MA 2B RECITATION 01/05/11

1. Administrivia

I'm Brian, and I'll be your Math 2b TA this quarter. Here's some basic info.

Email: bhwang at caltech
Website: http://hwang.caltech.edu/ma2b/
Office: 155 Sloan
Office Hours: Friday 4:00-5:00, or by appointment.

There's a short questionnaire that I'd like you to fill out if you are planning to attend this section, so that I can get to know you better. Please fill it out and return it to me in my office. It helps me to learn your names and make sure that you know where you can go if you need help. Since this is a core class with many different majors and motivations, I think it's important to know what people already know and what they'd like to know by the end of the quarter. If you cannot make my listed office hours, feel free to make an appointment.

As I said in the email, the first homework is up and it is due **Monday at 10am**. Late assignments will not be accepted without a note from the Dean, in which case you need to tell your TA by midnight before the due date. However, we will drop the lowest homework score when calculating your final grade.

Don't worry about writing down everything I say or write on the board. You are free to take notes, but lecture notes will be posted to the section website, so it's much more important that you try and follow the train of thought in each recitation. A good way to review the material is to take bare bones notes, then try and recreate the lecture yourself later, using the lecture notes as a reference if you get stuck.

Make sure to check the notes even if you feel like you understand everything in a recitation. I will sometimes put bonus material in there, or informal thoughts on how you can attack certain problems.

There's an anonymous feedback system at the bottom of the section website. Any constructive feedback is welcome. Since TQFR reports are not available until after the quarter is over, this is a good way to tell how the recitation is going throughout the course.

My philosophy towards recitations to focus on the practical skills you need to do the homework. All the beautiful theory and context will be left to Prof. Rains and your textbooks. Here, I will focus on the essentials and emphasize the quick, dirty techniques that you will need to succeed in the class. Other TAs will have slightly different approaches, so if you have some time and have a relatively flexible schedule, feel free to shop around to find a TA that best fits you. You are also free to attend other sections. Just remember to submit your assignments to your "official" TA.

Date: January 5, 2011.

2. Basics

2.1. First Formulas. The first part of the course deals with discrete probability. Recall the following standard facts about probability. Let A denote an event.

$$0 \leq P(A) \leq 1, \quad P(A^c) = 1 - P(A).$$

You want to keep, or draw a Venn diagram to make sure that you're proving things correctly.

In Pitman, AB denotes $A \cap B$, so we have the simple formula

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

The best way to deal with formulas like this is to use a Venn diagram.

For a slightly less trivial example, consider the formula

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).$$

How should you think about this? The term $P(A) + P(B) + P(C)$ does indeed give us $P(A \cup B \cup C)$, but we double-counted $A \cap B$, $A \cap C$, and $B \cap C$. Therefore, we need to subtract $P(AB) + P(AC) + P(BC)$. However, in all this adding and subtracting, we added and removed $A \cap B \cap C$ three times and subtracted it three times, so we need to add it back in.

One easy consequence of the first formula is that if two events A and B are disjoint, that is, $P(AB) = 0$, then $P(A \cup B) = P(A) + P(B)$.

2.2. Conditional probability. The conditional probability of A **given** B is written $P(A|B)$ and stands for

$$P(A|B) = \frac{P(AB)}{P(B)}.$$

This makes sense. You know that B happens, so what's the chance that B and A happen together? That's the probability of A given B. The way to think of this is that if you are given B, you're in the "world where B holds," that is, the part of the sample space where B happens, and you care about when A happens in this world, which is precisely when the event $A \cap B$ occurs.

A straightforward formula involving conditional probability is

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c).$$

Once again, this formula makes sense if you reason through it. One way to say when A happens to look at what happens when you are in the world of B and when you're not in the world of B. Since you must always be in one of these worlds, this covers all the times that A can occur.

2.3. Warning. The following is a common error. It is not true in general that

$$P(AB) = P(A)P(B).$$

The formula actually is

$$P(AB) = P(A|B)P(B) = P(B|A)P(A),$$

and the general formula is

$$P(A_1 \cdots A_n) = P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 \cdots A_{n-1}).$$

If $P(AB) = P(A)P(B)$, then we say that the events A and B are **independent**. In other words, A and B are independent if and only if $P(A|B) = P(A)$.

Remark 2.3.1. Note that being independent and disjoint are wholly different notions. Disjoint events are *mutually exclusive*, that is, if one occurs, the other cannot occur. If events are independent, they have *no influence on each other*, that is, learning that one event occurs does not give any information about whether the other event occurs. This is right way to interpret $P(A|B) = P(A)$.

Disjoint events cannot be independent, because knowing whether one happens gives you information about the other event, namely, that it cannot happen.

