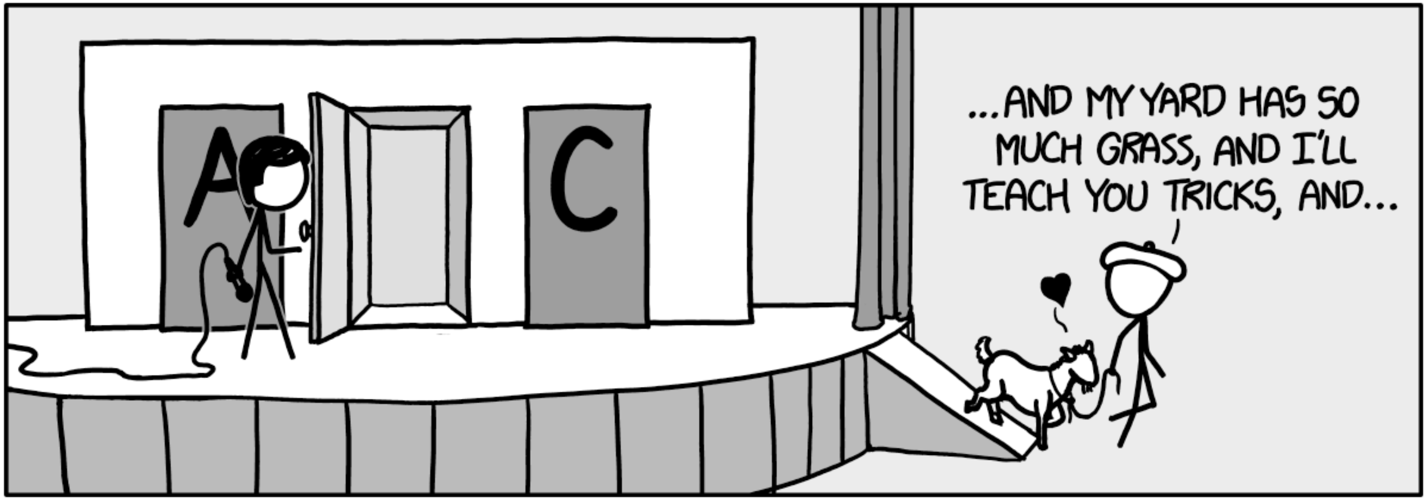
Stat 88: Probability and Mathematical Statistics in Data Science



https://xkcd.com/1282/

Lecture 5: 1/29/2021

Bayes' Rule: definition, use, and interpretation, Independence

Sections 2.2, 2.3, 2.5

Shobhana M. Stoyanov

Agenda

- Section 2.2: Symmetries in sampling: go over examples
- 2.3: Bayes' rule: complete tree diagram, and write down the formula
- 2.4: Use and interpretation of Bayes' rule
- 2.5: Independence

- Note: as mentioned last time, beware of the difference between conditional and unconditional probability

- P(2^{nd} card is QH) =1/52.
- P(2^{nd} card is QH | 1^{st} card is AD) = 1/51

Product rule and counting

- Recall product rule for counting, if there are sequences constructed in n stages, with k_i options at each stage, then the total number of sequences is $k_1 \times k_2 \times \cdots \times k_n$
- Count the number of outcomes for each stage and multiply them.
- **Example**: The English language has 26 letters. 5 letters are chosen with replacement. What is the chance that the middle three letters are all different and the first and last are the same as each other, and also the same as one of the three middle letters.

$3 \cdot 26 \cdot 25 \cdot 24 \cdot 1 \leftarrow$ # of sequences

$3 \cdot 26 \cdot 25 \cdot 24 \quad 1$

$\text{Prob} = \frac{3 \cdot 26 \cdot 25 \cdot 24}{\#(\Omega)}$

$\#(\Omega) = 26 \cdot 26 \cdot 26 \cdot 26 \cdot 26 = 26^5$

Symmetries in cards

w/o replacement.

std 52-card deck unless o/w stated

- Deal 2 cards from top of the deck.
  - How many possible sequences of 2 cards? $= 52 \cdot 51$
  - What is the chance that the second card is red? $\frac{26}{52}$
- P(5^{th} card from top is red) $= \frac{26}{52}$

Read this example in the text to see the multiple ways of computing this prob.

