Lecture 6 - Conditional Probability

Ex. Conditional probability.

Survey W&M student about political afil. and gender

gender \ afil	A	B	margin
men	501	238	739
women	782	123	905
		361	1644

Q: If I randomly select a student what is the prob. they are a woman?

$$\mathbb{P}(\text{woman}) = 905/1644 \approx 55\%$$

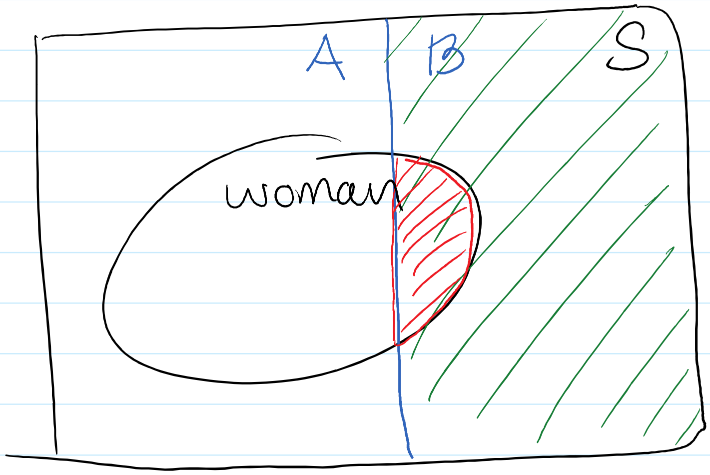
Q: Given that the student is a member of party B, what is the prob. they are a woman?

conditionally (given) this prob. on info that the student is in party B

conditional prob.

$$\mathbb{P}(\text{woman GIVEN B}) = \frac{123}{361} \approx 34\%$$

Venn Diagram



Q1: $\mathbb{P}(\text{woman}) = \dfrac{\text{area of woman}}{\text{area of } S}$

Q2: $\mathbb{P}(\underset{\sim\sim}{\text{Woman}} \ \underset{\sim\sim}{\text{Given}} \ B)$

$$\frac{\text{woman} \cap B}{B} = \frac{\text{area of } \cancel{D}}{\text{area of } B}$$

Defn : Conditional Probability

If $A, B \subset S$ then if $\mathbb{P}(B) > 0$,

$$\mathbb{P}(A \mid B) \overset{\text{def}}{=} \frac{\mathbb{P}(AB)}{\mathbb{P}(B)} .$$

read: "given"

all together: prob. of A given B

Facts: Henceforth assume $P(B) > 0$

① $\boxed{P(B|B) = 1}$

Pf. $P(B|B) = \frac{P(BB)}{P(B)} = \frac{P(B)}{P(B)} = 1.$

② If A and B are disjoint ($AB = \emptyset$)

$\boxed{P(A|B) = 0}$.

Pf. $P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(\emptyset)}{P(B)} = \frac{0}{P(B)} = 0.$

Ex. Roll two dice.

Q: What is the prob. the first die is a 2 given the sum of the dice is ≤ 5.

A (first die is a 2) $\quad$ B (the sum of the dice is ≤ 5)

$S = \{(i,j) \text{ where } 1 \le i,j \le 6\}$ ← all equally likely

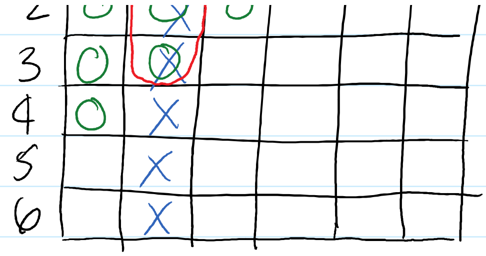
$|S| = 36 = 6^2$

first

second \ first	1	2	3	4	5	6
1	O	X	O	O		
2	O	X	O			
3		X				

$P(\ \ |\ \ )$ $\quad$ $P(AB)$ ← first is 2

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A\underbrace{B}_{\text{sum} \le 5})}{\mathbb{P}(B)}$$



$$= \frac{|AB|/|S|}{|B|/|S|} = \frac{|AB|}{|B|} = \frac{3}{10}$$

<u>Theorem:</u> Conditional probability defines a valid prob. fn.

