CS109A Spring 2022 Week 4 Notes

Ian Tullis

April 19, 2022

Since Quiz 1 is looming, this week we'll focus on practicing Week 1-2 stuff.

1 Double De Morgan

(*) Let A, B, and C be events, and let D be the event $(A \cap B) \cup C$. Use De Morgan's laws to produce an expression for D^c in terms of A^c, B^c, and C^c.

2 A common mistake

(**) When we roll 5 6-sided dice, the probability of seeing two of one number and three of a different number – e.g., 52522 – is $\frac{6 \cdot 5 \cdot \binom{5}{2}}{6^5}$. Why is it that when we roll 4 6-sided dice, the probability of seeing two of one number and two of a different number is **not** $\frac{6 \cdot 5 \cdot \binom{4}{2}}{6^4}$?

3 Questionable claims

(**) Let A, B, and C be three events, each of which has a probability *strictly* between 0 and 1. For each of the following claims, either give a short mathematical argument for why it is true, or give a counterexample.

(a) If A is independent of B, then A is independent of B^c.

(b) If A is conditionally independent of B given C, then A is conditionally independent of B given C^c.

(c) If A is independent of B, and B is independent of C, then A is independent of C.

(d) $P(A|B) = 1 - P(A^c|B)$.

(e) $P(A|B) = 1 - P(A|B^c)$.

(f) $P(A|B, C)P(B|C)P(C) = P(A)P(B|A)P(C|A, B)$.

4 When expectations don't meet expectations

If we roll a six-sided die six times, we would *expect* (in the sense of "expectation") to see every number once. But remember that expectations sometimes don't match our "expectations"! Should we actually be *surprised* if we don't see every number once? If we see the same number come up twice, should we be *suspicious*?

In the following problem, we will use 4-sided dice to make the math easier.[1] Suppose we roll a fair 4-sided die four times. Let C_1, C_2, C_3, C_4 be the counts of ones, twos, threes, and fours that we see (so $C_1 + C_2 + C_3 + C_4 = 4$). Let C_{max} be $\max(C_1, C_2, C_3, C_4)$, i.e., the largest number of instances of any one roll that we see. If we roll two 3s and two 4s, for example, then $C_{max} = 2$. Or if we roll one 1 and three 2s, then $C_{max} = 3$.

(***) What is the complete probability distribution of C_{max}? (Hint: Find the values one at a time, in whatever order makes it easiest.)

5 Hit point insecurity

In Dungeons and Dragons, I always dread rolling dice to determine my character's hit points. What if I get unlucky and roll a 1? What if someone else gets so lucky that their sneaky rogue has as many hit points as my beefy fighter?

Suppose that I determine my fighter character's hit points (H_F) by rolling ten 10-sided dice and adding them together, and my friend determines her rogue character's hit points (H_R) by rolling ten 6-sided dice and adding them together.

(a) (*) What is $\mathbf{E}(H_F)$?

(b) (***) What is the **exact** probability that my friend's rogue has **exactly** the expected number of hit points for a fighter, i.e., $P(H_R = \mathbf{E}(H_F))$? Your answer must be a single term with a single $\binom{n}{k}$ type expression in the numerator.

(c) (**) Suppose that my Dungeon Master allows me to reroll any 1s that come up when I roll my ten 10-sided dice, but only once each (i.e. even if a rerolled 1 comes up 1 again, I have to keep that 1). What is $\mathbf{E}(H_F)$ in this case?

[1] Believe me, this problem is *hot garbage* to solve with 6-sided dice, and I say this as a combinatorics fan...

6 Critical success and critical failure

(a) (*) Suppose you roll 20 20-sided dice. What is the probability that you will see at least one 20?

(b) (****) Suppose you roll 20 20-sided dice. What is the probability that you will see at least one 20 and at least one 1? (This is hard! Feel free to leave your answer in terms of summations etc. Also consider breaking the problem down into cases that you know are mutually exclusive and exhaustive, so that you don't have to worry about double-counting.)

7 Six-card poker

(This appeared in our Week 3 notes, but we didn't get to it. I'm including it here so that it doesn't fall through the cracks.) One dirty secret of Vegas is that the resorts are pretty much all the same. Suppose that one of them tries to spice things up by introducing six-card poker! The game uses a standard deck, but with hands of six cards. (All descriptions are of unordered sets of cards.)

(a) (**) A "three pair" is a hand of two cards of one rank, two of another rank, and two of a third rank. What's the probability of drawing a three pair?

