# CONDITIONAL PROBABILITY & INDEPENDENCE

So far considered unconditional probabilities:

I.e. $P(A)$ computed with respect to a sample space $S$ with $P(S) = 1$.

Can sometimes update $S$ based on new information.

E.g. Recall: 12 passengers riding in groups of 4, 5, and 3 in 3 vehicles.

"$|A|$" means "cardinality", i.e. # of elements in set A

$$S = \{\text{All possible partitions}\}, \quad |S| = 27{,}720$$

We had

$$P(\text{You ride with friend}) = \frac{\overbrace{7{,}980}^{\#\{\text{partitions w/ You with friend}\}}}{\underbrace{27{,}720}_{\#\text{ partions}}} = 0.2789$$

Suppose: We know your friend is in group of 5 (in the van)

Gives new sample space $S^{new}$:

$$S^{new} = \{\text{partitions with friend in van}\}$$

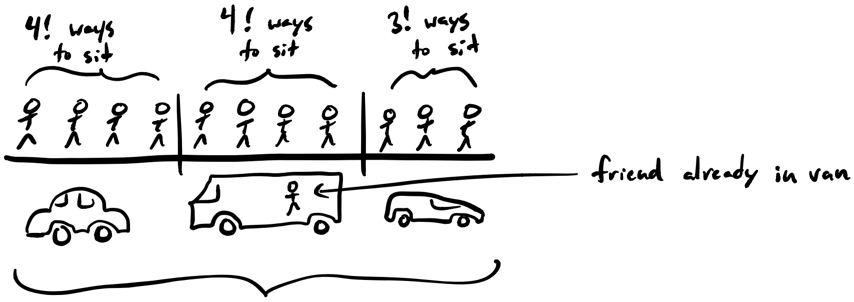
Now

$$P(\text{You ride with friend given friend in van}) = \frac{\#\left\{\begin{array}{c}\text{partitions with you with friend}\\ \text{and friend in van}\end{array}\right\}}{|S^{new}|},$$

where

$$|S^{new}| = \#\{\text{partitions with friend in van}\}$$

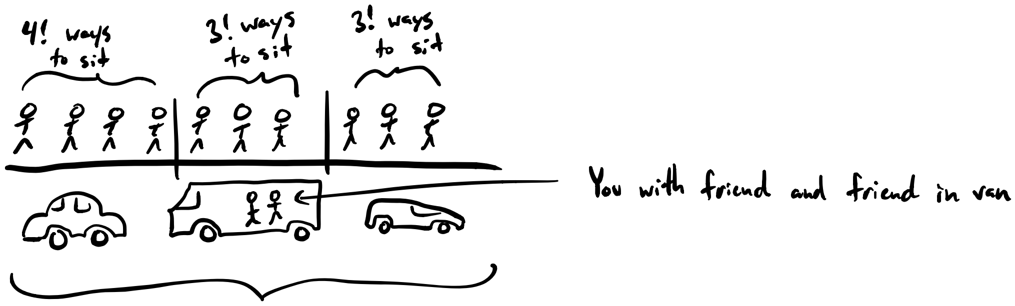
$$= \frac{11!}{4!\,4!\,3!} = \frac{11\cdot 10\cdot 9\cdot 8\cdot 7\cdot 6\cdot 5}{4\cdot 3\cdot 2\cdot 1\cdot 3\cdot 2\cdot 1} = 11\cdot 5\cdot 3\cdot 2\cdot 7\cdot 5 = 11{,}550$$




11! ways to sit in all 11 remaining passenger seats.

and

$$\#\left\{\begin{array}{c}\text{partitions with you with friend}\\ \text{and friend in van}\end{array}\right\} = \frac{10!}{4!\,3!\,3!} = \frac{10\cdot 9\cdot 8\cdot 7\cdot 6\cdot 5}{3\cdot 2\cdot 1\cdot 3\cdot 2\cdot 1} = 5\cdot 3\cdot 8\cdot 7\cdot 5 = 4{,}200$$




10! ways to sit in all 10! remaining passenger seats.