Assume we condition on $B \subseteq S$, then for any event $E \subseteq S$

$$\underbrace{\mathbb{P}_B(E) = \mathbb{P}(E|B) = \frac{\mathbb{P}(EB)}{\mathbb{P}(B)}}_{\text{fix } B\text{, this is a fn of } E \subseteq S}$$

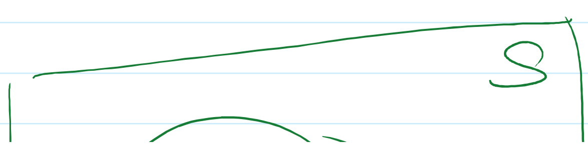
i.e. $\mathbb{P}_B : \mathcal{P}(S) \longrightarrow \mathbb{R}$

Convince yourself $\mathbb{P}_B$ satisfies the Kolmogorov axioms.

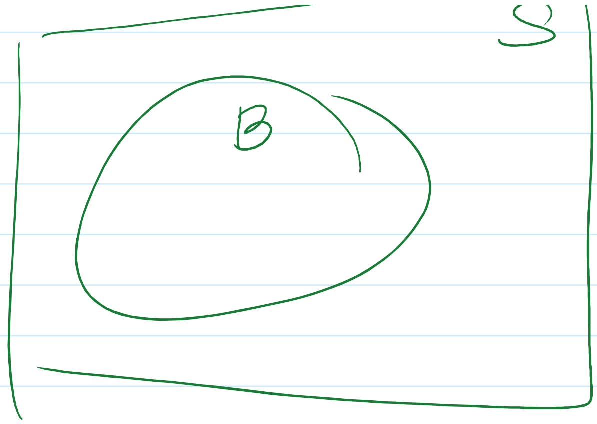
✓ (1) $\mathbb{P}_B(E) \geq 0$ <u>pf.</u> $\mathbb{P}_B(E) = \frac{\mathbb{P}(EB)}{\mathbb{P}(B)} > 0$

✓ (2) $\mathbb{P}_B(S) = 1$ <u>pf.</u> $\mathbb{P}_B(S) = \frac{\mathbb{P}(SB)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$

(3) . . . .



Basically:
re-define $\mathbb{P}$ so
that $\mathbb{P}(B) = 1$
$(\mathbb{P}(B|B) = 1)$



why do we care?
Fix B, I can manipulate $\mathbb{P}(\bullet \mid B)$
like any other prob. fn.

e.g. $\mathbb{P}(A^c|B) = 1 - \mathbb{P}(A|B)$

$$\mathbb{P}(A_1 \cup A_2|B) = \mathbb{P}(A_1|B) + \mathbb{P}(A_2|B) - \mathbb{P}(A_1A_2|B).$$

Theorem: Compound Probability

$$\mathbb{P}(AB) = \mathbb{P}(A|B)\,\mathbb{P}(B) = \mathbb{P}(B|A)\,\mathbb{P}(A).$$

pf. If $\mathbb{P}(B) > 0$ then $\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)}$

hence ... rearrange.

else $\mathbb{P}(B) = 0$ then $AB \subset B$

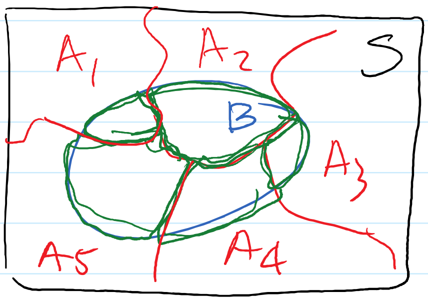
hence $\mathbb{P}(AB) \le \mathbb{P}(B) = 0$

So $\mathbb{P}(AB) = 0$.

So both sides are zero.

Theorem: Law of Total Probability

If (A_i) is a partition of S,



then for any $B \subset S$,

$$\mathbb{P}(B) = \sum_i \mathbb{P}(B|A_i)\,\mathbb{P}(A_i)$$

are of BA_i rel. to area of A_i

area of A_i rel. to area of S

product = area of BA_i rel. to area of S = $\mathbb{P}(BA_i)$

Pf. Compound prob theorem from above, says

$$P(B|A_i)P(A_i) = P(BA_i)$$

$$\sum_i P(B|A_i)P(A_i) = \sum_i P(BA_i)$$

$$= P(B)$$

partitions theorem in Lec. 2.

Ex. → Basket 1 → Basket 2

Game: (1) randomly select ball from Basket 1 and put in Basket 2

(2) randomly select from basket 2.

