Data 8 Midterm Review Sheet
Spring 2017

1 Probability

1. A probability is a number from 0 to 1 representing the chance of **something happening.**
2. The chance that **something doesn't happen** is equal to one minus the chance that it happens; in other words, if the chance that something happens is $P(\text{event happens})$, the chance that the event doesn't happen is $1 - P(\text{event happens})$
   (a) Example 1: A die roll. The chance that you roll a 4 is 1/6. The chance that you don't roll a 4 is 1 - 1/6 = 5/6.
3. Probabilities as a proportion of equally likely events.
   (a) $P(A) = \frac{\text{number of outcomes that make A happen}}{\text{total number of outcomes}}$
   (b) Example: I roll a die once. What's the probability that I roll an even numbered side? 2,4,6 are the even numbers represented on a six sided die. There are six sides. $P(\text{roll an even number}) = \frac{\text{roll 3 or 4 or 6}}{\text{roll 1 or 2 or 3 or 4 or 5 or 6}} = \frac{3}{6} = \frac{1}{2}$
4. Multiplication rule
   (a) The rule: the chance that event A **and** event B both happen is the probability that A happens **times** the probability that B happens **given** that A has happened.
      i. As a mathematical statement, $P(AB) = P(B) * P(A|B)$
   (b) Example: **If** it rains, I will wear a jacket with 80% probability. If it **doesn't** rain, I will wear a jacket with 40% probability. The chance of it raining on any given day is 25%.
      i. What is the probability that I'm wearing a jacket and it's raining?

$$P(\text{wearing jacket and raining}) = P(\text{raining}) * P(\text{wearing jacket given that it's raining})$$

$$P(\text{wearing jacket and raining}) = 25\% * 80\% = 20\%$$

      ii. Some things:
         A. Why was the second chance higher than the first? Doesn't it make more sense for wearing jacket and raining to be happening together? Intuitively yes, but since the probability for rain was so low, it made the event that had no rain more likely!
         B. What would have happened if I had tried to calculate the probability like this:

$$P(\text{wearing a jacket}) * P(\text{raining given that I'm wearing a jacket})$$

         Much harder to figure out, right? What's the probability that on any given day, I'm wearing a jacket? I can't be sure just from reading the question. The probability that I wear a jacket is **dependent** on whether or not it's raining. By contrast, the rain doesn't care whether I'm wearing a jacket or not; it happens with 25% probability, regardless. This is an important point to notice when you're doing your own calculations!

(c) Now let's apply the rule from bullet 2 to this problem.

i. What's the probability that I'm wearing a jacket and it's not raining?

$$P(\text{jacket, no rain}) = P(\text{no rain}) * P(\text{wearing jacket given no rain})$$

$$P(\text{jacket, no rain}) = (1 - P(\text{rain})) * P(\text{wearing jacket given no rain})$$

$$P(\text{jacket, no rain}) = 75\% * 40\% = 30\%$$

ii. What's the probability that I'm not wearing a jacket and it's not raining?

$$P(\text{no jacket, no rain}) = P(\text{no rain}) * P(\text{no jacket given no rain})$$

$$P(\text{no jacket, no rain}) = (1 - P(\text{rain})) * (1 - P(\text{wearing jacket given no rain}))$$

$$P(\text{no jacket, no rain}) = 75\% * 60\% = 40\%$$

Notice the difference between this example and what I covered in section: in this case, I know the probability of jacket wearing in both rain and no rain cases, so this is calculable. In section, we only had probabilities given rain: so if it didn't rain, we couldn't calculate a probability.

(d) And now, something a bit more complicated.

i. What's the probability of no rain and no jacket two days in a row?

$$P(\text{no jacket, no rain on day 1 and no jacket, no rain on day 2})$$

$$= P(\text{no jacket, no rain on day 2})$$

$$*P(\text{no jacket, no rain on day 1 given no jacket, no rain on day 2})$$

Do the jacket wearing and rain on day 1 say anything about jacket wearing and rain on day 2? Or vice versa? No, because we have not stated any conditions that say so. So,

$$P(\text{no jacket, no rain on day 1 given no jacket, no rain on day 2})$$

$$= P(\text{no jacket, no rain on day 1})$$

This tells us

$$P(\text{no jacket, no rain on day 1 and no jacket, no rain on day 2})$$

$$= P(\text{no jacket, no rain on day 1}) * P(\text{no jacket, no rain on day 2})$$

Furthermore, neither of these probabilities rely on which day it is, so

$$P(\text{no jacket, no rain on day 1 and no jacket, no rain on day 2})$$

$$= [P(\text{no jacket, no rain any day})]^2 = (40\%)^2 = 16\%$$

based on the calculation in 4.c.ii.