- 21^{st} card *and* 35^{th} cards are red = $P(R_{21} \cap R_{35})$ = (write it using conditional prob)

$$P(R_{21} R_{35}) = P(R_{21}) P(R_{35} | R_{21}) = \frac{26}{52} \cdot \frac{25}{51}$$

- P(7^{th} card is a queen) $\frac{4}{52}$

$$P(21^{st} \text{ is } R \mid 35^{th} \text{ is } R)$$

- $P(B_{52} \mid R_{21} R_{35})$

Exercise

Let A = 21^{st} card is Red
B = 35^{th} card is red

$P(AB)$

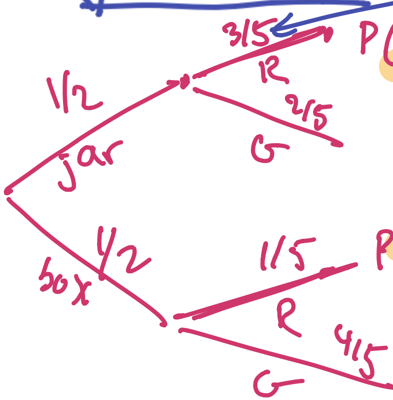
Prior and Posterior probabilities

- The ***prior*** probability of drawing the box = _0.5__ (before we knew anything about the balls drawn)

- The ***posterior*** probability of drawing the box = $\frac{1}{4}$ (this is after we *updated* our probability, *given* the information about which ball was drawn)

prior prob.

Select a container at random, and then select a ball from the container at random. If the ball is red, what is the probability that the box was picked?



LIKELIHOODS

$P(R / jar) = \frac{3}{5}$

$P(Jar \& R) = P(J)P(R|J) = (\frac{1}{2})(\frac{3}{5}) = \frac{3}{10}$

$P(box \& R) = P(B)P(R|B) = \frac{1}{2} \cdot \frac{1}{5} = \frac{1}{10}$

$P(R) = P(Box \& R) + P(Jar \& Red)$

$$P(box \mid R) = \frac{P(Box \,\&\, R)}{P(R)} = \frac{\frac{1}{10}}{\frac{1}{10} + \frac{3}{10}} = \boxed{\frac{1}{4}}$$

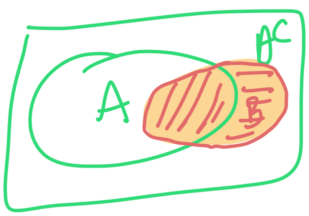
$$P(AB) = P(A)\,P(B \mid A) \qquad \underline{\text{Mult. rule}}$$

$$P(B \mid A) = \frac{P(AB)}{P(A)} \longrightarrow P(A) \neq 0$$

$$P(\overset{A}{Box} \mid \overset{B}{Red}) = \frac{P(Box \,\&\, R)}{P(R)} \qquad \frac{P(AB)}{P(B)}$$

$$P(A \mid B)$$

$$P(B) = P(AB) + P(A^cB) = \frac{P(Red \mid Box)\,P(Box)}{P(R \,\&\, Box) + P(R \,\&\, Jar)} \qquad \frac{P(B \mid A)\,P(A)}{P(AB) + P(A^cB)}$$



$$= \frac{P(B \mid A)\,P(A)}{P(B \mid A)P(A) + P(B \mid A^c)\,P(A^c)}$$

$$= \frac{P(R \mid Box)\,P(Box)}{P(R \mid Box)P(Box) + P(R \mid Jar)P(Jar)}$$

BAYES' RULE

Computing Posterior Probabilities: Bayes' Rule

- We want the *posterior* probability. That is, the conditional prob for the first stage, given the second. $P(\text{Box} \mid \text{Red})$

- Division rule (for conditional probability) = $P(A|B) = \frac{P(AB)}{P(B)}$, $P(B) \neq 0$

- Using the multiplication rule on P(AB), we get:

$$P(AB) = P(B|A)\,P(A)$$

- Rule first written down by Rev. Thomas Bayes in the 18th century. Helps us compute posterior probability, given prior prob. And likelihoods (which are conditional probabilities for the *second* stage given the first)

Exercise 2.6.9

A: Is event widget is acceptable
M_1 : Machine 1, M_2 : Machine 2.

A factory has two widget-producing machines. Machine I produces 80% of the factory's widgets and Machine II produces the rest. Of the widgets produced by Machine I, 95% are of acceptable quality. Machine II is less reliable – only 85% of its widgets are acceptable.

Suppose you pick a widget at random from those produced at the factory.

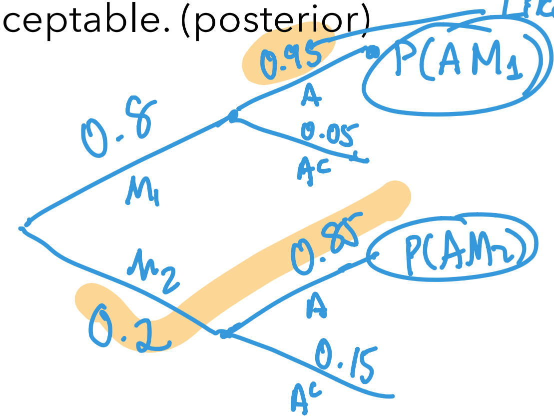
a) Find the chance that the widget is acceptable, given that it is produced by Machine I. (likelihood)

