

Question continue : September 26

a) $S = BC \cup B^c$ - disjoint

\therefore by the law of total Prob

$$P(\text{Pos}) = P(\text{Pos} | B_1) P(B_1) + P(\text{Pos} | B_2) P(B_2)$$

so replacing with (A, B)

$$P(A) = P(A | B_1) P(B_1) + P(A | B_2) P(B_2)$$

// we know

$$\begin{aligned} P(\text{Pos} | B_1) &= 0.95 \\ P(\text{Neg} | B_1^c) &= 0.95 \\ P(B_1) &= 5/1000 \end{aligned}$$

$$\therefore P(\text{Pos}) \text{ or } P(A) = 0.95 (5/1000) + (1 - 0.95) (1 - 5/1000)$$

ii // Here we want $P(B_1 | \text{Pos})$, which is a reversal of $P(\text{Pos} | B_1)$ [given above]

$$\begin{aligned} \text{we have } P(B_1 | \text{Pos}) &= \frac{P(\text{Pos} | B_1) P(B_1)}{P(\text{Pos} | B_1) P(B_1) + P(\text{Pos} | B_1^c) P(B_1^c)} \\ &= \frac{0.95 * 5/1000}{0.95 * 5/1000 + (1 - 0.95) * (1 - 5/1000)} \end{aligned}$$

NOTE: In word problems, conditional prob is suggested
Prof if you can reasonably use the word "given"
to describe the probability.

Notes on Bayes theorem:-

- 1) In a diagnostic test such that above (example $P(\text{Pos} | \text{Disease})$) is called the sensitivity of the test.
 $P(\text{Neg} | \text{Disease}^c)$ is called specificity of the test.
- 2) $P(\text{Disease} | \text{Positive})$ = positive predictive value of the test.
 - Depends on not only the sensitivity and specificity of test, but also on prevalence of the diseases.

Further Jargon:-

- In the description of Bayes's theorem, the probabilities B_1, B_2, \dots, B_m (ie $P(B_1), P(B_2), \dots, P(B_m)$) are called prior probabilities, since they are found in absence of current data.
- Thus we have prior probabilities of breast cancer before we see the result of the test.
- The probabilities $P(B_i | A)$ are called posterior prob. Think A as representative of current data.
- In this light, Bayes Theorem tells us how to update prior prob \rightarrow to \rightarrow posterior

Statistical Independence:-

- Sometimes, knowledge that an event A has occurred, does not affect our calculations of the probability that B will occur.
- For example in the marble problem, if we draw 2 marbles with replacement then $P(G_2 | R_1)$ is just $P(G_2)$; since knowing that drawing a red marble at first doesn't affect drawing green.
- This idea translates into an immediate definition (2)

Definition 1:- Two events A & B are said to be independent if $P(B|A) = P(B)$ or $P(A|B) = P(A)$

Notes:-

- 1) It's hard to extend this natural definition to more than two events.

o We therefore give a second definition that is easily extended to arbitrary set of events: $A_1, A_2, A_3, \dots, A_n$

↳ Definition 2:- The events A & B are independent if $P(A \cap B) = P(A) P(B)$

Theorem to show the Def 1-2 are equal

- Two events are independent according to Def 1 \Leftrightarrow they are independent according to Def 2

Proof

- Let A & B be independent by ~~Def 1~~ Def 1
Then $P(B|A) = P(B)$. But $P(B|A) = P(A \cap B) / P(A)$
 $\therefore P(B) = P(A \cap B) / P(A) \Rightarrow \underbrace{P(A \cap B)}_{\text{Def 2}} = P(B)P(A)$



- 2) We can now extend definition of Independence to more than 2 events.

Def:- The events A_1, A_2, \dots, A_n are mutually independent if

$$P\left(\bigcap_{j \in I} A_j\right) = \prod_{j \in I}^k P(A_j) \quad \text{symbol for product}$$

for all subsets $A_{i_1}, A_{i_2}, \dots, A_{i_k}$ of

all A_1, \dots, A_n for all k

Remember k are subsets.

3 We say A_1, A_2, \dots, A_n are pairwise independent if $P(A_i \cap A_j) = P(A_i)P(A_j)$ for all pairs.

o It is possible to show that pairwise independence doesn't imply mutual independence

4) Very IMP: Independence arises in two ways

1) We have data and we want to check if certain events are independent

Eg: Is recovery after abdominal surgery independent of the temp of operating room?

So Checking independence

Q) (Most IMP way) if we can assume independence, based on our knowledge of basic science of the problem, then, calculations of probability involving intersections are rendered much easier

o All we need to know is $P(A_1)P(A_2)\dots P(A_n)$

Example, when you toss a fair coin twice, it is reasonable to assume that the outcomes on the two tosses are independent. Because we can't think of any reasonable way in which they might be dependent.

Then we can easily compute $P(H_1 \cap T_2) = P(H_1)P(T_2) = 1/2 \cdot 1/2 = 1/4$, & all other outcomes are also easily found to have prob $1/4$. If you were to find $P(A \cap B)$