(b) (**) A "double straight" is two cards of one rank, followed by two cards of the next highest rank, followed by two cards of the next highest rank after that. What is the probability of drawing a double straight? (In CS109, straights can use an ace as a low card.)

(c) (**) A "pyramid" is three cards of one rank, two cards of another rank, and one card of a third rank. What is the probability of drawing a pyramid?

8 What even is LOTUS

(*) Consider a random variable X with the following probability distribution supported on the integers 1 through 6. (The ugly numbers here are needed to make this a valid probability distribution, i.e. to ensure that $\sum_{x=1}^{6} p(X = x) = 1$.)

$$p(X = x) = \frac{3600}{5369x^2}$$

What is $\mathbf{E}[X^2]$?

9 False positives

(***) Suppose that a certain disease occurs in 10% of the population, and that the probability that a patient has the disease, given that a test for the disease comes back positive, is $\frac{3}{5}$. Can you determine the false positive rate of the test?

Solution to Problem 1

According to De Morgan's laws, for general events W and X, $(W \cup X)^c = W^c \cap X^c$. Therefore $D^c = (A \cap B)^c \cap C^c$. Also per De Morgan's laws, $(A \cap B)^c = A^c \cup B^c$, so in the end we have $D^c = \boxed{(A^c \cup B^c) \cap C^c}$.

(Notice that the parentheses matter! $A^c \cup (B^c \cap C^c)$ would be a different expression.)

Solution to Problem 2

To explain the first expression:

- We will consider the rolls as an ordered string.
- There are 6 possible outcomes for each die, so there are 6^5 possible outcomes overall. This is the sample space.
- We have 6 choices for the number that will be the group of 3, and 5 choices for the number that will be the group of 2. (We could have also chosen the group of 2 first and then the group of 3, and it would still be $6 \cdot 5$)
- We have $\binom{5}{2}$ ways of assigning the locations (in the string) of the two rolls that will be our "group of 2" number.

So what goes wrong when we try to apply the same reasoning to the four-die case?

- We will consider the rolls as an ordered string. - still OK
- There are 6 possible outcomes for each die, so there are 6^4 possible outcomes overall. This is the sample space. - still OK
- We have 6 choices for the number that will be the group of 2, and 5 choices for the number that will be the group of 2. – oh no! Now the groups are of equal size, so saying "the first group of 2 will be 3s and the second group of 2 will be 4s" is the same as saying "the first group of 2 will be 4s and the second group of 2 will be 3s". So we are double-counting – an outcome like 4334 will be counted once as "two fours and two threes", and again as "two threes and two fours"! What we really want here is 'one group of 2 will be 4s and the other group of 2 will be 3s". We can fix the double-counting by dividing by 2.

Another equivalent way to frame the answer is as $\frac{\binom{6}{2}\binom{4}{2}}{6^4}$. We just need to choose two of the six possible numbers to be our repeated values, but there is no notion that one is "first" and one is "second".

Solutions to Problem 3

(a) If A is independent of B, we can write $P(A|B) = P(A)$, using one definition of independence.

Now, by the Law of Total Probability, $P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$. But using $P(A|B) = P(A)$, this becomes $P(A) = P(A)P(B) + P(A|B^c)P(B^c)$. With some rearrangement, and using $P(B^c) = 1 - P(B)$, this becomes $P(A)(1-P(B)) = P(A|B^c)(1-P(B))$. Dividing out by $1-P(B)$ (which is nonzero because we are told that $P(B) \neq 1$), we get $P(A) = P(A|B^c)$, i.e., A is independent of B^c. So this is true!

(b) This is not necessarily true. Suppose that we have the following:

- $P(A \cap B \cap C) = \frac{1}{8}$
- $P(A \cap B \cap C^c) = \frac{1}{4}$
- $P(A \cap B^c \cap C) = \frac{1}{8}$
- $P(A \cap B^c \cap C^c) = 0$
- $P(A^c \cap B \cap C) = \frac{1}{8}$
- $P(A^c \cap B \cap C^c) = 0$
- $P(A^c \cap B^c \cap C) = \frac{1}{8}$
- $P(A^c \cap B^c \cap C^c) = \frac{1}{4}$

We see that $P(A \cap B|C) = \frac{1/8}{4/8} = \frac{1}{4}$, $P(A|C) = \frac{2/8}{4/8} = \frac{1}{2}$, and $P(B|C) = \frac{2/8}{4/8} = \frac{1}{2}$. So $P(A \cap B|C) = P(A|C) \cdot P(B|C)$, i.e., A and B are conditionally independent given C.

