Stat310

Bayes' Rule

Hadley Wickham

1. Homework comments
2. Independence practice & proof
3. Bayes' rule & natural frequencies
4. Toolbox
5. Intro to random variables
   (the next big topic)

Homework

Homework

Please turn in to **DH mailbox** by 4pm!

(Webpage was wrong because of a bug in my template so I'll accept in class today, but not in the future)

Make sure to vote for office hours:
http://www.doodle.com/mceugkzvu8k9wets
Votes close 9am Friday.

Sorry about the mixup with the homework help session rooms.

Use your best judgement

(and write down any assumptions that you make)

Independence


THE
DENZEL WASHINGTON
VENN DIAGRAM
GLASSES
HAT
FACIAL HAIR
HAT & GLASSES
GLASSES & FACIAL HAIR
HAT & FACIAL HAIR
ALL THREE!
GLASSES
Manchurian Candidate
Out of Time
American Gangster
HAT & GLASSES
Mo' Better Blues
Man on Fire
HAT
Preacher's Wife
John Q
Courage Under Fire
Glory
GLASSES & FACIAL HAIR
Hurricane
Pelham 123
Book of Eli
ALL THREE!
Malcolm X
HAT & FACIAL HAIR
Devil in a Blue Dress
Training Day
FACIAL HAIR
Cry Freedom
He Got Game
Philadelphia
Heart Condition
Antwone Fisher
Inside Man
MAXIM.COM

Are the wearing glasses and wearing hat events independent?

21 Dennis Washingtons in total

Calculations

P(glasses) = 9 / 21 = 3 / 7

P(hat) = 9 / 21

P(glasses and hat) = 3 / 21 = 0.14

P(glasses) P(hat) = 9 / 49 = 0.18

Wearing a glasses and hat together is (slightly) less likely than we'd expect if they were independent.

Another proof

Assume that A and B are independent.
Show that A and B' are independent.

In homeworks and exams, correctly turning a word problem into a math problem will get you a few points. Writing down possible strategies will get you a few more.

Bayes' rule

Motivation

Imagine you select a person at random and give them an HIV test. If the HIV test is positive what's the probability that they have HIV? (An HIV test is >99.9% accurate)

a) 95-100%

b) 90-94%

c) 80-89%

d) 65-79%

e) < 64%

It's **50%!**

More precisely

Imagine you select a person at random and give them an HIV test. If the HIV test is positive what's the probability that they have HIV?

If you have HIV, the probability that the test is positive is 99.9%

If you don't have HIV, the probability that the test is negative is 99.99%

In Texas, about 13 people per 100,000 have HIV.

Your turn

Convert the previous word problem into mathematical probability problem by defining the appropriate events and writing down what we know about them.

What are we trying to find out?

Events

T = {test is positive}

H = {has HIV}

$P(T \mid H) = 0.999$

$P(T' \mid H') = 0.9999$

$P(H) = 1.3 * 10^{-4}$

We want to find $P(H \mid T)$

Bayes' rule

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

(Often need to use law of total probability down here)

Alternative approach

Use natural frequencies. Pick a large number of people and break into categories.

Much more intuitive, and helps you understand why this surprising result is true.

Toolbox

Complements

Convert union to sum

Convert intersection to conditioning
(and vice versa)

Convert intersection to product
(if independent)

Use law of total probability

Switch conditioning (Bayes' rule)

Random variables

Intro to rv's

Probability is a set function. Kind of tricky to deal with. Easier to deal with functions of numbers.

Want to ignore details of problem (e.g. specific events) and focus on essence.

Real world ➙ mathematical world

Definitions

A **random variable** is a random experiment with a numeric sample space. Usually given a capital letter like X, Y or Z.

(More formally a random variable is a function that converts outcomes from a random experiment into numbers)

The **space** (or **support**) of a random variable is the range of the function (cf. sample space)

Definitions

If the size of the support is **finite** or **countably infinite**, then the random variable is **discrete**.

If the size of the support is **uncountably infinite**, then the random variable is **continuous**.

Your turn

The random experiment is to go to Vegas with $100 and play blackjack until you make $1,000 or lose all your money.

How many different random variables can you generate from this random experiment? Brainstorm in pairs.

Some ideas

How many games you play

The most amount of money you have at any point. Or the least amount.

The average amount of money you win or lose in each game.

Review for next week

We're going to start working with sums and a little bit of calculus.

http://tutorial.math.lamar.edu/Classes/CalcI/SummationNotation.aspx

Basic calculus: differentiation & integration of (e.g.) polynomials. Practice using wolfram alpha