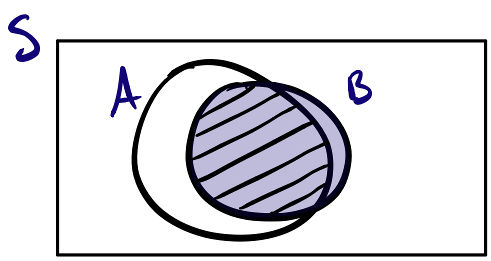
So

$$P\left(\text{You with friend given friend in van}\right) = \frac{4{,}200}{11{,}550} = .3636$$

**Defn:** If $A$ and $B$ are events in $S$ and $P(B) > 0$, then the <u>conditional probability</u> of $A$ given $B$ is

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

* Event $B$ becomes new sample space:




In this picture it looks like if $B$ happens, $A$ is more likely to happen

Eg: We have

$$P(\text{You with friend} \mid \text{friend in Van}) = \frac{P(\text{You with friend} \cap \text{friend in van})}{P(\text{friend in Van})},$$

where from before

$$P(\text{You two in van}) = \frac{4{,}200}{27{,}720} = .1515$$

and

$$P(\text{Friend in Van}) = \frac{\#\{\text{Partitions w/ friend in Van}\}}{\#\{\text{all partitions}\}}$$

$$= \frac{11{,}550}{27{,}720} = .4167$$

So

$$P(\text{You with friend} \mid \text{friend in Van}) = \frac{.1515}{.4167} = .3636$$

(Compare to unconditional probability of riding w/ friend)

Eg: Exp: Wake up on first 2 days of semesters

A: Oversleep on day 1

B = Oversleep on day 2

Suppose
$$p(A) = \frac{1}{4}$$
$$p(B) = \frac{1}{3}$$
$$p(A \cap B) = \frac{1}{10}$$

Then

Oversleep on day 2 given you overslept on day 1

$$p(B \mid A) = \frac{1/10}{1/4} = \frac{4}{10} > p(B)$$

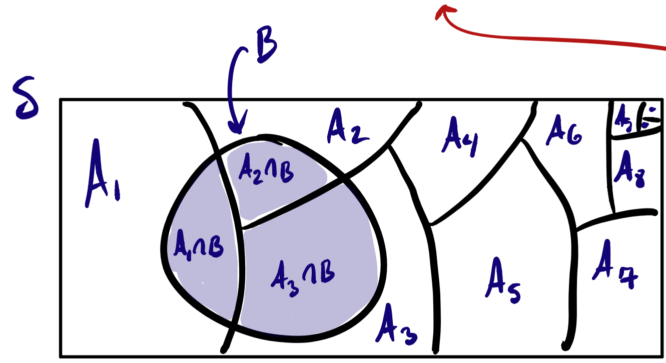
**Theorem (Bayes' Rule):** Let $A_1, A_2, \ldots$ be a partition of the sample space and let $B$ be any set with $P(B) > 0$. Then

$$P(A_i \mid B) = \underbrace{\frac{P(B \mid A_i)\,P(A_i)}{\sum_{j=1}^{\infty} P(B \mid A_j)\,P(A_j)}}$$

← A reformulation of $P(A_i \mid B) = \dfrac{P(A_i \cap B)}{P(B)}$

$\sum_{j=1}^{\infty} P(B \cap A_j)$

**Illustration**




We know by defn of conditional probability

$$P(A_i \mid B) = \frac{P(A_i \cap B)}{P(B)},$$

and $$P(A_i \cap B) = P(B \mid A_i)\,P(A_i)$$

Also

$$P(B) = \sum_{j=1}^{\infty} P(B \cap A_j) = \sum_{j=1}^{\infty} P(B \mid A_i)\,P(A_i)$$

keep on board during next example

**Application (Simpler version) of Bayes' Rule:**

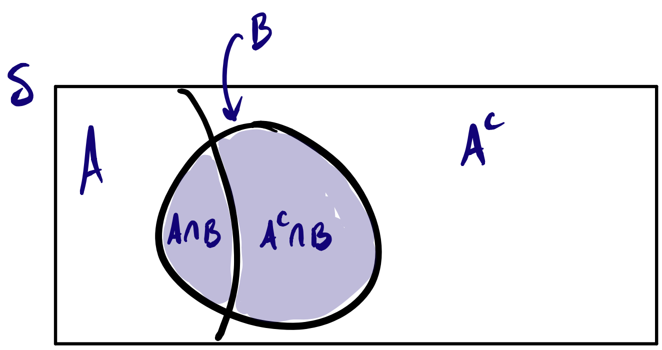
For any events $A$ and $B$ in $S$,

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

← reformulation of $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$

* Instance of Bayes' Rule because $A, A^c$ make a partition of $S$

**Illustration:**




Note that

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

$$P(B) = P(B \cap A) + P(B \cap A^c)$$

$$= P(B|A)P(A) + P(B|A^c)P(A^c)$$

Eg Suppose we have "noisy" Morse Code transmission such that

$$P(\text{'}-\text{' rec} \mid \text{'}\cdot\text{' sent}) = 1/8$$

$$P(\text{'}\cdot\text{' rec} \mid \text{'}-\text{' sent}) = 1/8$$

Want to know

(i) $P(\text{'}\cdot\text{' sent} \mid \text{'}\cdot\text{' rec})$ ← Given that we receive '·', what is the probability that '·' was sent?