However, the inverse statement—that non-disjoint are independent—is false. Events that are not disjoint may or may not be independent. For instance, breast cancer occurs more often in women than men. Let A be the event that a person is man and B be the event that a person has breast cancer. Men can get breast cancer, so A and B are not disjoint. However, the probability that a man get breast cancer is much lower than that of women, so A and B are not independent. For non-disjoint independent events, consider a roll of two dice. Let A be the event that the first die rolls a 3 and let B the event that the second die rolls a 3. Since we can roll a 3 and a 3, A and B are not disjoint. Neither die influences outcome of the other, so A and B are independent.

3. Examples

Let's start with a simple one.

Example 1. Suppose that we are flipping 5 coins. We get the following sequence for the first four flips:

$$HHHH.$$

What is the probability the next flip is heads?

Actually, this is a trick question. The question is ambiguous. Are we talking about the whole sequence or simply a single coin flip? The wording is unclear. It can be either $\frac{1}{2^5}$ or $\frac{1}{2}$ depending on the context. I was probably asking about just the coin flip, but how do you know for sure? When you come across something like this, ask for clarification!

Thus, context and wording is important. Probability problems can be worded in a peculiar way, so it is critically important to reread the problem carefully so that you are answering the right question.

OK. Let's do a real coin example.

Example 2. I have two fair coins. One is double-headed and one of which is normal (has heads and tails). I toss you a coin at random with equal probability. You flip it 10 times. You get 10 heads in a row. What is probability that you have the double-headed coin?

The short way (appealing to intuition): The only ways to get 10 heads is if you either pick the double-headed coin or if you pick the normal coin and get 10 straight heads. Therefore,

$$P(\text{Double}|10 \text{ Heads}) = \frac{1}{1 + 0.5^{10}} = \frac{1024}{1025}.$$

The long way (the kind we expect in a full solution):

$$P(10 \text{ Heads}) = P(10 \text{ Heads} \mid \text{Double})P(\text{Double}) + P(10 \text{ Heads} \mid \text{Normal})P(\text{Normal})$$
$$= 1 \cdot \frac{1}{2} + \left(\frac{1}{2}\right)^{10} \cdot \frac{1}{2}$$
$$= \frac{1}{2}\left(1 + \left(\frac{1}{2}\right)^{10}\right).$$

We have

$$P(\text{Double} \mid 10 \text{ Heads}) = \frac{P(\text{Double} \cap 10 \text{ Heads})}{P(10 \text{ Heads})}.$$

We need to find $P(\text{Double} \cap 10 \text{ Heads})$, but this is not immediate, since getting the double-headed coin and getting 10 heads are not independent events. We must therefore do case analysis. We

$$1 = P(\text{Double}^c) + P(\text{Double} \cap 10 \text{ Heads}) + P(\text{Double} \cap 10 \text{ Heads}^c).$$

Since $P(\text{Double}^c) = \frac{1}{2}$ and $P(\text{Double} \cap 10 \text{ Heads}^c) = 0$, we have $P(\text{Double} \cap 10 \text{ Heads}) = \frac{1}{2}$. Therefore,

$$P(\text{Double} \mid 10 \text{ Heads}) = \frac{P(\text{Double} \cap 10 \text{ Heads})}{P(10 \text{ Heads})} = \frac{1/2}{\frac{1}{2}(1 + (\frac{1}{2})^{10})} = \frac{1}{1 + (\frac{1}{2})^{10}} = \frac{1024}{1025}.$$

This is an example of a problem to which you could appeal to your intuition to check your answer. However, always verify your thinking by putting your thought into symbols and working through the reasoning. On an exam, if we ask you a question like this (say as a single part of a multi-part question) and say explicitly that you do not need to justify your argument, then the short solution is OK. However, if we do not say this explicitly in the instructions, assume that you need to justify your steps, in which case the long solution is recommended.

Example 3. You are at Sorting Day at Hogwarts. There are four houses and the Sorting Hat assigns a student to each house with equal probability. What are the chances that you will end up in the Ravenclaw house?

$$P(\text{Ravenclaw}) = 1/4.$$

Just checking.

Suppose that if your last name is "Malfoy," you will be sorted into Slytherin with probability 1. If your last name is Malfoy, what is the probability that you will be sorted into Hufflepuff?

$$P(\text{Hufflepuff} \mid \text{Malfoy}) = 0.$$

Similarly, note that

$$P(\text{Malfoy} \mid \text{Hufflepuff}) = 0.$$

Given such examples, you might think that $P(B|A) = P(A|B)$, but this is another common mistake! This is NOT true in general. Consider the following example:

$$P(\text{Slytherin} \mid \text{Malfoy}) = 1$$

$$P(\text{Malfoy} \mid \text{Slytherin}) = 1/n$$

where n is the number of students sorted into Slytherin that year.

