Bayes' Theorem

Bayes can do magic!

Ever wondered how computers learn about people?

Advanced Example

An internet search for "movie automatic shoe laces" brings up "Back to the future"

Has the search engine watched the movie? No, but it knows from lots of other searches what people are **probably** looking for.

And it calculates that probability using Bayes' Theorem.

Bayes' Theorem is a way of finding a probability when we know certain other probabilities.

The formula is:

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

Which tells us:	how often A happens *given that B happens*, written **P(A|B)**,
When we know:	how often B happens *given that A happens*, written **P(B|A)**
	and how likely A is on its own, written **P(A)**
	and how likely B is on its own, written **P(B)**

Let us say P(Fire) means how often there is fire, and P(Smoke) means how often we see smoke, then:

P(Fire|Smoke) means how often there is fire when we can see smoke
P(Smoke|Fire) means how often we can see smoke when there is fire

So the formula kind of tells us "forwards" P(Fire|Smoke) when we know "backwards" P(Smoke|Fire)

Example: If dangerous fires are rare (1%) but smoke is fairly common (10%) due to barbecues, and 90% of dangerous fires make smoke then:

$$P(\text{Fire}|\text{Smoke}) = \frac{P(\text{Fire})\,P(\text{Smoke}|\text{Fire})}{P(\text{Smoke})}$$

$$= \frac{1\% \times 90\%}{10\%}$$

$$= 9\%$$

So the "Probability of dangerous Fire when there is Smoke" is 9%

Example: Picnic Day

You are planning a picnic today, but the morning is cloudy

- Oh no! 50% of all rainy days start off cloudy!
- But cloudy mornings are common (about 40% of days start cloudy)
- And this is usually a dry month (only 3 of 30 days tend to be rainy, or 10%)

What is the chance of rain during the day?

We will use Rain to mean rain during the day, and Cloud to mean cloudy morning.

The chance of Rain given Cloud is written P(Rain|Cloud)

So let's put that in the formula:

$$P(\text{Rain}|\text{Cloud}) = \frac{P(\text{Rain})\,P(\text{Cloud}|\text{Rain})}{P(\text{Cloud})}$$

- P(Rain) is Probability of Rain = 10%
- P(Cloud|Rain) is Probability of Cloud, given that Rain happens = 50%
- P(Cloud) is Probability of Cloud = 40%

$$P(\text{Rain}|\text{Cloud}) = \frac{0.1 \times 0.5}{0.4} = .125$$

Or a 12.5% chance of rain. Not too bad, let's have a picnic!

Remembering

First think "AB AB AB" then remember to group it like: "AB = A BA / B"

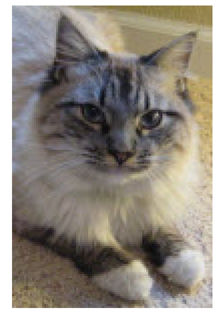
$$P(A|B) = \frac{P(A)\ P(B|A)}{P(B)}$$

"A" With Two Cases

One of the famous uses for Bayes Theorem is False Positives and False Negatives.

For those we have two possible cases for "A", such as **Pass**/**Fail** (or Yes/No etc)

Example: Allergy or Not?



Hunter says she is itchy. There is a test for Allergy to Cats, but this test is not always right:

- For people that **really do** have the allergy, the test says "Yes" **80%** of the time
- For people that **do not** have the allergy, the test says "Yes" **10%** of the time ("false positive")

If 1% of the population have the allergy, and **Hunter's test says "Yes"**, what are the chances that Hunter really has the allergy?

We want to know the chance of having the allergy when test says "Yes", written **P(Allergy|Yes)**

Let's get our formula:

$$P(\text{Allergy}|\text{Yes}) = \frac{P(\text{Allergy})\ P(\text{Yes}|\text{Allergy})}{P(\text{Yes})}$$

- P(Allergy) is Probability of Allergy = 1%
- P(Yes|Allergy) is Probability of test saying "Yes" for people with allergy = 80%
- P(Yes) is Probability of test saying "Yes" (to anyone) = ??%

Oh no! We **don't know** what the **general** chance of the test saying "Yes" is ...

... but we can calculate it by adding up those **with**, and those **without** the allergy:

- 1% have the allergy, and the test says "Yes" to 80% of them
- 99% do **not** have the allergy and the test says "Yes" to 10% of them

Let's add that up:

$$P(\text{Yes}) = 1\% \times 80\% + 99\% \times 10\% = 10.7\%$$

Which means that about 10.7% of the population will get a "Yes" result.

So now we can complete our formula:

$$P(\text{Allergy}|\text{Yes}) = \frac{1\% \times 80\%}{10.7\%} = 7.48\%$$

$$P(\text{Allergy}|\text{Yes}) = \text{about } \mathbf{7\%}$$

This is the same result we got on False Positives and False Negatives.

In fact we can write a special version of the Bayes' formula just for things like this:

$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(\text{not } A)P(B|\text{not } A)}$$

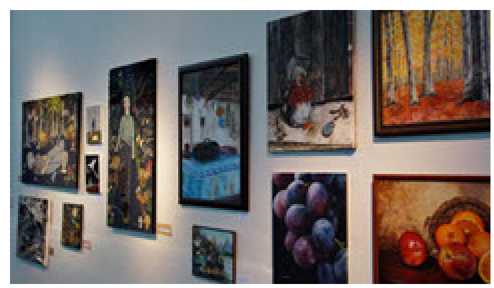
"A" With Three (or more) Cases

We just saw "A" with two case (A and not A), which we took care of in the bottom line.

When "A" has 3 or more cases we include them all in the bottom line:

$$P(A1|B) = \frac{P(A1)P(B|A1)}{P(A1)P(B|A1) + P(A2)P(B|A2) + P(A3)P(B|A3) + \text{...etc}}$$

Example: The Art Competition has entries from three painters: Pam, Pia and Pablo



- Pam put in 15 paintings, 4% of her works have won First Prize.
- Pia put in 5 paintings, 6% of her works have won First Prize.
- Pablo put in 10 paintings, 3% of his works have won First Prize.

What is the chance that Pam will win First Prize?

$$P(\text{Pam}|\text{First}) = \frac{P(\text{Pam})P(\text{First}|\text{Pam})}{P(\text{Pam})P(\text{First}|\text{Pam}) + P(\text{Pia})P(\text{First}|\text{Pia}) + P(\text{Pablo})P(\text{First}|\text{Pablo})}$$

Put in the values:

$$P(\text{Pam}|\text{First}) = \frac{(15/30) \times 4\%}{(15/30) \times 4\% + (5/30) \times 6\% + (10/30) \times 3\%}$$

Multiply all by 30 (makes calculation easier):

$$P(\text{Pam}|\text{First}) = \frac{15 \times 4\%}{15 \times 4\% + 5 \times 6\% + 10 \times 3\%}$$

$$= \frac{0.6}{0.6 + 0.3 + 0.3}$$

$$= 50\%$$

A good chance!

Pam isn't the most successful artist, but she did put in lots of entries.

So now you know how search engines can guess what you want: they simply keep track of what lots of people type in and what websites they eventually click on.

Then using Bayes they figure which ones are probably the best to show first.

It makes them look like they can read your mind!

Question 1 Question 2 Question 3 Question 4 Question 5
Question 6 Question 7 Question 8 Question 9 Question 10

Copyright © 2017 MathsIsFun.com