Lecture 7

Conditional Probability

Up until now, we have discussed probabilities of events when our sample space consists of all possible outcomes. Conditional probability discusses probabilities of events given that another event has already happened first.

We denote the word "given" by using

For example, if we are looking at the example of rolling 1 fair die, the sample space is

Intuitively:

If Event A was that the roll was even and Event B was that the roll was less than or equal to 3, we can write

Now, consider the scenario where we are interested in the probability of Event B given that Event A already occurred. This new event is written ______ where the event that already happened is on the right of the vertical bar.

In practice, this changes ______________. To find the P(B|A), we are finding the probability that the outcome satisfies Event B, given that we already know the outcome satisfies Event A. So, we are effectively now shrinking the sample space to just those values in Event A. The new sample space is now

So, now we can find the probability of Event B using the relative frequency formula, and claim that

Formulaically:

Note: The formula is just a relative frequency formula, except that you've changed the sample space to just the possible outcomes in the event that you already know occurred.

Note: Think of this formula as equaling the probability of both occurring divided by the probability of the event that already occurred. Don't get caught up in where A or B is in the formula. Focus instead on truly understanding the formula and you'll have an easier time.

Example – Favorite Music Artist

We are studying favorite music artists at the university. We know that 13% of students prefer Kygo, 17% of students prefer Weezer, and 2% prefer both.

a) What are the 3 pieces of information given in terms of the probabilities of events K and W?

b) What is the probability that someone prefers Kygo given that they prefer Weezer?

Note: Another possible wording for this question is "we know someone prefers Weezer. What is the probability that he/she will (also) prefer Kygo?"

Note: (Interpretation) Out of all the people who prefer Weezer, ____ % also prefer Kygo.

c) Draw a probability table for this problem

d) What is the probability that someone prefers Weezer given that they do not prefer Kygo?

e) What is the probability that someone does not prefer Kygo given that they do not prefer Weezer?

f) We know someone prefers Kygo. What is the probability that he/she will also prefer Weezer?

Comments Regarding Conditional Probability

1.

2. The complement rule also applies when handling conditional probability AS LONG AS WE CONDITION ON THE SAME EVENT

Formulaically:

Note: Careful with this property. It's a very useful property but VERY easily misused.

Let's look at the music example from above…

We know that $P(K^c|W) =$

and that $P(K|W) =$

Thus,

3. When we calculated conditional percents in the C → C case for EDA, we essentially calculated conditional probabilities.

   Example – Smoking Status by Sex

   To determine if they can predict whether someone is a smoker based on their sex, researchers gathered the following information summarized in the two-way table below:

	Sex		
Smoker	**Male**	**Female**	**Total**
Yes	77	404	481
No	16	122	138
Total	93	526	619

   From this table we can fill in a contingency table by calculating the conditional percents. It's important here to remember from Lecture 2 that we always condition over the ***explanatory*** variable's total. In this case, the explanatory variable is

The appropriate contingency table is as follows:

	Sex		
Smoker	**Male**	**Female**	**Total**
Yes	77/93	404/526	
No	16/93	122/526	
Total	1	1	

Let's take a deeper look at how the box where Male and Yes cross was calculated.

Let Event M be the event that the person is male, and the Event Y the event that the person is a smoker.

which matches our contingency table.

Independence

This is a very powerful property in probability. Calculations become much easier, as you'll see…

Intuitively:

What does it mean for 2 events to be independent?

So, let's look at a couple examples:

Example 1 – Music Preference

We took a random sample of 400 students from 400 different universities to try to see what music preferences are for those within the age bracket of 18 to 29. Event A is that person 1 preferred Kygo over Weezer. Event B is that person 2 preferred Weezer over Kygo.

Are the 2 events independent?

Example 2 – Package Delivery

Going back to the example about Fed-Ex and Prime Delivery for delivering my taxes, we can consider three simple events. Event A is that the documents were delivered on time by Fed Ex. Event B is that the documents were delivered on time by Prime Delivery. Event C is that the documents were delivered on time by both.

Assuming the weather and traffic were fine, is Event A independent of Event B?

Is Event A independent of Event C?

Formulaically:

We have 4 conditions that prove independence between 2 events:

1.

2.

3.

4.

