All About that Bayes

1. What are the limitations of chi-square?
2. Predict the probability of disease in a pedigree's next-born child
3. Use Bayesian networks to improve on chi-square...

χ^2: I KNEW YOU WERE TROUBLE

Remember using chi-square to diagnose cancer?

Gene FGFR2	Gene TOX3	Gene ESR1	Gene CASC16
0.63	0.51	0.81	0.71

Expected, model cancer profile compared against our patient?

Gene FGFR2	Gene TOX3	Gene ESR1	Gene CASC16
0.60	0.45	0.75	0.92

1. How many cancer patients are there in the world?
2. Which one should be our “model” patient profile?
3. Is every cancer caused by the same set of genes?

What happens to the chi-square value if we add “Does Patient Have Cat?” to profile? Does χ^2 value go up or down?

#TurnUp the Bayes

Somehow we need to smartly (1) aggregate the thousands of possible cancer profiles and (2) choose which genes should be considered in our profile (can't have no cat genes...)

For that, we're going to use Bayes' Theorem!

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

Great. You got it. Three practice problems:

1a. P(my cat is meowing | it's sunny) = 0.65
P(sunny) = 0.8
P(my cat is meowing) = 0.75
P(it's sunny | my cat is meowing) = ???

1b. P(my cat is meowing | it's sunny) = 0.30
P(sunny) = 0.8
P(my cat is meowing | it's not sunny) = 0.30
P(it's sunny | my cat is meowing) = ???

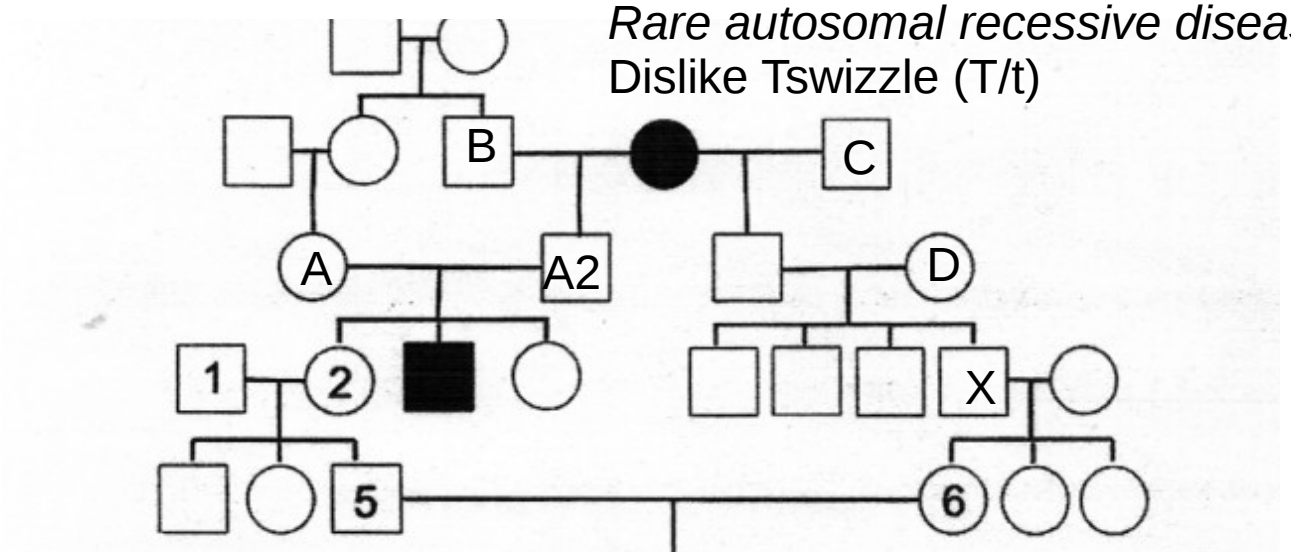
2. P(heads | tails) = 0.5
P(heads) = 0.5
P(tails) = 0.5
P(tails | heads) = ???

hint use common sense too

3. The blood tests came in. You tested positive for Senioritis. About 20% of seniors have Senioritis, and 80% of all blood tests turn out positive. And if you have Senioritis, the probability the test actually shows it is 95%. What's the probability you got senioritis, given your blood test?

