Lecture: Bayes Theorem

We Know that

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

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

This suggests that you should be able to solve for $P(A|B)$ if you know $P(B|A)$, $P(A)$, and $P(B)$. In fact,

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

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

which is ~~the~~ Bayes' Theorem.

Note: We already saw $P(B|A) \neq P(A|B$
in general. Now we know

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

Example. We have already seen Bayes' Theorem in our snapchat example.

$$P(A) = P(\text{uses iPhone}) = 0.69$$
$$P(B) = P(\text{uses Snapchat}) = 0.5$$
$$P(B|A) = \text{fraction of iPhone users which use snapchat} = 0.6$$

(re)Compute $P(A|B)$ = probability that a Snapchat user is an iPhone user with Bayes Theorem. (3 Students)

Example. Suppose we don't know $P(B|A)$, (fraction of iPhone users who use Snapchat), exactly, but can only estimate it.

What can we conclude about $P(A|B)$ (the probability that the person who sent you a snap was using an iPhone)

1) If <u>all</u> iPhone users are Snapchat users?
2) If <u>no</u> iPhone users are Snapchat users?

(Do this in groups. Discuss. Why are these answers different?)

We know that for any A, A and A^c partition S. So for any B,

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

written in terms of conditional probabilities (again), we recall

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

Since $P(A^c) = 1 - P(A)$, we know $P(A^c)$ if we know $P(A)$. We can now prove

Theorem (Bayes Theorem, 2nd form) If $P(B) \neq 0$, and $0 < P(A) < 1$,

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

Proof. (This is just substituting the formula for $P(B)$ in.)

The usual use for this is called a "Bayesian update." We have an estimate for $P(A)$, ~~when a new fact, B, comes to our~~ and we learn a new fact, that B is true.

If B is true, then our estimate of $P(A)$ should be replaced by the more accurate $P(A | B)$.

Example. Lycanthropy is a rare condition, affecting 1 person in ~~10,000~~ 10 million. 90% of lycanthropes are so allergic to silver that they suffer a rash on skin contact.

On the other hand, 1% of humans display a similar skin allergy.

~~When greeting your~~
In 2011, CNN host Wolf Blitzer broke out into a rash after handling a silver tie clip.

What is your best estimate of the probability that Blitzer is a lycanthrope:

a) before the tie clip incident?
b) after the tie clip incident?

(Students should discuss in groups, then share answers.

S = sample space = people.
A = lycathropes A^c = humans
B = people with silver allergy

N.b. This whole thing is made up, but you're free to pretend that you (or I) believe it. It's not a real medical example b/c that can upset people needlessly.

~~If there's more time, have them write up "double dice" from two classes ago.~~

Next, students should prove

$$P(A \cap B \cap C) = P(A)\,P(B \mid A)\,P(C \mid A \cap B)$$

~~[illegible]~~

Example. An urn contains N red balls and M black balls. We draw w/o replacement 3 balls.

Compute P(all three are red) as P(ball 1 is red $\cap$ ball 2 is red $\cap$ ball 3 is red) using formula above.

Example. Generalize to k balls drawn from urn (all red).