However, $P(A \cap B|C^c) = \frac{1/4}{2/4} = \frac{1}{2}$, $P(A|C^c) = \frac{1/4}{2/4} = \frac{1}{2}$, and $P(B|C) = \frac{1/4}{2/4} = \frac{1}{2}$. Since $\frac{1}{2} \neq \frac{1}{2} \cdot \frac{1}{2}$, A and B are not conditionally independent given C^c. Notice the contrast with the answer to the previous part!

(c) The easiest way to see that this is not necessarily true is to notice that A and C might be *exactly the same event*, in which case A and C are clearly not independent.

(d) This is always true. $P(A|B) + P(A^c|B) = \frac{P(A \cap B)}{P(B)} + \frac{P(A^c \cap B)}{P(B)} = \frac{P(A \cap B) + P(A^c \cap B}{P(B)} = \frac{P(B)}{P(B)} = 1$.

(e) This is not necessarily true. Consider our frequent example of rolling an 8-sided die, with event A being that the roll is even and event B being that the roll is 6, 7, or 8. Then $P(A|B) = \frac{2}{3}$, $P(A|B^c) = \frac{2}{5}$, and $\frac{2}{3} + \frac{2}{5} \neq 1$.

(f) This is always true; these are just two ways of writing $P(A, B, C)$. For instance, use the definition of conditioning to see that $P(A|B, C) \cdot P(B|C) \cdot P(C) = \frac{P(A,B,C)}{P(B,C)} \cdot \frac{P(B,C)}{P(C)}\ cdot P(C) = P(A, B, C)$.

Solutions to Problem 4

- The only way to get $C_{max} = 1$ is for every value from 1 to 4 to come up exactly once. We can solve this using sample and event spaces: the sample space is of size 4^4 (all possible outcomes, treating the dice as distinct) and the event space is of size 4! (the number of ways for 1, 2, 3, 4 to come up in some order on the four rolls). So $P(C_{max} = 1) = \frac{4!}{4^4} = \frac{24}{256}$.

- The only way to get $C_{max} = 4$ is for all the dice to come up the same value. We can view this as the probability that the second through fourth dice all match whatever came up on the first die. This probability is $(\frac{1}{4})^3 = \frac{1}{64}$. Just for ease of comparison with the other values, let's call it $\frac{4}{256}$.

- To get the size of the event space for $P(C_{max} = 3)$, we can start by picking the value to be repeated (there are 4 choices) and the other value (there are 3 choices left over), then use the binomial coefficient $\binom{4}{3} = 4$ to get the number of ways to distribute these relative to one another. Therefore the size of the event space is $4 \cdot 3 \cdot 4 = 48$, so the overall probability is $\frac{48}{256}$.

- Now we can note that 1, 2, 3, and 4 are mutually exclusive and exhaustive outcomes, so $P(C_{max} = 2) = \frac{256-24-4-48}{256} = \frac{180}{256}$. But it's good to review how to get this value directly, so let's do that.

  We can do something like what we did for $P(C_{max} = 3)$, but now there are two possibilities for the leftover two: either they are the same (Scenario 1) or different (Scenario 2).

  - In Scenario 1, we have one 2×-repeated value and two (different) singleton values. The multinomial coefficient gives us the number of ways to order these relative to each other: $\binom{4}{2,1,1} = \frac{4!}{2!1!1!} = 12$. There are $4 \cdot \binom{3}{2} = 12$ to choose the three values while avoiding double-counting. (See problem 2 for why it's $\binom{3}{2}$ here and not $3 \cdot 2$. Yes, it is annoying and I also have to double-check and think through it every time!) So the total number of ways here is 144.
  - In Scenario 2, we have we have two 2×-repeated values. There are $\binom{4}{2} = 6$ ways to pick these while avoiding double-counting. Then there are $\binom{4}{2} = 6$ ways to order these relative to each other. So we get $6 \cdot 6 = 36$.

  Since Scenario 1 and Scenario 2 are mutually exclusive, the total event space size is $144 + 36 = 180$, which is what we wanted. Yay!

  (If you get flustered over situations like Scenario 2, it helps to just write out some values. For example, writing out 1122, 1212, 1221, 2112, 2121, 2211, then 1133, 1313... makes it much clearer that the number of ways for Scenario 2 to happen is 36.)