I've got a blank square, baby...

Reviewing your pedigree game: how do you read one of these?



Rare autosomal recessive disease
Dislike Tswizzle (T/t)

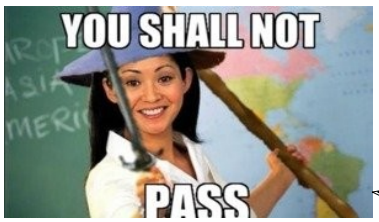
Warm-up:

What does it mean for disease to be rare, autosomal, rec.?
What is genotype of D, A, A2, and X?
Probability A is a carrier? 1 is a carrier? 2 is a carrier?

What's the probability 6 is a carrier?

What's the probability 5 is a carrier?

What's the probability that 5 and 6's child likes TSwizzle?

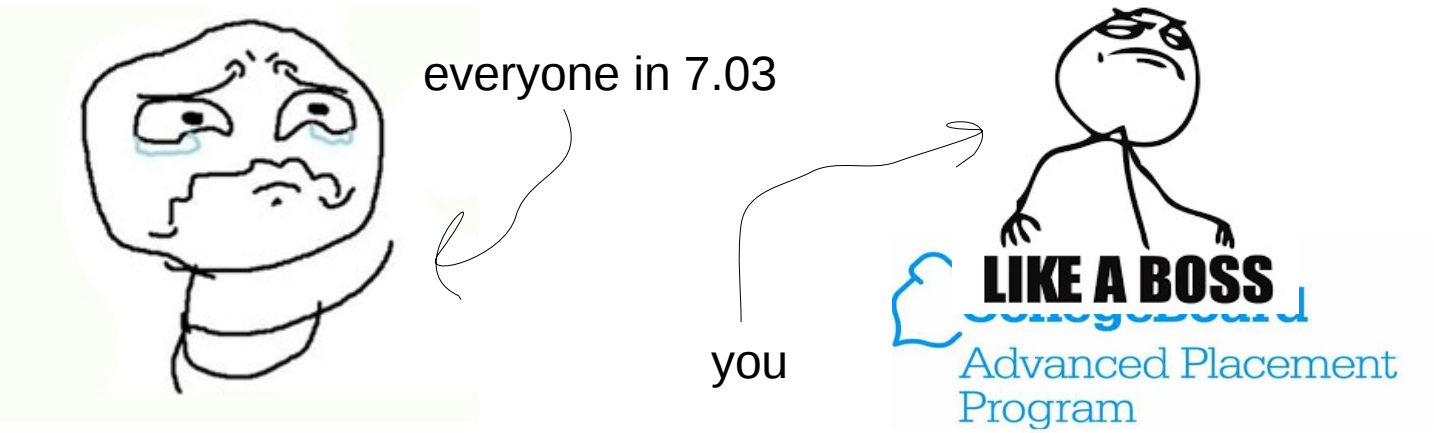


my genetics professor
this exact pedigree was on our midterm
really made people cry

5 and 6 (Are Never Ever Getting Back Together)

