Session 78: Bayes’ Theorem

- Bayes’ Theorem
- Generalized Bayes’ Theorem

Motivation for Bayes’ Theorem

- Bayes’ theorem allows us to use probability to answer questions such as the following:
  - Given that someone tests positive for having Corona, what is the probability that they actually do have Corona?
  - Given that someone tests negative for Corona, what is the probability, that in fact they do have Corona?
- Bayes’ theorem has applications to medicine, law, artificial intelligence, engineering, and many diverse other areas.

Bayes' Theorem

Bayes' Theorem: Suppose that *E* and *F* are events from a sample space S such that $p(E) \neq 0$ and $p(F) \neq 0$. Then:

$$p(F \mid E) = \frac{p(E \mid F)p(F)}{p(E \mid F)p(F) + p(E \mid \bar{F})p(\bar{F})}$$

or alternatively (a more common formulation)

$$p(F \mid E) = \frac{p(E \mid F)p(F)}{p(E)}$$

Typical use of Bayes Theorem

E = Testing positive on Corona

F = Having Corona

$$P(F|E) = \frac{P(E|F)\,P(F)}{P(E)}$$

posterior probability	$P(F\|E)$ what is the probability that I have Corona (don't know)
$\uparrow$	
likelihood	$P(E\|F)$ probability of testing positive when having Corona
$\uparrow$	
evidence	$P(E)$ probability of testing positive (both known from clinical studies)
$\uparrow$	
prior probability	$P(F)$ probability of having Corona (known from monitoring) (what I assume before the test)

Proof of Bayes Theorem

- $P(F|E) \overset{\text{Def.}}{=} \dfrac{P(E\cap F)}{P(E)} = \dfrac{P(E\cap F)\cdot P(F)}{P(E)\cdot P(F)} = \dfrac{P(E\cap F)}{P(F)}\cdot\dfrac{P(F)}{P(E)} \overset{\text{Def.}}{=} \dfrac{P(E|F)\cdot P(F)}{P(E)}$

- $P(E) = P\left((E\cap F)\cup(E\cap \bar{F})\right) =$ (set identity)

  $= P(E\cap F) + P(E\cap \bar{F}) =$ ($F, \bar{F}$ are disjoint)

  $= P(E|F)P(F) + P(E|\bar{F})P(\bar{F})$ (Def. cond. prob.)

- $$P(F|E) = \frac{P(E|F)\cdot P(F)}{P(E|F)P(F) + P(E|\bar{F})P(\bar{F})}$$

Example

Suppose that 2 persons in 100 have a Covid-19. There is a test for the disease that gives a positive result 95% of the time when given to someone with the disease. When given to someone without the disease, 98% of the time it gives a negative result. Find

a) the probability that a person who test positive has Covid-19.
b) the probability that a person who test negative does not have Covid-19.

Should someone who tests positive be worried?

Example

Let *D* be the event that the person has the disease, and *E* be the event that this person tests positive.

We know that

$$p(D) = 0.02,\ p(E \mid D) = 0.95, p\left(E \mid \bar{D}\right) = 0.02$$

$$p(\bar{E} \mid D) = 0{,}05,\ p(\bar{E} \mid \bar{D}) = 0.98$$

We want to know p(D|E)

$$p(D \mid E) = \frac{p(E \mid D)\, p(D)}{p(E)} = \frac{0{,}95 \cdot 0{,}02}{0.95 * 0.02 + 0.02 * 0.98} \approx 0.49$$

therefore $p(\bar{D} \mid E) \approx 0.51$ (false positives)

Example

And if the result is negative?