(e) This is now reminiscent of coin flip problems. Remember that our calculation for two (fair) coin flips being heads was

$$P(\text{H on toss 1, H on toss 2}) = P(\text{H on any toss})^2 = (1/2)^2 = 1/4$$

Why is this the case? In both of these examples, my two events are **independent** of each other. That is, knowing whether it rained and whether I wore a jacket day one doesn't tell me anything about the rain and jacket wearing of day 2, and vice versa. Similarly, getting heads or tails on one toss of a coin doesn't tell me anything about my second toss, and vice versa. The conclusion of all this?

i. Given two independent events A and B, $P(B|A) = P(B)$

(f) The statement above allows us to more easily talk about probabilities where there's a bunch of events involved. For example, the chance of all heads in 20 coin tosses ($(1/2)^{20}$).

i. **Warning: this is a bit technical. Skip if short on time.** I think it's helpful intuition to have though. To motivate this, let's say that A is the event of coin toss 1 is heads, B is the event that coin toss 2 is heads, and C is the event that coin toss 3 is heads. This means to calculate the chance that all 3 coins are heads, we are calculating $P(ABC)$. We don't have a rule for 3 events, so we split ABC into two groups: A and (BC). BC is the event that both toss 2 and toss 3 are heads. Now we can write our multiplication rule:

$$P(ABC) = P(BC) * P(A|BC) \tag{1}$$

Since A is independent of what BC is, we use our conclusion from above to say

$$P(A|BC) = P(A) \tag{2}$$

We can also rewrite $P(BC)$ using multiplication rule as usual:

$$P(BC) = P(C) * P(B|C) \tag{3}$$

B is independent of C, so

$$P(B|C) = P(B) \tag{4}$$

Plugging (4) into (3) gives us

$$P(BC) = P(C) * P(B) \tag{5}$$

Plugging (5) and (2) into (1), we finally get

$$P(ABC) = P(C) * P(B) * P(A) \tag{6}$$

The tl;dr of it all? If every event is independent of every other event, then the chance of every event happening is the product the probabilities of the individual events.

(g) Remember how handy this can be when calculating the chance of something happening at least once. Recalling the example of "at least one faulty syringe in 5 randomly drawn syringes, given that any syringe has a 1/100 chance of being faulty":

i. "At least one faulty" is another way of saying "something is faulty"
ii. "Something is faulty" is the opposite of "nothing is faulty".
iii. "Nothing is faulty" is the same as "the first is not faulty and the second is not faulty and the second is not faulty and the third is not faulty and the fourth is not faulty and the fifth is not faulty"
iv. We now express the statement in (iii) as:

$$(1 - 1/100) * (1 - 1/100) * (1 - 1/100) * (1 - 1/100) * (1 - 1/100)$$
$$= (99/100) * (99/100) * (99/100) * (99/100) * (99/100)$$
$$= (99/100)^5$$

v. Since our original question asked for the opposite of this, as we noted in (ii), we take one minus this probability to be our final answer: $1 - (99/100)^5$.

(h) I think it's interesting to note that in the above example, we use the "one minus" rule in two different places: in calculating the probability of each individual event, as well as for calculating the answer to the whole question. To illustrate the difference, consider these equations (still based off of the syringe example)

$$(1 - 1/100)^{10} \tag{7}$$
$$1 - (1/100)^{10} \tag{8}$$
$$1 - (1 - 1/100)^{10} \tag{9}$$

What is each equation saying? (7) is the probability that **in 10 draws, all 10 are good**. (8) is the probability that **at least one is good**. (9) is the probability that **at least one is bad**. These are subtle but important differences!

5. The Addition Rule

(a) As the slides say, if some event A can happen in exactly one of two ways, then

$$P(A) = P(\text{first way}) + P(\text{second way})$$

The most important thing to note about this rule is that "can happen in exactly one of two ways" means that if it happens in one way, then **it doesn't happen in the second way** I think perhaps a better way of saying this is that if I'm adding P(A) and P(B) then A and B **cannot simultaneously happen** (this is known as being **mutually exclusive**.

To illustrate this: say we flip a coin twice, and try to find the chance that the two tosses had different results: There are two sequences of tosses that satisfies this condition: first a heads, then a tails (HT) **or** first a tails, then a heads (TH). We can quickly see that these are the only events that work (HH and TT are the only two other combinations of coin tosses, and neither works).

We check that the addition rule works by considering the following: if my first toss is a heads (as in HT) then it is obviously impossible for that first toss to also be a tails (as it is in TH). A similar argument applies to the second toss. By this, HT and TH cannot simultaneously be true, so the addition rule works. $1/4 + 1/4 = 1/2$, which is the same result that we got earlier by counting results.

(b) The condition that the events must be mutually exclusive is the reason that we can't use simple multiplication (which is just repeated addition) to solve "at least one" problems. A common error was to solve the problem "Find the chance that at least one syringe in 5 draws is faulty, given that the chance of any given syringe being faulty is 1/100" with the following formula: $5 * 1/100$. What this expression is saying is this:

$$P(\text{at least one faulty}) = P(\text{first is faulty}) + P(\text{second is faulty}) + ...$$

What's the problem with this? Well, it's certainly possible that my first is faulty **and** my second is faulty in the same draw of 5 syringes. Therefore, these events are not mutually exclusive and I can't use the addition rule like this.

(c) An example of where the addition rule **is** valid was the question "Find the chance that **exactly** one syringe is faulty. Our equation for this looks like

$$P(\text{exactly one faulty}) = P(\text{only the first is faulty}) + P(\text{only the second is faulty}) + ...$$

This is different from the equation in 5b because saying "only the first is faulty" **excludes** the possibility that the second is faulty. In excluding all other possibilities when describing the event, we've made sure that each of the probabilities we add together is mutually exclusive. So the final calculation for this looks like

$$P(\text{exactly one is faulty}) = (1/100) * (99/100)^4 + (99/100) * (1/100) * (99/100)^3 + ...$$

Which is then equal to

$$P(\text{exactly one is faulty}) = 5 * (1/100) * (99/100)^4$$