$P(A \mid M_1) = 0.95$

b) Find the chance that the widget is produced by Machine I, given that it is acceptable. (posterior)

Likelihood

$$P(A \mid M_1) = \frac{P(AM_1)}{P(M_1)}$$



$$P(M_1 \mid A) = \frac{P(AM_1)}{P(A)}$$

posterior

$$= \frac{P(A \mid M_1)P(M_1)}{P(A \mid M_1)P(M_1) + P(A \mid M_2)P(M_2)}$$

$$P(M_1 \mid A) = \frac{(0.8)(0.95)}{(0.8)(0.95) + (0.2)(0.85)}$$

Example: Binge drinking & Alcohol related accidents

(This example is from the text *Intro Stats* by De Veaux, Velleman, and Bock)

For men, binge drinking is defined as having 5 or more drinks in a row and for women as having 4 or more drinks in a row. (The difference is because of the average difference in weight.)

According to a study by the Harvard School of Public Health (H.Wechsler, G. W. Dowdall, A. Davenport, and W. Dejong, "*Binge Drinking on Campus: Results of a National Study*"):

- 44% of college students engage in binge drinking,37% drink moderately, and 19% abstain entirely.
- Another study, published in American journal of Health Behavior, finds that among binge drinkers aged 21 to 34, 17% have been involved in an alcohol-related automobile accident, while among nonbingers of the same age, only 9% have been involved in such accidents.
- Given that a student has been in a car crash, what is the chance that they were a binge drinker?

2.4: Use and interpretation of Bayes' rule

- Harvard study: 60 physicians, students, and house officers at the Harvard Medical school were asked the following question:
- "If a test to detect a disease whose ***prevalence*** is 1/1,000 has a false positive rate of 5 per cent, what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person's symptoms or signs?"

- What is your guess – without any computations?

- Prevalence aka Base Rate = fraction of population that has disease.
- False positive rate: fraction of positive results among people who don't have the disease
- Positive result: test is positive

Tree diagram for disease and positive test

- P(D|pos. test) or *posterior* probability =
- Recall that prior probability = 0.001 = 0.1%

Base Rate Fallacy

- P(D|pos. test) or *posterior* probability =
- Recall that prior probability = 0.001 = 0.1%

- P(+ test) = P(+ & disease) +P(+ & no disease)

- Base rate fallacy: Ignore the base rate and focus only on the likelihood. (Moral of this story: ignore the base rate at your own peril)

- Note: Want P(D|+) but most people focus on the test giving correct results for negative tests 95% of the time, that is P(no disease| neg)

- What happens to posterior probability if we change prior probability?

Section 2.5: Independence

If A & B are indep

$P(A|B) = P(A)$

$P(B|A) = P(B)$

- Multiplication rule: $P(AB) = P(A|B)P(B) = P(A)P(B)$

$$P(A \text{ and } B) = P(A) \times P(B|A)$$
$$P(A \text{ and } B) = P(B) \times P(A|B)$$

- Note that we are assuming that *A* and *B* have nonzero probabilities.

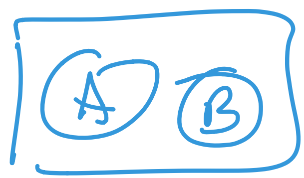
- Events A and B are ***independent*** if the information that one of them occurred does not change the chance of the other.

- To check if two events are independent (of each other), we can check if:

$$P(B|A) = P(B) \text{ or } P(A|B) = P(A)$$

- Or we can check if:

$$P(AB) = P(A) \times P(B)$$



Suppose $AB = \emptyset$, $P(A|B) = 0$

Case of Sally Clarke and SIDS

- Around 2003, Sally Clark, in a famous murder trial had two children one year apart who both died mysteriously. Sally Clarke's defence was that the babies both died of Sudden Infant Death Syndrome (SIDS)
- A = event the first child dies of SIDS
- B = event the second child dies of SIDS.
- Assumption: P(A) = P(B) = 1/8543 (based on stats, unconditional probability)

Mutually Exclusive vs. Independent

- Make sure you understand the difference; these are *very* different ideas, though both apply to pairs of events.

- "Mutually exclusive" events means that the occurrence of one **prevents** the occurrence of the other. (This means that it reduces the chance of the other occurring to 0.)

- "Independent" events means that the occurrence of one **does not change** the chance of the other occurring.

- Do NOT assume independence without justification.

Exercise: Do events have to be either Mut. Exc. or indep? Yes/No.

Let's make a deal!: The Monty Hall Problem

There are 3 doors, A, B, C, behind one is a new car (a Ferrari, say), and behind the other two are goats.

Now suppose you are the contestant, and you choose door A. Then Monty Hall opens one of the other two doors, say B, to show you a goat!

He asks you if you want to switch to C or stick with your original choice A, you say...?