Esssentials of Applied Data Analysis
IPSA-USP Summer School 2017

Handout - The Basics of Set Theory

Leonardo Sangali Barone
leonardo.barone@usp.br

jan/17

Colored balls, boxes and political parties.

Bayes' rule

Suppose there are A_1, A_2,..., A_n events and $A_1 \cup A_2 \cup ... \cup A_n$ = S. A_i events are mutually exclusive. Suppose there is another event B. For each A_i, B has a different joint conditional probability $P(A_i|B)$. B can be described as:

$$P(B) = P(A_1) * P(B|A_1) + P(A_2) * P(B|A_2) + ... + P(A_n) * P(B|A_n)$$

The probability that an specific A_i will occur given B is defined as:

$$P(A_i|B) = \frac{P(A_i) * P(B|A_i)}{P(A_1) * P(B|A_1) + P(A_2) * P(B|A_2) + ... + P(A_n) * P(B|A_n)}$$

Bayes' rule - a simple example

There are two political parties, A_1, with 60 members and A_2, with 40 members. All members of party A_1 are against the country engaging in a new war. In party A_2, one third of the members are against the new war, but the others are in favor of it. B is the event of being against enganging in war.

You turn on the radio and there is a politician speaching against war. You don't know her and you want to figure out her party. What is the probability that she belongs to party A_2 given she is against war?

Bayes' rule - a simple example

Let's think about the problem for a second. Whithout any calculations, is she more likely to belong to which party? Make your guess.

What is the probability that a a politician at random belongs to party A_1? And party A_2?

$$P(A_1) = 40/100 = 2/5; P(A_2) = 1 - P(A_1) = 1 - 2/5 = 3/5;$$

So without knowing that she is against the war, it is more likely that she belgons to party A_2, right? 3/5 would be a good first guess. But we do know she is against war and we can **update** our "guess" with this information.

Bayes' rule - a simple example

What is the probability that a politician from A_1 is against war? $P(B|A_1) = 1$. And from A_2? $P(B|A_2) = 1/3$.

Since we know that being or not in favor of war is conditioned on the party, we can **update** our guess in that way.

$$P(A_2|B) = \frac{P(A_2) * P(B|A_2)}{P(A_1) * P(B|A_1) + P(A_2) * P(B|A_2)} =$$
$$= \frac{3/5 * 1/3}{2/5 * 1 + 3/5 * 1/3} = \frac{1}{3} < P(A_2) = 3/5$$

Bayes' rule - colored balls

There are 3 boxes, each containing 10 colored balls. In the first box (A_1) there are 8 blue balls and 2 red balls. In the second box (A_2) there are 3 blue balls and 7 red balls. There are only red balls in the third box (A_3) I have randomly chosen a box and pick up randomly one ball from it. The ball was blue ($Blue$). What is the probability that I got the first box (A_1)?

Bayes' rule - colored balls

There are 3 boxes, each containing 10 colored balls. In the first box (A_1) there are 8 blue balls and 2 red balls. In the second box (A_2) there are 3 blue balls and 7 red balls. There are only red balls in the third box (A_3) I have randomly chosen a box and pick up randomly one ball from it. The ball was blue ($Blue$). What is the probability that I got the first box (A_1)?

$$P(A_1|B) = \frac{P(A_1) * P(Blue|A_1)}{P(A_1) * P(Blue|A_1) + P(A_2) * P(Blue|A_2) + P(A_3) * P(Blue|A_3)}$$

$$= \frac{1/3 * 8/10}{1/3 * 8/10 + 1/3 * 3/10 + 1/3 * 0} = \frac{8/30}{8/30 + 3/30 + 0} = \frac{8}{11} = 0.73$$

Bayes' rule - colored balls

There are 3 boxes, each containing 10 colored balls. In the first box (A_1) there are 8 blue balls and 2 red balls. In the second box (A_2) there are 3 blue balls and 7 red balls. There are only red balls in the third box (A_3) I have randomly chosen a box and pick up randomly one ball from it. The ball was blue ($Blue$). What is the probability that I got the first box (A_1)?

Bayes' rule - positive medical exam

$$P(A_1|B) = \frac{P(A_1) * P(Blue|A_1)}{P(A_1) * P(Blue|A_1) + P(A_2) * P(Blue|A_2) + P(A_3) * P(Blue|A_3)}$$

$$= \frac{1/3 * 8/10}{1/3 * 8/10 + 1/3 * 3/10 + 1/3 * 0} = \frac{8/30}{8/30 + 3/30 + 0} = \frac{8}{11} = 0.73$$

Bayes' rule - another way to think

A theorem that explains how one ought to change one's subjective probability in response to empirical evidence.

Assume:

T = theory or hypothesis

E = represents a new piece of evidence that seems to confirm or disconfirm the theory.

P(T)= Prior Belief = Probability of (T)

Then, the Posterior Probability of T:

$$P(T|E) = \frac{P(T) * P(E|T)}{P(T) * P(E|T) + P(T^c) * P(E|T^c)} =$$

Bayes' rule - another way to think

A theorem that explains how one ought to change one's subjective probability in response to empirical evidence.

Assume:
T = theory or hypothesis
E = represents a new piece of evidence that seems to confirm or disconfirm the theory.

P(T)= Prior Belief = Probability of (T)

Then, the Posterior Probability of T:

$$P(T|E) = \frac{P(T) * P(E|T)}{P(T) * P(E|T) + P(T^c) * P(E|T^c)} =$$

Bayes' rule - textbook example

Stokes campaing example - Moore and Siegel, chap 9, pp 188-190