Example – Vacation Preference

Travelers is interested in studying which type of vacation their customers prefer, a cold or a warm one. The probability someone does not prefer a warm vacation is 68%. The probability that someone prefers a cold vacation is 71%. The probability that someone does not prefer a cold vacation and does not prefer a warm vacation is .01 (and just don't like fun…).

a) Write down the 3 pieces of information given in terms of the probabilities of events W and C.

b) Fill in a probability table from the information given.

c) Are the events disjoint?

d) Are the events independent?

Disjoint vs Independent

Recall the verbal definition of independent events:

Two events are independent if knowing that one event occurred has no impact on the probability that the other event occurred.

Now, we can recall the formulaic definition of disjoint:

This means that if A occurs, then B cannot occur, and vice versa. This means that knowing if A occurred has a direct impact on the probability that the other event occurred.

Formulaically,

for independence →

However, if $P(A \cap B) = 0$, then this implies that either P(A) or P(B) is equal to 0 which means that knowing one event occurs instantly dictates that the other did not. This implies that the events are not independent.

So, if two events are disjoint, then they MUST be not independent.

What is the case when 2 events are not disjoint?

They can be either independent or not independent:

Example 1 – Not Disjoint and Independent

Example 2 - Not Disjoint and Not Independent

Multiplication Principle

If we know that 2 events are ***independent***, and we are interested in finding the probability of their intersection, we can apply the principle of independence that we just multiply the two probabilities to find the probability of the intersection.

So, let's consider an example

Example – Urn With Replacement

*An urn has 6 green balls and 6 yellow balls. If 2 balls are chosen randomly **with replacement**, what is the probability that both are green?*

Example – Dog Breeding

Dogs are inbred for such desirable characteristics as blue eyes color; but an unfortunate by-product of such inbreeding can be the emergence of characteristics such as deafness. A 1992 study by Dalmatians (Strain, et al as reported in <u>The Dalmatian Dilemma</u>*) found that 38% of all Dalmatians are deaf, 31% of all Dalmatians have blue eyes, and 42% of blue-eyed Dalmatians are deaf.*

a) Write down the information that is given in terms of the probabilities of the events A (the Dalmatian is deaf) and B (the Dalmatian is blue-eyed).

b) What is the probability that a randomly-selected Dalmatian is deaf and blue-eyed?

Example – Weather

Researchers at Brown are studying local weather and the effect of the temperature on precipitation. We know that the probability of it being cold is .50, given that it's cold there is a 40% chance it's raining, and there is a 40% chance it is raining.

a) Write down the information given in terms of the probabilities of the events R and C.

b) What is the probability that it's cold ***and*** it's raining?

c) What is the probability that it's not raining ***and*** it's cold?

d) What is the probability that it's not raining ***and*** it's warm?

e) If it is not raining, what is the probability that it's warm?

Some Notes on Independence

1. Sometimes we are told that events are independent, although this is rare but very useful.

2. If Event A and Event B are independent, then this implies that A^c and B are independent events, A and B^c are independent events, and A^c and B^c are independent events.

3. When we take a random sample from a LARGE population, we can assume independence (if the population is more than ___________ larger than the sample, then random selection implies independence)

   Example – Hair Color

   Say that the distribution of hair color in the U.S. is as follows:

Hair Color	Brown	Black	Blonde	Red
Percentage	48	19	25	8

   Two U.S. adults are selected at random. What is the probability that both have Black hair?

4. ___________ is key in using the multiplication property. If the 2 events are not independent, then you CANNOT use this property!

5. We only need to satisfy one of the 4 properties in order to show 2 events are independent. Conversely, we also only need to show that one of the 4 properties does not hold in order to show that the 2 events are not independent.

6. The multiplication principle for independent events can extend to more than 2 independent events.

Example – Coin Tosses

A fair coin is tossed 8 times. Which of the 2 outcomes is more likely?

H H H H H H H H   H T H T H T H T H T

Example – Children

A couple wants to have children until they have 1 boy and 1 girl, but won't have more than 3 children.

Recall the sample space for this random experiment was

However, recall that we said that the outcomes are not all equally likely. The multiplication principle is the reason why:

Example – Hair Color (continued)

6 U.S. adults are selected at random.

a) Which is more likely, all 6 having blonde hair or all 6 having red hair?

b) What is the probability that none of the 6 have blonde hair?

c) What is the probability that ***at least one of the 6*** has blonde hair?