MAT 105 - Summary of Lecture 12
Thursday, February 11, 2016

This branch of mathematics is the only one, I believe, in which good writers frequently get results which are entirely erroneous. (Charles Pierce)

Conditional probability: more examples

If we know event A occurs, and $P(A) > 0$, then the probability of B given A is defined as

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

$B|A$ reads "B given A" or "B conditioned on A", and $A \cap B$ stands for "A and B". Multiplying both sides by $P(A)$ gives another useful equation

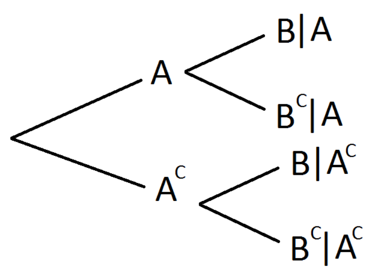
$$P(A \cap B) = P(A)P(B|A).$$

Think of this as a two step experiment: for both A and B to occur, first A must occur, and then given that A occurs, B must occur too.

To find the probability of some event B, we can consider two possibilities: A occurs or A^c occurs (here A^c denotes the complement of the event A). We write it as

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

This is just a formal way to compute probabilities along the branches of a tree. Recall that we multiply horizontally in the tree (depth) and add vertically (breadth).



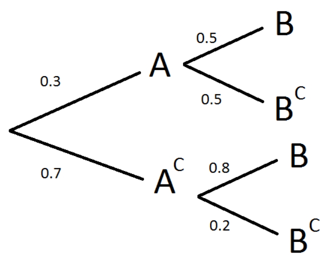
Here were derived **Bayes Formula**, which is at the base at Bayesian statistics:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A^c)P(B|A^c)}.$$

Why is this useful? Suppose we know how the occurrence of A affects the probability of B, but we would like to find out how the occurrence of B affects the probability of A. Bayes Formula allows us to do exactly that.

Examples:

1. Consider the following tree, with probabilities along the branches.



   We read the tree as follows:

   (a) $P(A) = 0.3$

   (b) $P(A^c) = 0.7$

   (c) $P(B|A) = 0.5$

   (d) $P(B^c|A) = 0.5$

   (e) $P(B|A^c) = 0.8$

   (f) $P(B^c|A^c) = 0.2$

   And we compute

   (a) $P(A \cap B) = P(A)P(B|A) = (0.3)(0.5) = 0.15$

   (b) $P(A^c \cap B) = P(A^c)P(B|A^c) = (0.7)(0.8) = 0.56$

   (c) $P(B) = P(A \cap B) + P(A^c \cap B) = 0.15 + 0.56 = 0.71$

   (d) $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0.15}{0.71} = \dfrac{15}{71}$

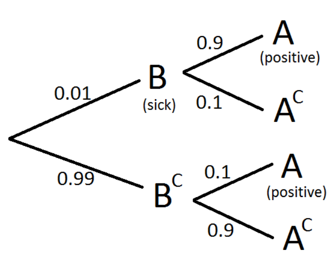
2. Suppose two events A and B have probabilities $P(A) = 0.2$, $P(B) = 0.3$ and $P(A|B) = 0.5$. We find

   (a) $P(A \cap B) = P(A|B)P(B) = (0.5)(0.3) = 0.15$

   (b) $P(B|A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{0.15}{0.2} = \dfrac{3}{4}$

3. Approximately 1% of women aged 40-50 have breast cancer. A woman with breast cancer has a 90% chance of a positive test from a mammogram, while a woman without has a 10% chance of a false positive result. What is the probability a woman has breast cancer given that she just had a positive test?

   Let B denote the event that a randomly selected woman 40-50 years old has breast cancer. Let A denote the event that the mammogram is positive for the disease.



Then we know $P(B) = .01$, $P(A|B) = .9$ and $P(A|B^c) = .1$, where B^c denotes the complement of B, namely the event that the woman does not have breast cancer. Then $P(B^c) = 1 - P(B) = .99$ and we need to find $P(B|A)$.

$$\begin{aligned} P(B|A) &= \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A \cap B) + P(A \cap B^c)} \\ &= \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)} \\ &= \frac{(.9)(.01)}{(.9)(.01) + (.1)(.99)} = \frac{9}{108} \end{aligned}$$

Another way to look at this problem is to find the proportion of women who have breast cancer and a positive test, to the women who have a positive test. The probability of a woman to have breast cancer and a positive test is $(.01)(.9) = .009$, or .9% of women tested are in this category. A woman can have a positive test if she has breast cancer, with probability .009 as we saw above, or she can be disease free but have a false positive test with probability $(.99)(.1) = .099$ or 9.9% of population. So the percentage of the population with a positive test is 10.8%. Thus, given a positive test, the proportion of women who are sick is $\frac{.9}{10.8} = \frac{9}{108}$.