Therefore the must common outcome, *by far*, is to see at most two of one number. So even though we really do expect (in the mathematical sense) to see each value 1 time in the 4 rolls, we should not confuse ourselves and expect to see this happen often on any given set of 4 rolls. The mathematical expectation is what happens on average *over multiple sets of rolls*, and is often not super useful for thinking about what happens on any particular roll.

Solutions to Problem 5

(a) For a single roll R, the expected value $\mathbf{E}(R)$ is $\frac{1+2+...+10}{10} = \frac{11}{2}$. So by linearity of expectation, $\mathbf{E}(H_F)$ is 10 times that, i.e., $\boxed{55}$.

(b) This is tricky, but we can observe that in order for the rogue to get exactly 55 out of the possible 60 hit points, my friend must roll something close to all sixes, but with exactly 5 one-point deductions distributed among the ten dice. We can safely use the divider method here because even if all five deductions end up on the same die, this is still a valid result (a roll of 1). So we need to distribute 5 things among 10 buckets, and per the usual divider method formula, there are $\binom{5+10-1}{10-1} = \binom{14}{9}$ ways to do so.

This gives us the event space, and the sample space is all ways of rolling the 10 dice: 10^{10}. (Notice that in both the numerator and the denominator, the order of the rolls matters.) So the overall answer is $\boxed{\frac{\binom{14}{9}}{10^{10}}}$, which is very small (about 1 in 5 million).

Remark: This would not have worked with 54, for instance, since then the divider method could potentially assign all six deductions to the same die, which is not possible.

(c) As usual, let's start by thinking about just one die roll. $\frac{9}{10}$ of the time, we keep the original value on the die, and the other $\frac{1}{10}$ of the time, it is like a single new standard roll. Therefore the expected value of the single die is $\frac{1}{10}(2) + ... + \frac{1}{10}(10) + \frac{1}{10}(\mathbf{E}(R)) = \frac{54}{10} + \frac{1}{10}(\frac{11}{2}) = \frac{119}{20}$. Then $\mathbf{E}(H_F) = \boxed{\frac{119}{2}} = 59.5$. This is a noticeable improvement over the mean of 55 without the DM's generosity.

Solutions to Problem 6

(a) The probability of seeing at least one 20 is 1 minus the probability of seeing no 20s, i.e., every roll comes up something other than 20. That probability is $(\frac{19}{20})^{20}$, so the answer is $\boxed{1-(\frac{19}{20})^{20}}$.

We implicitly used a binomial distribution there, and we would get the same answer by using one directly, with $n = 20$, and $p = \frac{1}{20}$ being the probability of rolling a 20 on any given roll: $1-P(X = 0) = 1-\binom{20}{0}(\frac{1}{20})^0(\frac{19}{20})^{20}$.

(b) This is hard to solve directly, but we can break it up into four mutually exclusive and exhaustive cases:

(a) There are no 1s or 20s.

(b) There is at least one 1, but there are no 20s.

(c) There is at least one 20, but there are no 1s.

(d) There is at least one 1 and at least one 20.

Case i. is similar to part (b), but with each die having a $\frac{18}{20}$ probability of not coming up 1 or 20. So the probability of that case is $(\frac{18}{20})^{20}$.

For Case ii., we can further subdivide this (into mutually exclusive and exhaustive subcases) based on where in the sequence we see our first 1:

- If we get a 1 on the first die, then it doesn't matter what the remaining dice are (they could even be more 1s), as long as they are not 20. The probability of this is $\frac{1}{20} \cdot (\frac{19}{20})^{19}$.
- We could get something other than 1 or 20 as the first die, then get 1 as the second die. Then it doesn't matter what the remaining dice are, as long as they are not 20. The probability of this is $\frac{18}{20} \cdot \frac{1}{20} \cdot (\frac{19}{20})^{18}$.

And so on. An expression for the sum of these probabilities is $\sum_{i=0}^{19}(\frac{18}{20})^i(\frac{1}{20})(\frac{19}{20})^{19-i}$.

Case iii. can be handled the same way as Case ii. Therefore the probability of Case iv., which is what we want, is 1 minus the sum of the other three Cases, i.e.,

$$\boxed{1-(\frac{18}{20})^{20} - 2\sum_{i=0}^{19}(\frac{18}{20})^i(\frac{1}{20})(\frac{19}{20})^{19-i}}.$$

As is often the case in combinatorics, there may be a nicer way to do this! Let me know if you find one.