Q: what is the prob. of choosing a black ball on step (2)?

Let W = choose white on step (1)

W^c = " black "

Let B = choose black on step (2)

B^c " white "

Want: $P(B)$. Use Law of total prob.

Notice that $\{W, W^c\}$ partitions S.
(use in role of A_i)
$i = 1, 2$

Law of total prob:

$$P(B) = \underbrace{P(B|W)}_{1/2}\underbrace{P(W)}_{3/5} + \underbrace{P(B|W^c)}_{2/3}\underbrace{P(W^c)}_{2/5}$$

$$= \frac{3}{10} + \frac{4}{15}$$

$$= \frac{9}{30} + \frac{8}{30}$$

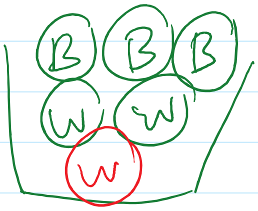
$$= \frac{17}{30}$$

$$> \frac{1}{2}$$

$P(B|W)$ = prob. choose Black on (2) given choose White on (1)

Basket 2 | W

Basket 2 $/W$



$$\boxed{P(B|W) = \tfrac{1}{2}}$$

$> \tfrac{1}{2}$

$P(B|W^c) =$

Basket 2 $/W^c$

$= 4/6 = 2/3 = P(B|W^c)$

<u>Theorem:</u> <u>Bayes' Theorem</u>

Way of calc. $P(A|B)$ using $P(B|A)$.

If $A, B \subset S$, $P(A), P(B) > 0$ then

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

<u>pf.</u>

$$P(A|B) \overset{\text{def.}}{=} \frac{P(AB)}{P(B)} = \frac{P(BA)}{P(B)} \overset{\text{compound prob.}}{=} \frac{P(B|A)P(A)}{P(B)}.$$

<u>Ex,</u> Continue Previous.

<u>Given</u> I choose a black ball on second,

what is the prob. I choose a white on (1)

$$P(W|B) = \frac{P(B|W)\,P(W)}{P(B)}$$

$$= \frac{(1/2)(3/5)}{(17/30)}$$

Theorem: Law of Total Prob + $\boxed{\text{Bayes'}}$

If (A_i) is a partition of S and $B \subset S$,

then

$$P(A_i|B) = \frac{P(B|A_i)\,P(A_i)}{\sum_j P(B|A_j)\,P(A_j)}$$

Pf: Law of Total Prob

Know: $P(B|A_j)$, $P(A_j)$

Calc:

$$\boxed{P(B) = \sum_j P(B|A_j)\,P(A_j)}$$

Bayes': Know: $P(B|A_i)$, $P(B)$, $P(A_i)$

Calc: $P(A_i|B) = \dfrac{P(B|A_i)\,P(A_i)}{P(B)}$.

$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i)\,\mathbb{P}(A_i)}{\mathbb{P}(B)}$.

Ex. → Covid 19 has a prevalence rate of 1%

D = have disease D^c = I don't

→ We test for covid 19 and get either a + or a −.

$\mathbb{P}(D) = .01$

→ The test accurately reports a + 95% of the time.

$\mathbb{P}(+|D) = .95$ (sensitivity)

→ The test accurately reports a − 99% of the time.

$\mathbb{P}(-|D^c) = .99$ (specificity)

I go and get a Covid 19 test. I get a +.

what is the prob this is correct?

$$\mathbb{P}(D|+) = \frac{\mathbb{P}(+|D)\,\mathbb{P}(D)}{\mathbb{P}(+|D)\,\mathbb{P}(D) + \mathbb{P}(+|D^c)\,\mathbb{P}(D^c)} = \cancel{\frac{(.95)(.01)}{(.95)(.01)+(.95)(.99)}}$$

correction:

$$= \frac{\mathbb{P}(+|D)\cdot\mathbb{P}(D)}{\mathbb{P}(+|D)\,\mathbb{P}(D) + (1-\mathbb{P}(-|D^c))(1-\mathbb{P}(D))} = \frac{(.95)(.01)}{(.95)(.01)+(.01)(.99)}$$

$= .49$

---

partition: $\{D, D^c\}$ partition S (play role of A_i). $\cancel{= .49}$

"$B = +$" $A_1 = D$, $A_2 = D^c$

$\mathbb{P}(A_i|B)$