Conditional Probability

- Important concept, especially for inference and decision making.
- Often counterintuitive so it is important to rely on the formal definitions, especially initially.
- Motivating questions:
  - → After observing a certain radar signature, how likely is it that an aircraft is approaching?
  - → Test for a disease comes back positive, what is the likelihood the patient has the disease?
  - → Given a set of structural integrity sensor readings, what is the chance that the bridge fails in 5 years?

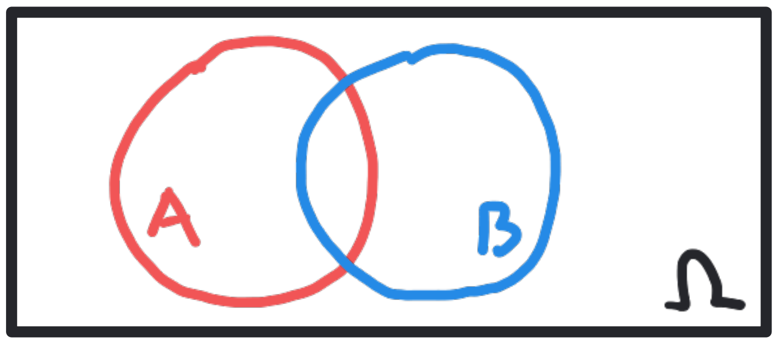
- <u>Simple Example:</u>
  - → Roll a six-sided die.
  - → I tell you the outcome was even.
  - → What is the probability that it was a 4?
  - → Intuitively, we can guess that the probability of seeing a 4 given that the outcome is even is $\frac{1}{3}$.
  - → This is correct.
  - → Unpacking our reasoning a bit: There are three possible even outcomes and all outcomes are equally likely so $\mathbb{P}[\{4\} \mid \{\text{roll even}\}] = \frac{1}{3}$.
  - → What if outcomes are not equally likely?

- The conditional probability $\mathbb{P}[A \mid B]$ of event A given that event B occurs is

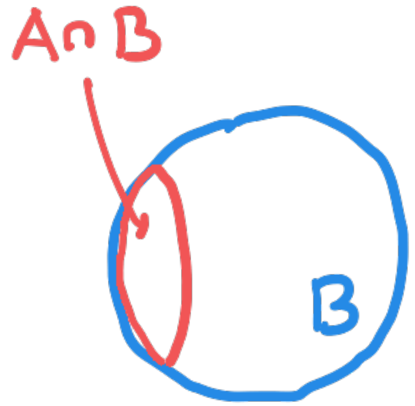
$$\boxed{\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}}$$

→ Recall that B can only occur if $\mathbb{P}[B] > 0$. For $\mathbb{P}[B] = 0$, $\mathbb{P}[A \mid B]$ is undefined.

→ Intuition: Conditioning on B is like restricting the universe of outcomes to those in B. Thus, given B, only outcomes in $A \cap B$ are possible out of those in A.



$\xrightarrow{\text{Condition on } B}$



Only outcomes in B are now possible.

- Let's see how this formula works in the simple example:

$$\mathbb{P}[\{4\} \mid \{\text{roll even}\}]$$
$$= \mathbb{P}[\{4\} \mid \{2, 4, 6\}]$$
$$= \frac{\mathbb{P}[\{4\} \cap \{2, 4, 6\}]}{\mathbb{P}[\{2,4,6\}]}$$
$$= \frac{\mathbb{P}[\{4\}]}{\mathbb{P}[\{2\}] + \mathbb{P}[\{4\}] + \mathbb{P}[\{6\}]}$$
$$= \frac{\frac{1}{6}}{\frac{1}{6} + \frac{1}{6} + \frac{1}{6}}$$
$$= \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

- This is an instance of a special case for conditional probability: If all outcomes are equally likely, then

$$\mathbb{P}[A \mid B] = \frac{\#\text{ outcomes in } A \cap B}{\#\text{ outcomes in } B}$$

(assuming finite number of outcomes)

- all of the axioms and properties for probability carry over to conditional probability.

axioms

Non-negativity: $\mathbb{P}[A \mid B] \geq 0$

Normalization: $\mathbb{P}[\Omega \mid B] = \mathbb{P}[B \mid B] = 1$

Countable additivity: For any countable collection $A_1, A_2, \ldots$ of mutually exclusive events,

$$\mathbb{P}\left[\bigcup_{i=1}^{\infty} A_i \,\middle|\, B\right] = \sum_{i=1}^{\infty} \mathbb{P}[A_i \mid B]$$

Complement: $\mathbb{P}[A^c \mid B] = 1 - \mathbb{P}[A \mid B]$

Inclusion - Exclusion: $\mathbb{P}[A \cup B \mid C] = \mathbb{P}[A \mid C] + \mathbb{P}[B \mid C] - \mathbb{P}[A \cap B \mid C]$

- **Multiplication Rule:** For any events $A_1, A_2, \dots, A_n$,

$$\boxed{\mathbb{P}\left[\bigcap_{i=1}^{n} A_i\right] = \mathbb{P}[A_1] \cdot \mathbb{P}[A_2 \mid A_1] \cdot \dots \cdot \mathbb{P}\left[A_n \mid \bigcap_{i=1}^{n-1} A_i\right]}$$

→ This assumes $\mathbb{P}\left[\bigcap_{i=1}^{n-1} A_i\right] > 0$ so the conditional probabilities are defined.