If you got a different expression and want to check it on Wolfram Alpha: that answer is

Result

$$\frac{21\,212\,944\,650\,652\,080\,893\,863\,087}{52\,428\,800\,000\,000\,000\,000\,000\,000} \approx 0.404605$$

Solutions to Problem 7

Let's do this treating outcomes as unordered sets rather than ordered lists.

(a)
- Choose the three ranks involved: $\binom{13}{3}$.
- For each rank, choose the suits of the two cards: $\binom{4}{2}^3$.

Therefore the answer is $\frac{\binom{13}{3}\binom{4}{2}^3}{\binom{52}{6}} = \boxed{\frac{594}{195755}}$.

(b) First, there are 12 possible ranks that a straight of 3 cards can start on: A, 2, ..., Q. Once we pick the starting rank, it forces the ranks of all of the cards. E.g., picking 9 as the starting rank forces the other two ranks to be 10 and J.

Now, for each of the three ranks involved, our two cards of that rank can have any two of the four suits. So the answer is $\frac{12\cdot\binom{4}{2}^3}{\binom{52}{6}} = \boxed{\frac{324}{2544815}}$.[2]

(c) Now we must pick one rank for the set of three, one of the remaining ranks for the set of two, one of the remaining ranks for the remaining lone card. Then, when choosing suits, we pick three of the four cards of the first rank, two of the four cards of the second rank, and one of the four cards of the third rank. This is $\frac{\binom{13}{1}\binom{12}{1}\binom{11}{1}\binom{4}{3}\binom{4}{2}\binom{4}{1}}{\binom{52}{6}} = \boxed{\frac{1584}{195755}}$.

[2]Note that this hand is actually a subset of the previous one. If this were an actual casino game, the rules would have to specify that "three pair" does not include "double straight"s.

Solution to Problem 8

This problem is supposed to illustrate why the Law of the Unconscious Statistician can be useful. It says that for a random variable X and a function g, $\mathbf{E}[g(X)] = \sum_{\text{all supported } x}(g(x) \cdot P(X = x))$. In this problem, $g(x) = x^2$ and $P(X = x) = \frac{3600}{5369x^2}$, so $g(x) \cdot P(X = x) = \frac{3600}{5369}$, conveniently. Then

$$\mathbf{E}[X^2] = \sum_{x=1}^{6}(x^2 \cdot P(X = x)) = \sum_{x=1}^{6} \frac{3600}{5369} = \boxed{\frac{21600}{5369}}$$

The point of that was that we never had to look in detail at the ugly probability distribution $p(X = x)$. (We never had to evaluate it for any particular value of x.)

Solution to Problem 9

Let T be the event of a positive test, and let D and D^c represent having and not having the disease. We are given that $P(D) = 0.1$ and $P(D|T) = 0.6$. The false positive rate of the test is $P(T|D^c)$; using Bayes' Rule we can write this as:

$$P(T|D^c) = \frac{P(D^c|T)P(T)}{P(D^c)}$$

We know $P(D^c) = 1 - P(D) = 0.9$. We also know $P(D^c|T) = 1 - P(D|T) = 0.4$. We still need $P(T)$; using the Law of Total Probability, this is $P(T|D)P(D) + P(T|D^c)P(D^c)$. But now we have an even bigger mess! Something feels wrong. Is there actually a unique solution to this problem?

The issue is that (as we showed when proving that the claim in Problem 3(e) was false) $P(T|D)$ is a separate parameter from $P(T|D^c)$, and we would need to know one of $P(T|D)$ and $P(T|D^c)$ to determine the other.

For instance, if $P(T|D) = 1$, then we have

$$P(D|T) = \frac{P(T|D)P(D)}{P(T|D)P(D) + P(T|D^c)P(D^c)} = \frac{1 \cdot 0.1}{1 \cdot 0.1 + P(T|D^c) \cdot 0.9} = 0.6$$

and we can solve $0.1 = 0.6(0.1 + 0.9 \cdot P(T|D^c)$ to get $P(T|D^c) = \frac{2}{27}$.

But if $P(T|D) = 0.1$, on the other hand, then we get a very different answer of $P(T|D^c) = \frac{1}{135}$.

So this is kind of a trick question – the answer to "can you determine the false positive rate of the test?" is no! More information is needed.