(ii) $P(\text{'}-\text{' sent} \mid \text{'}-\text{' rec})$

In Bayesian statistics (which some of you will learn more about later) "prior knowledge" is information possessed by the researcher which can be weighed against observed data when making conclusions.

**Prior knowledge:**

3 '·'s for every 4 '–'s

'·' and '–' sent in proportion 3:4

So 3 of 7 symbols are '·'
4 of 7 symbols are '–'
$\Rightarrow P(\text{'}\cdot\text{' sent}) = 3/7$, $P(\text{'}-\text{' sent}) = 4/7$

(i) By Bayes' Rule

$$P(\underset{\uparrow\,A}{\text{'}\cdot\text{' sent}} \mid \underset{\uparrow\,B}{\text{'}\cdot\text{' rec}}) = \frac{\overbrace{P(\text{'}\cdot\text{' rec} \mid \text{'}\cdot\text{' sent})}^{1-1/8}\overbrace{P(\text{'}\cdot\text{' sent})}^{3/7}}{\underbrace{P(\text{'}\cdot\text{' rec} \mid \text{'}\cdot\text{' sent})}_{1-1/8}\underbrace{P(\text{'}\cdot\text{' sent})}_{3/7} + \underbrace{P(\text{'}\cdot\text{' rec} \mid \text{'}-\text{' sent})}_{1/8}\underbrace{P(\text{'}-\text{' sent})}_{4/7}}$$

$$= \frac{(7/8)(3/7)}{(7/8)(3/7) + (1/8)(4/7)}$$

$$= \frac{21}{25}$$

(ii) $P(\text{'−' sent} \mid \text{'−' rec}) =$ Similar, left as exercise $= \frac{28}{31}$

E.g. Disease screening with imperfect assay

$A$ : positive disease assay , $A^c$ : negative disease assay
$D$ : presence of disease , $D^c$ : no disease

Suppose

$$P(A \mid D) = \underset{\text{"sensitivity" of assay}}{.92} \qquad P(A^c \mid D^c) = \underset{\text{"specificity" of assay}}{.98}$$

How sure can we be of results? Want to find

(i) $P(D \mid A)$

(ii) $P(D^c \mid A^c)$

Aside:
$A \cap D$: "True positive"
$A \cap D^c$: "False positive"
$A^c \cap D^c$: "True negative"
$A^c \cap D$: "False negative"

Prior knowledge: 5% of pop. has disease $\Rightarrow P(D) = .05$

(i) By Bayes' rule

$$P(D \mid A) = \frac{\overbrace{P(A \mid D)\,P(D)}^{P(A \cap D)}}{\underbrace{P(A \mid D)P(D) + P(A \mid D^c)\,P(D^c)}_{= P(A \cap D) + P(A \cap D^c) = P(A)}}$$

$1 - P(A^c \mid D^c) = 1 - .98 = .02$

$1 - P(D) = .95$

$$= \frac{(.92)(.05)}{(.92)(.05) + (.02)(.95)}$$

$$= .7077$$

(ii) $P(D^c | A^c)$ = similar, left as exercise = .9957

**Defn:** Two events $A$ and $B$ in $S$ are called independent if

$$P(A \cap B) = P(A)P(B)$$

* Note that the following statements are equivalent:

(i) $P(A \cap B) = P(A)P(B)$

(ii) $P(A) = P(A|B)$ ← probability of A unaltered by occurence of B

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

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

* If $A, B$ independent, $A, B^c$, $A^c, B$, and $A^c, B^c$ are independent

**E.g.** Exp: Flip a coin twice

$H_1$: head on first flip
$H_2$: head on second flip

$$P(\text{two heads}) = P(H_1 \cap H_2) = P(H_1)P(H_2) = \left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right) = \tfrac{1}{4}$$

**Defn:** A collection of events $A_1, \dots, A_n$ are mutually independent if for any subcollection $A_{i_1}, \dots, A_{i_k}$, we have

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j})$$

**E.g.** Exp: Flip a coin $n$ times,

$H_i$: head on $i^{th}$ flip, $i = 1, \dots, n$

* $H_1, \dots, H_n$ are mutually independent

(i) $P(\text{All heads}) = P\left(\bigcap_{i=1}^{n} H_i\right) = \prod_{i=1}^{n} P(H_i) = \left(\frac{1}{2}\right)^n$

(ii) $P(\text{At least one tail}) = P\left(\{\text{All heads}\}^c\right)$

$$= 1 - P(\text{All heads})$$

$$= 1 - \left(\frac{1}{2}\right)^n$$