

Solutions to In-Class Problems Week 13, Wed.

Problem 1. The following two parts are not related. Try them, to make sure you understand the jargon of random variables, distributions, probability density functions, etc. Ask your TA if you don't understand/remember what some phrase means.

(a) Suppose X_1 , X_2 , and X_3 are three mutually independent random variables, each having the uniform distribution

$$\Pr \{X_i = k\} \text{ equal to } 1/3 \text{ for each of } k = 1, 2, 3.$$

Let M be another random variable giving the maximum of these three random variables. What is the density function of M ?

Solution. This can be hashed out by counting the possible outcomes. Alternatively, we can reason as follows:

The event $M = 1$ is the event that all three of the variables equal 1, and since they are mutually independent, we have

$$\Pr \{M = 1\} = \Pr \{X_1 = 1\} \cdot \Pr \{X_2 = 1\} \cdot \Pr \{X_3 = 1\} = \left(\frac{1}{3}\right)^3 = \frac{1}{27}.$$

To compute $\Pr \{M = 2\}$, we first compute $\Pr \{M \leq 2\}$. Now the event $[M \leq 2]$ is the event that all three of the variables is at most 2, so by mutual independence we have

$$\Pr \{M \leq 2\} = \Pr \{X_1 \leq 2\} \cdot \Pr \{X_2 \leq 2\} \cdot \Pr \{X_3 \leq 2\} = \left(\frac{2}{3}\right)^3 = \frac{8}{27}.$$

Therefore,

$$\Pr \{M = 2\} = \Pr \{M \leq 2\} - \Pr \{M = 1\} = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}.$$

Finally,

$$\Pr \{M = 3\} = 1 - \Pr \{M \leq 2\} = 1 - \frac{8}{27} = \frac{19}{27}.$$

■

(b) Suppose X , Y are two independent binomial random variables with parameters (n, p) and (m, p) , respectively. What is $\Pr \{X + Y = k\}$?

Solution. The PDF of X is the probability of tossing k Heads out of n independent flips of a coin with bias p . Likewise for Y and m flips. Since, X and Y are independent, the PDF of $X + Y$ corresponds to $n+m$ independent flips, i.e., $X+Y$ is a Binomial variable with parameters $(n+m, p)$. Hence,

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

■

Problem 2. I know everything. Seriously. So, I know everything that everybody thinks. In particular, I know who each one of the 750 billion members of the galaxy want to vote for in the upcoming elections for the Intergalactic president. I know that a fraction $p = 0.52$ of them want to vote for the current president.

You are mortal. An insignificant dot in space-time. But a quite significant dot among dots. You work closely to the current Intergalactic president and, within a week, you must answer his agonizing question: “Am I winning?” Or, in math jargon (but with the same agony): “Is $p > 1/2$?”

Your *first* idea is to ask me (I know everything). But you haven’t talked to me for a long time, so you know I won’t tell you. Your *second* idea is to call every person in the galaxy, ask them, then divide the yes’s by 750 billion. But you soon realize there is not enough time (there is a reason for representative democracy). Your *third* idea... You have no third idea! In your panic as the week is almost over, you start picking galaxy members at random, call them, and ask!

Amazingly, that’s the correct approach. But you should be careful what you are going to say to the president! Let’s see.

(a) In your first phone call, you pick one galaxy member *uniformly at random*, call, and ask whether he/she will vote for the president. What is the probability that the answer is going to be “yes”... (i) from my perspective? (ii) from your perspective? How would you model this in terms of coin flips?

Solution. From my perspective, it’s 0.52. From your perspective, it is also 0.52. The only problem is that you don’t know that, so you just call it p . Clearly, from your perspective, the first phone call is the same as flipping a coin with an *unknown* bias, which you call p (and I know is 0.52). ■

(b) In your second phone call, you again pick a galaxy member *uniformly at random*, call, and ask whether he/she will vote for the president. But wait! When selecting the second voter, shouldn’t you exclude the guy that you asked in the first phone call? Maybe, but what complication comes from excluding him/her?

Solution. If you do this, you alter the coin that you are flipping. The bias will decrease or increase, depending on whether the first guy said “yes” or “no”, respectively. The analysis gets messy, so you don’t want to do this. ■

(c) So, in each one of n phone calls, you pick a galaxy member *uniformly at random* and ask. Your plan is to eventually divide the number, Y , of “yes” answers by n to get $P ::= Y/n$. An MIT friend tells you that, as the numerical outcome of a random experiment, this P is a random variable, and that, according to his calculations,

$$\Pr \{|P - p| \leq 0.03\} \geq 0.95.$$

When you are done calling people, you divide to get P , and it's 0.53. You call the president up and... what do you say?

- (1) Mr. President, $p = 0.53$!
- (2) Mr. President, with probability at least 95%, p is within 0.03 of 0.53!
- (3) Mr. President, either p is within 0.03 of 0.53 or something very strange (less than 5-in-100) has happened.

For each statement answer: (i) Are you justified to claim it? (ii) Is it true?

Solution. Statement (1) is clearly off the mark.

- (i) Since you haven't asked all galaxy members, you can only make probabilistic statements about p .
- (ii) Statement (1) is also false, since $p = 0.52$. However, with a different choice of voters, it could have been true. Of course, even in that case, you wouldn't be justified in making it.

It would be wrong for you to make statement (2).

- (i) The unknown fraction, p , is a *constant*, not a random variable. It is *either* within 0.03 of 0.53, or more than 0.03 away of 0.53. It doesn't make much sense to talk about the probability that it has this, or any other, property: it does or it doesn't. And since you don't know which case holds, it would be wrong for you to make statement (2) about p .
- (ii) Now you could argue that statement (2) is actually true in this case. Namely, since the constant p is 0.52, it is indeed within 0.03 of 0.53. So you might claim that the probability in question is actually 1, and therefore is indeed at least 95%. But introducing probability in this way is misleading, at best.

Statement (3) is the correct one.

- (i) You are justified in making statement (3). To see why, start with the statement

$$\text{either } |0.53 - p| \leq 0.03 \quad \text{or} \quad |0.53 - p| > 0.03.$$

which is clearly true. Now read it as follows: *Either* p is within 0.03 of 0.53 *or* it is not and therefore my random random variable, P , took a value from a set that is hit only 5 times in 100. So, clearly *either* p is within 0.03 of 0.53, *or* something strange has happened.

- (ii) Statement (3) is true. In this particular case, it is true because the first half of it is true. But it would still be true even if p was not within 0.03 of 0.53. That's because statement (3) is an assertion about the behavior of the random variable, P , that is true no matter what the value of p is, so you can legitimately make the statement without knowing what p really is.



A third problem that originally appeared here has been deleted. Few, if any, teams got to this problem; it will be postponed for a later class problem session.