→ Special case of Two Events A and B: $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B \mid A]$

(can condition in any order) or $\mathbb{P}[A \cap B] = \mathbb{P}[B]\,\mathbb{P}[A \mid B]$

- Example: Draw three consecutive playing cards. Probability all hearts?

$A_i = \{i^{th} \text{ card is a heart}\}$

$$\mathbb{P}[A_1] = \frac{13}{52} \qquad \mathbb{P}[A_2 \mid A_1] = \frac{12}{51} \qquad \mathbb{P}[A_3 \mid A_1 \cap A_2] = \frac{11}{50}$$

$$\mathbb{P}[A_1 \cap A_2 \cap A_3] = \mathbb{P}[A_1] \cdot \mathbb{P}[A_2 \mid A_1] \cdot \mathbb{P}[A_3 \mid A_1 \cap A_2]$$

$$= \frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50} = \frac{11}{850}$$

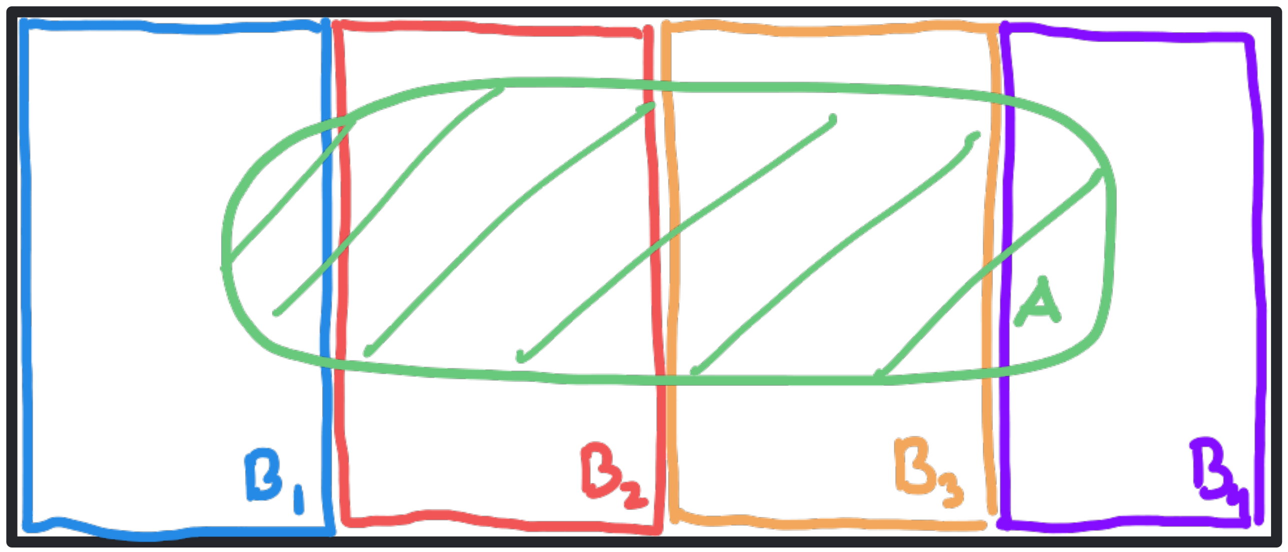
- **Law of Total Probability:** For a partition $B_1, B_2, \dots$ satisfying $\mathbb{P}[B_i] > 0$ for all i,

$$\boxed{\mathbb{P}[A] = \sum_{i=1}^{\infty} \mathbb{P}[A \mid B_i]\,\mathbb{P}[B_i]}$$

→ Why does this work?

$$\mathbb{P}[A \mid B_i]\,\mathbb{P}[B_i] = \mathbb{P}[A \cap B_i] \qquad \text{Multiplication Rule}$$

$$\Rightarrow \sum_{i=1}^{\infty} \mathbb{P}[A \mid B_i]\,\mathbb{P}[B_i] = \sum_{i=1}^{\infty} \mathbb{P}[A \cap B_i] \underset{\substack{\uparrow \\ \text{Additivity}}}{=} \mathbb{P}\left[\bigcup_{i=1}^{\infty} A \cap B_i\right] = \mathbb{P}[A]$$



- **Bayes' Rule:** Think of this as a technique to "flip" conditioning.

$$\boxed{\mathbb{P}[B \mid A] = \frac{\mathbb{P}[A \mid B]\,\mathbb{P}[B]}{\mathbb{P}[A]}}$$

→ Why? $\mathbb{P}[A \mid B]\,\mathbb{P}[B] = \mathbb{P}[A \cap B]$ and $\mathbb{P}[B \mid A] = \dfrac{\mathbb{P}[A \cap B]}{\mathbb{P}[A]}$.

→ Very important to remember that $\mathbb{P}[A \mid B] \neq \mathbb{P}[B \mid A]$ in general!

→ Can combine with the Law of Total Probability.
Let $B_1, B_2, \dots$ be a partition with $\mathbb{P}[B_i] > 0$ for all i.

$$\mathbb{P}[B_j \mid A] = \frac{\mathbb{P}[A \mid B_j]\,\mathbb{P}[B_j]}{\mathbb{P}[A]} = \frac{\mathbb{P}[A \mid B_j]\,\mathbb{P}[B_j]}{\sum_{i=1}^{\infty} \mathbb{P}[A \mid B_i]\,\mathbb{P}[B_i]}$$