The Monty Hall Problem

We play the following game show. The game show host (Monty Hall) shows us 3 doors. A car is hidden behind one door and goats behind the other two. We are asked to pick a door. For simplicity, let's assume we pick Door #1. Then the game show host opens Door #3 and reveals a goat. We are asked if we would like to switch our pick from Door #1 to Door #2. Should we switch? What is the probability of winning the car behind Door #1, now that we are left with only 2 doors? Is is 1/2? Or is it 1/3 as it was when we started the game? Or is it something else? To compute the probability that the car is behind Door #1, we look at 3 possible cases.

Case	Door #1	Door #2	Door #3	Host action
1	Car	Goat	Goat	reveal Door #2 or Door #3 (each with prob. 1/2)
2	Goat	Car	Goat	reveal Door #3
3	Goat	Goat	Car	reveal Door #2

Once we know *the host picked Door* #3, we cannot be in Case 3. So we are left with two equally likely cases: Case 1 and Case 2. Note that it is twice as likely he picked Door #3 in Case 2 than in Case 1, so the probability of being in Case 1 (hence winning the car if we do not switch) is 1/3. More precisely,

$$P(\text{Case 1}|\text{reveal Door \#3}) = \frac{P(\text{Case 1 \& reveal Door \#3})}{P(\text{reveal Door \#3})} = \frac{(1/3)(1/2)}{(1/3)(1/2) + (1/3)(1) + (1/3)(0)} = \frac{1}{3},$$

where we found

$$P(\text{reveal Door \#3}) = P(\text{Case 1 \& reveal Door \#3}) + P(\text{Case 2 \& reveal Door \#3}) + P(\text{Case 3 \& reveal Door \#3}).$$

Therefore switching at this point to Door #2 would increase our probability of winning the car to 2/3 from the 1/3 probability we started with. Intuitively, your first choice has 1/3 chance of winning and all the other probability of 2/3 is hidden behind the other unopened doors. Once a door is revealed, the door that remains keeps the 2/3 probability.

Now you can imagine what happens if we have more doors: you select one and the host reveals all other but one. What happens to your probability of winning if you stay with your initial choice? It stays the same, it does not improve. But switching brings you a big advantage. For example, if there are 10 doors, you pick the winner with probability 1/10. Now the game show host reveals 8 doors with goats and you have to decide whether to keep your door or switch. The other door has 9/10 chance of being a winner, so you should certainly switch. In fact, even if the host reveals only one door, you should still switch to increase your winning probability from 1/10 to 9/80.

Group Worksheet

1. Jack and Jill are trying a new game. The probability that Jack likes a new game is 0.45, and the probability that Jill likes a new game is 0.4. The probability that Jill will like the game, given that Jack does is 0.6. Find the probability that

   (a) both like the game

   (b) Jack likes the game, given that Jill does.

   Solution: Let A be the event that Jack likes the game and B the event that Jill likes the game. We know $P(A) = 0.45$, $P(B) = 0.4$, and $P(B|A) = 0.6$.

   (a) $P(A \cap B) = P(B|A)P(A) = (0.6)(0.45) = 0.27$ by the multiplication rule.

   (b) $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0.27}{0.4} = 0.675$ by definition of conditional probability, and using result from (a).

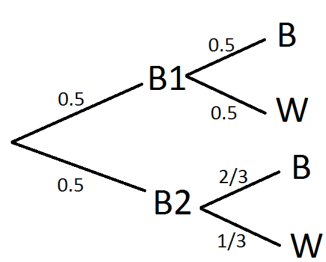
2. Consider 2 boxes. Box 1 contains 1 black and 1 white marble. Box 2 contains 2 black and 1 white marble. A box is selected at random and a marble is drawn at random from the selected box.

   (a) What is the probability that the marble is black?

   (b) What is the probability that Box 1 was selected, given that the marble is white?

   Solution:

   (a) Let B be the event that the marble is black and B1, B2 be the events that Box 1 and Box 2 respectively were selected. Note that the chance to pick a black marble from Box 1 is 1/2 and from Box 2 is 2/3.



$$P(B) = P(B \cap B1) + P(B \cap B2) = P(B|B1)P(B1) + P(B|B2)P(B2) = \frac{1}{2} \cdot \frac{1}{2} + \frac{2}{3} \cdot \frac{1}{2} = \frac{7}{12}.$$

   (b) Let W be the event we pick a white marble. So we need to find $P(B1|W)$.

$$P(B1|W) = \frac{P(B1 \cap W)}{P(W)} = \frac{P(W|B1)P(B1)}{1 - P(B)} = \frac{\frac{1}{2} \cdot \frac{1}{2}}{1 - \frac{7}{12}} = \frac{3}{5}.$$