PROBABILITY 7 : Bayes' Theorem

In this video I will review Bayes' Theorem. With the Bayes' Theorem you can find probabilities when you know other probabilities.

Bayes' Theorem & Exercisers

Here is the formula for Bayes' Theorem. We'll now calculate given we know someone exercises, what is the probability that they are a man? We'll use the exercise chart we have used in previous videos.

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

	NO EXERCISE	DID EXERCISE	TOTAL
MEN	78	22	100
WOMEN	83	17	100
TOTAL	161	39	200

P(A) will represent P(Man), or the probability we will pick a man at random. This is 100/200 or .5. P(B) represents P(Exercisers), or the percentage of people who exercise being 39/200 or .195. P(B|A) or the probability of getting an exerciser given we know we have a man is 22/100 or .22.

To calculate the probability you have a man given you know the person exercises you fill in the results to find P(A | B) and find that the probability is .564 or 56.4%. If we double check our work by dividing 22/39 we get .564.

$$P(A \mid B) = \frac{.5 \times .22}{.195} = .564$$

Probability that Person has Diabetes if Obese

Here I want to calculate the probability that someone has type 2 diabetes if they are obese.A represents people with type 2 diabetes. B represents people who are obese.

P(A | B) = P(Diabetes | Obese)
P(B | A) = P(Obese | Diabetes) = 62.4% of Type 2 Diabetics are Obese
P(A) = P(Having Diabetes) = 10.5% of Americans have Type 2 Diabetes
P(B) = P(Obesity) = 42.4% of Americans are Obese

P(A | B) = .105 * .624 / .424 = .155 or 15.5%

Bayes' Theorem Vs. Disease Testing

Now we will calculate something more complex. We will calculate the odds that someone has a disease knowing that testing is sometimes inaccurate. Let's say that 10% of the people in an area are expected to actually have the disease and this is what we know about testing accuracy.

	TEST YES	TEST NO
HAVE	96.5%	24%
DON'T HAVE	2.5%	97.5%

P(A) = Has Disease = .10
P(B | A) = Test says they Have it and they do = .965
P(B | Not A) = Test says they Have it and they Don't = .025
P(Not A) = Person Doesn't have Disease = .90

Previously we divided by the percentage P(A) x P(B | A) meaning we multiplied the percentage of the people who were sick by the percentage positive test given the person had the disease.

Now we must also account for the 2.5% of people that are not sick that get false positive results. We multiply those who test positive by the 2.5% false positive and add that to the bottom.

Then we plug in the numbers and find that .81 or 81% of people who test positive are expected to have the disease.

$$P(A \mid B) = \frac{P(A) \times P(B \mid A)}{P(A) \times P(B \mid A) + P(nA) \times P(B \mid nA)}$$

$$P(A \mid B) = \frac{.1 \times .965}{.1 \times .965 + .90 \times .025} = \frac{.0965}{.0965 + .0225} = .81$$

I hope you found this tutorial interesting. In the next video I'll cover Discrete Random Variables.