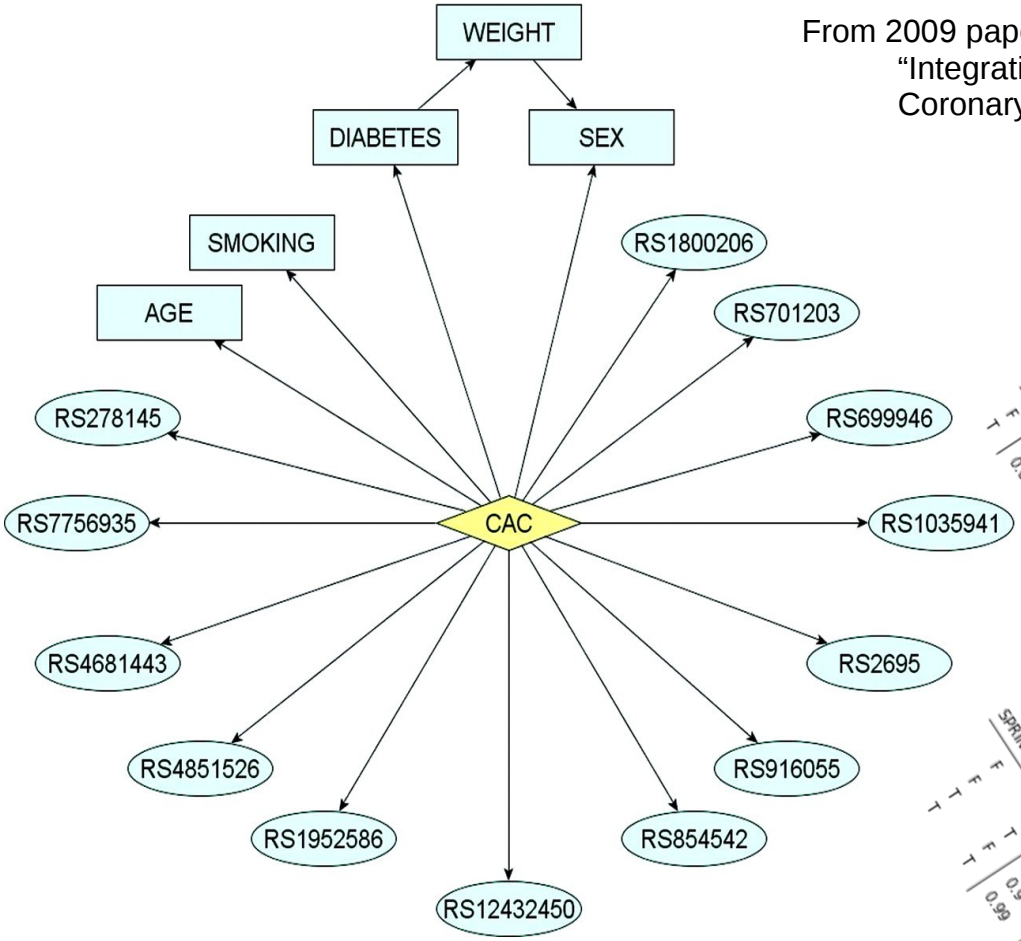
If 5 and 6 have three daughters that like TSwizzle, what is the probability that their next child will dislike TSwizzle?

hint hint hint
use Bayes Theorem P(A|B)
A = both parents are carriers, B = three unaffected daughters

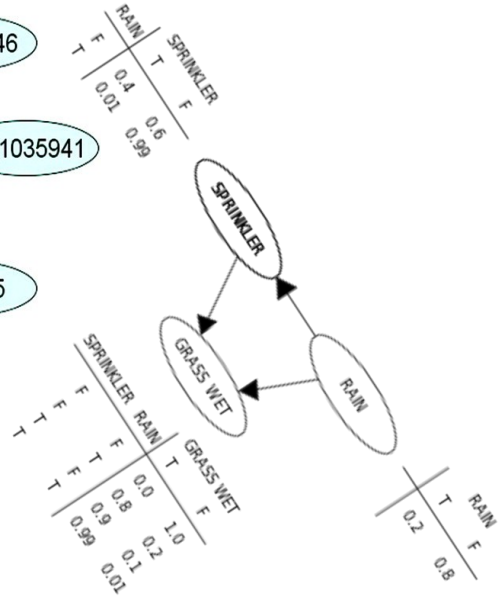


Back to the Bayesics...

Imagine probabilistically relating causal factors to a outcome. Picture this as a network:



From 2009 paper:
"Integrative Predictive Model of Coronary Artery Calcification"



What is great about bayes net:

- Graph topology can capture biological meaning
- Easy to detect genes with little correlation to phenotype
- Powerful prediction P(disease | factor values = …)
- Easy to use, plug and chug

What is not so great:

- Requires extensive data sets to train and get network created

---

2-second Intro on How To Get Started:

1. Grab a dataset for your disease of choice:

http://www.ncbi.nlm.nih.gov/gds/

2. Plug into a bayes-net creation framework:

http://docs.opencv.org/modules/ml/doc/normal_bayes_classifier.html
http://www.bayespy.org/intro.html