* Announcement
① HW 1 due today 11:59 pm.
② HW 2 due next Wed (9/9)
③ Quiz 1 next Thursday (9/10)
(Chapter 1 and 2)
④ HW3
* Today: Lecture note 4 & 5 (stat88.org)

STAT 88: Lecture 5

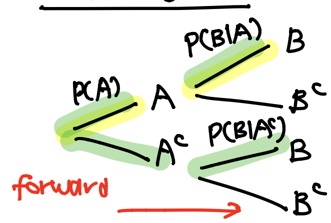
Contents

Last time

Sec 2.3 Bayes' Rule:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A^c)P(B|A^c)}.$$

Tree diagram



- $P(A)$ is called prior probability;
- Forward conditional → $P(B|A)$ is called likelihood;
- Backward conditional → $P(A|B)$ is called posterior probability.

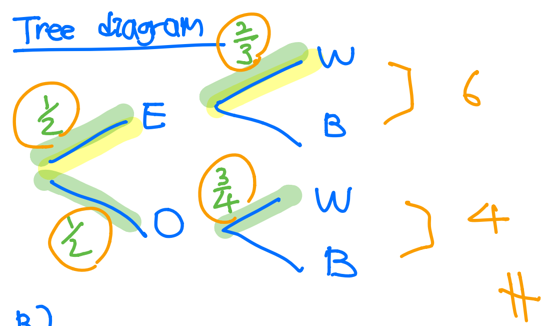
Base rate fallacy — according to Bayes' rule you can't ignore the base rate (i.e. prior probability) when computing the posterior probability.

Warm up: There are two boxes, the odd box containing 1 black marble and 3 white marbles, and the even bot containing 2 black marbles and 4 white marbles. A box is selected at random, and a marble is drawn at random from the selected box.

(a) What is the probability that the marble is black?

(b) Given that the marble is white, what is the probability that it came from the even box?

- O = Odd box
  E = Even box
  B = Black marbles
  W = White "



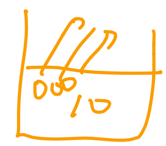
(a) $P(B) = P(E \cap B) + P(O \cap B)$

Multiplication $= P(E)P(B|E) + P(O)P(B|O)$

$= \frac{1}{2} * \frac{1}{3} + \frac{1}{2} * \frac{1}{4}$

$= \frac{1}{6} + \frac{1}{8} = \frac{7}{24} \neq \boxed{\frac{3}{10}}$

Assume



(b) $P(E|W) = \dfrac{\frac{1}{2} * \frac{2}{3}}{\frac{1}{2} * \frac{2}{3} + \frac{1}{2} * \frac{3}{4}} = \boxed{0.47}$

posterior.

$\|$

$$\frac{P(E \cap W) = P(E)P(W|E)}{P(E)P(W|E) + P(O)P(W|O)}$$

2.4. Use and Interpretation (continued)

Example: There are three boxes, each of which contains two coins. One box has two gold coins, one has two silver coins, and one has a gold coin and a silver coin. A box is picked at random and then a coin is picked at random from the box. Given that the coin is gold, what is the chance that the other coin in the box is silver?

Exercise (3 events)

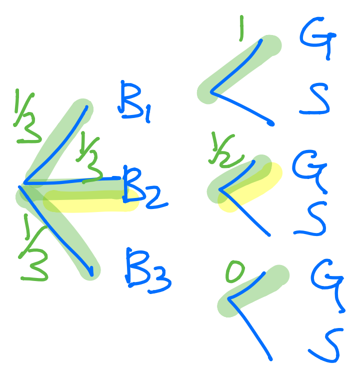
B_1 = box 1

B_2 = box 2

B_3 = box 3

G = gold coins

S = silver coins



$$P(B_2|G) = \frac{\frac{1}{3} * \frac{1}{2}}{\frac{1}{3} * 1 + \frac{1}{3} * \frac{1}{2} + \frac{1}{3} * 0} = \frac{1}{3}$$

2.5. Independence

Events A and B are independent if the information that one of them occurred does not change the chance that the other occured, in other words,

