator and $n-x$ terms in the second piece of the denominator for a total of n terms. Then there are n terms in the denominator.

- (h) [harder] Reason in English that the denominator looks like a bunch of $N - c_i$ where the c_i 's are all constants which are negligible as $N \rightarrow \infty$.

- (i) [harder] Reason in English that the numerator looks like a bunch of $Np - c_i$ where the c_i 's are all constants which are negligible as $N \rightarrow \infty$ as well as a bunch of $(1 - p)N - c_i$ where the c_i 's are all constants which are negligible as $N \rightarrow \infty$.
- (j) [harder] Match each $Np - c_i$ term in the numerator to one $N - c_i$ term in the denominator and take the limit of each one individually. Show that you wind up with $p \times p \times \dots$ for a total of x times, i.e. p^x .
- (k) [harder] Match each $(1 - p)N - c_i$ term in the numerator to one $N - c_i$ term in the denominator and take the limit of each one individually. Show that you wind up with $(1 - p) \times (1 - p) \times \dots$ for a total of $n - x$ times, i.e. $(1 - p)^{n-x}$.

- (l) [easy] Using your answers from the previous parts to write the binomial r.v.'s PMF.

$$p(x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

- (m) [E.C.] Imagine you are drawing n successes with and without replacement. Derive a function of N and n which gives this percentage difference for N and n generally when the number of successes $K = \frac{1}{2}N$.

Problem 3

We will now look at the binomial in general.

- (a) [easy] Show that success and failure is arbitrary by letting the number of successes x equal the number of failures $n - x$ and the probability of success p equals the probability of failure $1 - p$ using the PMF of the binomial distribution. This is similar to last homework where we illustrated the same fact for the hypergeometric distribution.

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

- (c) [difficult] [OPTIONAL — this question will not be graded] Let $X_1, X_2, X_3, X_4 \stackrel{iid}{\sim} \text{Bernoulli}(p)$ and $T_4 = \sum_{i=1}^4 X_i$. Use a tree structure like we did in class to show that $\mathbb{P}(T_4 = 2) = \binom{4}{2} p^2 (1 - p)^2$.

(d) [easy] In (c) explain why you need the $\binom{4}{2}$ term *in English*.

(e) [easy] In (c) explain what the function of the $p^2(1-p)^2$ term is *in English*.

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

Problem 4

Imagine you are flipping the same bundle of coins from the midterm. The probability of the coin bundle landing on its side is $\mathbb{P}(S) = 1/11$. Let's call landing on its side a "success."



(a) [easy] I flip the coin bundle once. Model a success as a "1." Show that the r.v. modeling this event outcome is Bernoulli and define its parameter. Write " $X \sim$ " something below.

- (b) [easy] Let's say we flip 10 times. What is the probability that we get one (and only one) success? I want to see a probability model. Write " $X \sim$ " something below. Then I want to see a probability statement. Then I want to see a computation. Answer then in decimal rounded to two digits.
- (c) [easy] Let's say we flip 10 times. What is the probability that we get 5 (and only 5) successes?
- (d) [easy] Let's say we flip 10 times. What is the probability that we get 8 (and only 8) successes?
- (e) [harder] Let's say we flip 10 times. What is the probability we get one or two successes?
- (f) [difficult] Let's say we flip 10 times. What is the probability we get 3 or less successes? That is, solve for $\mathbb{P}(X \leq 3) = F(3)$.

Problem 5

Imagine you are playing roulette again this time in America. The probability of winning a bet on black is $18/38$. Call this a “success.”



- (a) [easy] Let's say we spin 15 times. What is the probability that we get 10 successes?
- (b) [harder] Let's say we spin 30 times. Write a summation expression for getting 15 or more successes. Do not compute the answer explicitly.
- (c) [difficult] Preview of statistics. You are now the casino floor manager for roulette. You witness 40 spins and it comes out black 18 times. Is this a “weird” or “unexpected” outcome? Explain using a calculation and a few sentences *in English*.

(d) [difficult] You witness 40 spins and it comes out black 38 times. Is this a “weird” or “unexpected” outcome? Explain using a calculation and a few sentences *in English*.

(e) [E.C.] You witness 40 spins. How many times should black occur “normally?” At which large values of number of blacks do get concerned by? At which small values of number of blacks do you get concerned by?

Problem 6

Now that we understand both the binomial and the concept of $\overset{iid}{\sim}$, we will ask some conceptual questions.

(a) [harder] Recall X_1, X_2 from problem 1(d) which were the two sorcery-controlled coins. Let $T_2 = X_1 + X_2$. Is $T_2 \sim \text{Binomial}(2, \frac{1}{2})$? Why or why not?

- (b) [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?

Negative Binomial A related distribution to the binomial. We will be doing geometric next week.

Problem 7

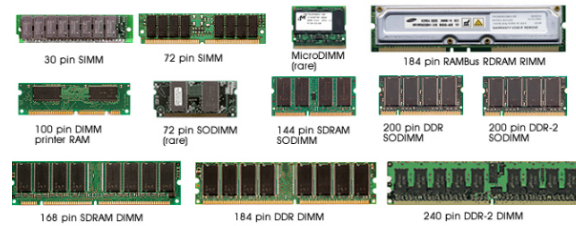
We will rederive the negative binomial PMF as we did in class. The probability of success is p and the number of successes we wish to find is r .

- (a) [easy] If we are waiting x trials to finally see exactly r successes, what does the outcome result of the last trial *need* to be?
- (b) [easy] How many trials do we witness in order to witness $r - 1$ successes not counting the last trial?
- (c) [easy] Can these $r - 1$ successes happen anywhere within these $x - 1$ trials?
- (d) [easy] If you get $r - 1$ successes in $x - 1$ trials, how many failures do you get?

- (e) [easy] How many ways is there to get $r - 1$ successes among $x - 1$ trials?
- (f) [easy] What is the probability of getting $r - 1$ successes and $x - r$ failures *in that order* if successes and failures are independent?
- (g) [easy] Use the answers in (e) and (f) to find the probability of getting $r - 1$ successes in $x - 1$ trials.
- (h) [easy] Use the answers in (g) and the probability of a final success to finally derive the full PMF of the Negative Binomial distribution.
- (i) [easy] Let $X \sim \text{NegBin}(r, p)$. What is the support of X ?
- (j) [harder] What is the parameter space of r and p ? Be careful not to allow degenerate cases.

Problem 8

You are testing RAM. The manufacturing process is near perfect. The probability of finding faulty RAM is about 1 in 300. We assume all RAM chips are independent with respect to whether they are faulty.



- (a) [easy] What is the probability you get three faulty RAM chips in a row?
- (b) [harder] What is the probability you have to investigate 100 RAM chips in order to find exactly 3 faulty chips? Compute explicitly.
- (c) [harder] What is the probability you have to investigate 500 RAM chips in order to find exactly 3 faulty chips? You can leave in choose notation and use exponents as well.
- (d) [difficult] What is the probability you have to investigate more than 500 RAM chips to see exactly 3 faulty chips? You can leave in choose notation and use exponents as well.