$$P(D|\bar{E}) = \frac{P(\bar{E}|D) \cdot P(D)}{P(\bar{E}|\bar{D}) \cdot P(\bar{D}) + P(\bar{E}|D) \cdot P(D)} =$$

$$= \frac{0.05 * 0.02}{0.98 * 0.98 + 0.05 * 0.02} \approx 0.001$$

false negatives are rare

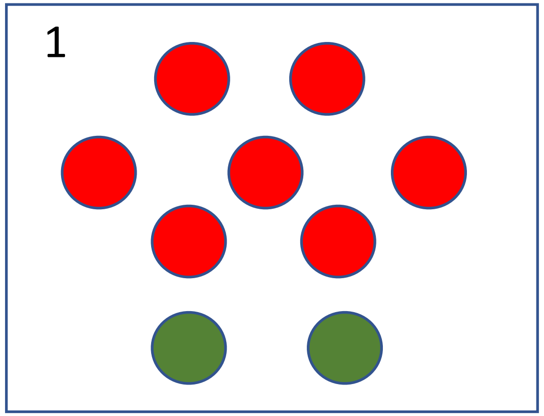
Example

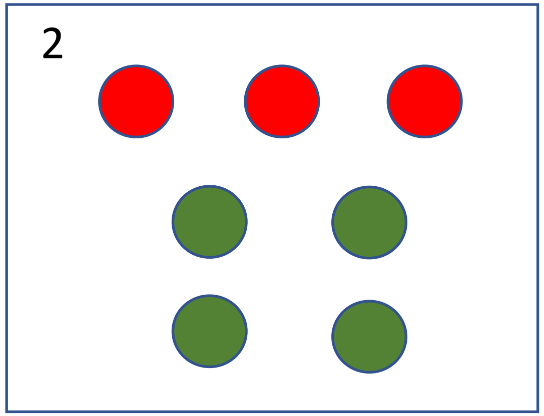
We have two boxes. The first box contains two green balls and seven red balls. The second contains four green balls and three red balls.

Bob selects one of the boxes at random.

Then he selects a ball from that box at random.

If he has a red ball, what is the probability that he selected a ball from the first box?





Evidence E : red ball , $p(E) = ?$

Unknown probability F : Box 1? $\quad p(F) = \frac{1}{2}$ (prior)

Likelihood : $P(E|F) = \frac{7}{9}$

$P(E|\overline{F}) = \frac{3}{7}$ $\qquad P(E) = P(E|F)P(F) + P(E|\overline{F})P(\overline{F})$

Determine the posterior probability:

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|\overline{F})P(\overline{F})} = \frac{\frac{7}{9}\cdot\cancel{\frac{1}{2}}}{\frac{7}{9}\cdot\cancel{\frac{1}{2}} + \frac{3}{7}\cdot\cancel{\frac{1}{2}}} =$$

$$= \frac{\frac{7}{9}}{\frac{49+27}{63}} = \frac{63}{76}\cdot\frac{7}{9} \approx 0.645$$

Generalized Bayes' Theorem

Generalized Bayes' Theorem: Suppose that *E* is an event from a sample space *S* and that F_1, F_2, …, F_n are mutually exclusive events such that $\bigcup_{i=1}^{n} F_i = S$

Note: before $S = F \cup \overline{F}$

Assume that $p(E) \neq 0$ for $i = 1, 2, \ldots, n$. Then

$$p(F_j \mid E) = \frac{p(E \mid F_j)p(F_j)}{\sum_{i=1}^{n} p(E \mid F_i)p(F_i)}$$

Summary

- Bayes’ Theorem
  - Testing diseases
- Generalized Bayes’ Theorem

Monty Hall Puzzle

3 doors – 1 price

① Player selects one door

② Host opens one of the remaining doors that has no price

③ Player can switch the door. Should she?

Let pr denote the event that the door the player selected contains a price

Winning probability without switching : $P(win_1) = P(pr) = \frac{1}{3}$

Winning probability with switching :

$$P(win_2 \mid pr) = 0 \qquad \text{(both other doors have no price)}$$

$$P(win_2 \mid \overline{pr}) = 1 \qquad \text{(the host opens the door without price)}$$

$$P(win_2) = P(win_2 \mid pr)\cdot P(pr) + P(win_2 \mid \overline{pr})\, P(\overline{pr}) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}$$