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

Example: Deal 2 cards from a deck. Let A be the event that 1st Card is an ace and B be the event that 2nd Card is an ace. Then $P(B|A) = \boxed{3/51}$ ~~and~~ $\neq P(B) = \boxed{4/52}$, so A and B are not independent. However, if deal 2 cards with replacement, then A and B are independent.

symmetry in simple random sampling

$\parallel$ $P(A)$

A Special Case of the Multiplication Rule If $A, B \subseteq \Omega$ are independent,

$$P(A \cap B) = \underbrace{P(B|A)P(A)}_{\text{Multiplication rule}} = P(B)P(A).$$

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

Mutually exclusive VS Independent

Two events cannot occur at same time, i.e. $A \cap B = \emptyset$

Occurence of A not affected by occurence of B

Ex1 A = Event that it rains in Berkeley in CA

B = Event that it rains in Chicago

~~Mutually exclusive~~ or Independent ?

Ex2 A = Event that it rains in Berkeley in CA

B = Event that it doesn't rain in Berkeley

Mutually exclusive

Ω ← Venn Diagram

Mutually exclusive

Note. If A and B are mutually exclusive, then A,B are dependent since if you have A, you don't have B which is a dependency between A and B.

Ex3 A = it rains today
B = " tomorrow. > independence : no
mutually exclusive : no
$A \cap B$ = it rain today "and it will rain tom.

Example: Around 2003, Sally Clarke, 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.

Assume $P(A) = P(B) = 1/8543$. A medical expert witness said the chance of two babies dying of SIDS is $1/8543^2 = 1/72982849 \approx 0$ and hence Sally Clarke must have murdered her babies. "$P(A)P(B) = P(A \cap B)$

What is the problem with argument? The assumption of independence is hard to justify

There may be genetic or environmental factors that predispose families to SIDS so that the second case within the family becomes much more likely. Hence $P(B|A) \cancel{\approx} > 1/8543$ and $P(A \cap B) = P(A)P(B|A) \cancel{\approx} > 1/8543^2 = 1/72982849$.

$P(B)$ small $\rightarrow$ $P(B|A) > P(B)$

Moral of the story

Don't just assume independence w/o justification.

Example: A population consists of equal numbers of students in 4 Categories: freshman, sophomore, junior and senior. 4 people are drawn with replacement from the population. What is the chance that a member from each category is chosen?

F = freshman, So = Sophomore

Method 1

J = junior, Se = senior

$$P(\text{1st F, 2nd So, 3rd J, 4th Se})$$

$$= P(\text{1st F})\,P(\text{2nd So})\,P(\text{3rd J})\,P(\text{4th Se})$$

independence

$$= \frac{1}{4} * \frac{1}{4} * \frac{1}{4} * \frac{1}{4} = \frac{1}{4^4}$$

permutations = 4! (factorial)

$$\Rightarrow P(\text{Member from each category}) = \frac{4!}{4^4}$$

Method 2

$$\frac{4}{4} * \frac{3}{4} * \frac{2}{4} * \frac{1}{4} = \frac{4!}{4^4}$$

Example: (Exercise 2.6.5) There are n students in a class. Assume that each student's birthday is equally likely to be any of 365 days of the year, regardless of the birthdays of others.

(a) What is the chance that at least one of the students was born on January 1?

(b) What is the chance that at least two students have the same birthday?

(a) 1 − P(no one born Jan 1.)

$$= 1 - \left(\frac{364}{365}\right)^n$$

(b) 1 − P(no student same b-day)

$$= 1 - \frac{365}{365} * \frac{364}{365} * \cdots * \frac{365-n+1}{365}$$

Example: Suppose A and B are two events with

$$P(A) = 0.5 \text{ and } P(A \cup B) = 0.8.$$

For what values of $P(B)$ would A and B be independent?

1. $P(B) = 0$
2. $P(B) = 0.3$
3. $P(B) = 0.6$
4. none of the above

Inclusion - Exclusion.

$$\underbrace{P(A \cup B)}_{= 0.8} = \underbrace{P(A)}_{= 0.5} + P(B) - \underbrace{P(A \cap B)}_{= P(A)P(B)}$$

Independent

$P(A) = 0.5$

$\Rightarrow \quad 0.3 = P(B) - 0.5\,P(B)$

$\Rightarrow \quad P(B) = \frac{3}{5} = 0.6$