Example 4. You have been captured by pirates. However, the captain, who loves gambling and thrills and probability agrees to let you go if you play him in a round of Russian roulette, just you and the boss. Seeing no other way out, and it being almost time for your recitation, you agree.

You see him load two bullets into consecutive chambers of the gun, points at his head and pull the trigger, but you all you hear is a click. It was an empty chamber. He then gives you the gun to fire at your head, giving you the option to either spin the barrel before shooting or to take the next shot. What should you do? (Remember, you want it *not* to fire.)

You should not spin the barrel. With spinning, you have a $2/6 = 1/3$ chance of killing yourself. Without spinning and given that the first shot was nonlethal, there are five unknown chambers, two of which have bullets. We have

$$P(\text{first empty}) = 1 - P(\text{first kills}) = 1 - \frac{1}{3} = \frac{2}{3}.$$

Therefore,

$$P(\text{next kills} \mid \text{first empty}) = \frac{P(\text{next kills} \cap \text{first empty})}{P(\text{first empty})} = \frac{1/6}{2/3} = \frac{1}{6} \cdot \frac{3}{2} = \frac{1}{4}.$$

Hence, we should *not* spin the barrel.

You might have guessed these correctly. But try this one for size, it screwed with me the first time I saw it. This is just bonus material, not needed for your homework, but it is an example of the problems you can run into if you rely solely on your intuition and don't use the language of probability.

Example 5. Suppose you are an bug on an infinitely long stick. (Or a particle moving in 1 dimension, if you prefer.) Let's model this mathematically as the (integral) number line (considered as a subset of the real line, so the number line consisting of integers $\ldots, -2, -1, 0, 1, 2, \ldots$), and say that we begin at 0.

At each time step, we can go one left with probably p_ℓ, right with probability p_r, and stays still with probability $1 - p_\ell - p_r$.

These probabilities are dependent are dependent on the bug's position, so what we're actually studying is $p_\ell(x)$ and $p_r(x)$ where x is some position on the line. I'm also assuming that the line is finite of length L, but this doesn't change the problem much. This kind of problem is called a "discrete one-dimensional random walk" and is a classical type of problem in probability.

If $p_\ell(x) = p_r(x) = c$ for some constant $0 \leq c < 0.5$, then we have the classical random walk with a well-known solution that you will learn sometime later. (I don't want to say this because it is likely that this will be a homework or exam problem.) However, consider this slight variation, we have

$$p_\ell(x) > p_r(x),$$

but $p_\ell(x+1) = p_r(x)$.

So what is this saying? It says that at any given point x, the bug is more likely to go left than right. However, if it does go left, it will come back with the same probability (although it's more likely to go left).

Question: what is the equilibrium probability distribution for the bug's position $W(x)$? In other words, if you let this run for a long enough time until the probability stabilizes, what are the chances that the bug is at the number x?

Let's simplify even further. Is it more likely that you will be on the left or the right of zero?

Guess #1. At any given point, chances are, you are more likely to go left than right. Therefore, this introduces a bias towards going left, so the walker is more likely to be on the left side of the line after a long time.

Guess #2. If you go left, it is much more likely that you will return right. If the bug happens to go right, then it's going to take a long time to return. Therefore, you're likely to find it on the right.

The Answer. It's actually a little of both. The correct answer is that there's an equal probability of finding the walker absolutely anywhere! This is highly unintuitive if you are not used to thinking about these problems, and the correct answer comes by following the math and not our intuition. However, once we analyze the answer in the language of probability, we can find a way of thinking about the problem.

At equilibrium, we require that there not be any probability "current." That is, the probability of being found at a site must not change with respect to time evolution. This can be expressed mathematically by the equation:

$$\frac{\partial W(x)}{\partial t} = W(x+1)p_\ell(x+1) + W(x-1)p_r(x-1) - W(x)[p_\ell(x) + p_r(x)] = 0.$$

For this to work, think of the little p's as rates. In our case, the rate for going left or right is the same as making the reverse move. Our equation thus reduces to

$$p_r(x-1)[W(x-1) - W(x)] = p_r(x)[W(x) - W(x+1)]$$

and the solution is that $W(y) = \frac{1}{L}$ for all y on the stick, where L is the length of the stick. A constant!

One way to think about this is to pretend that you are sitting at the halfway point of x and $x+1$. You will see flux coming from the left and from the right. In equilibrium, these rates should be the same. (Technically, this is a condition called "detailed balance" that is slightly stronger than the notion needed for the above result to be true.) In other words, we have

$$W(x)p_r(x) = W(x+1)p_\ell(x+1).$$

Since $p_r(x) = p_\ell(x+1)$, we have $W(x) = W(x+1)$.