Lecture 3: Bayes Theorem and Inference

Lecturer: Jie Fu, Ph.D.

Outline

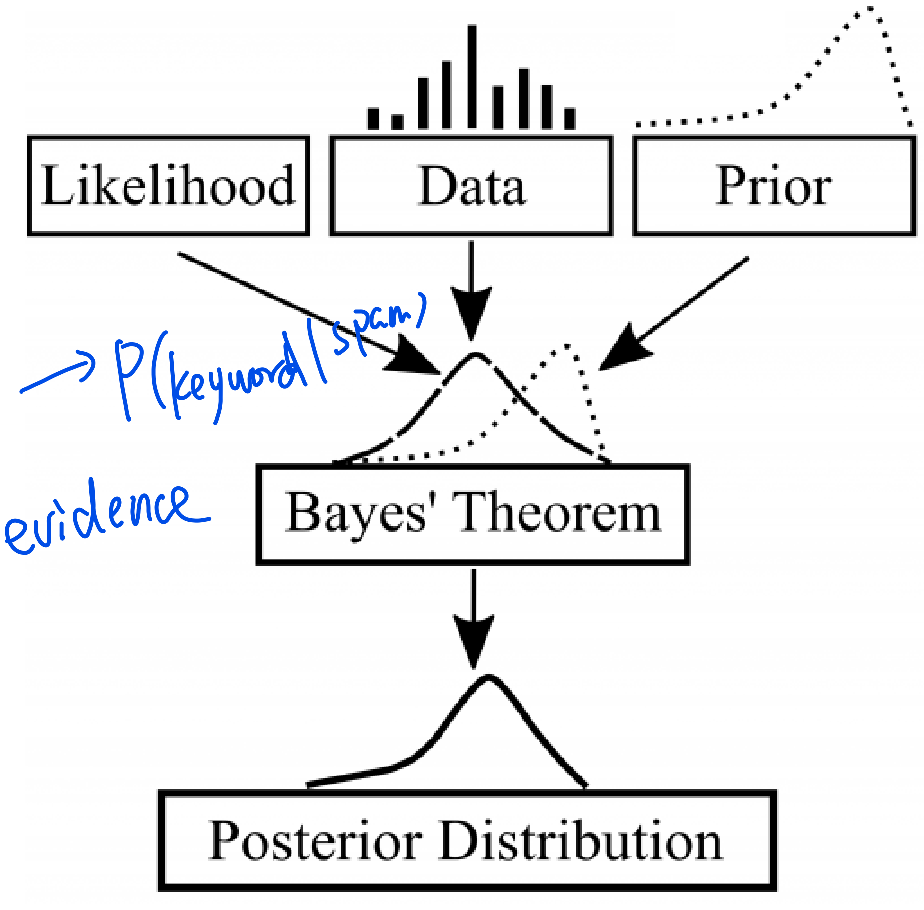
- Bayes Theorem
- Statistical Independence

Related applications

- **Disease Diagnosis**: • Detecting cancer based on medical imaging results. • Estimating the probability of infection based on symptoms and diagnostic test outcomes.
- **Spam Filtering**: Bayesian spam filters calculate the probability of an email being spam based on keywords and patterns. $P(\text{spam} \mid \text{keywords})$
- **Bayesian Neural Networks**: Incorporating uncertainty in predictions by treating weights as distributions instead of fixed values.
- **Sensor Fusion**: Combining data from multiple sensors (e.g., GPS and IMU) to improve state estimation.

evidence

$P(\text{keyword} \mid \text{spam})$



A motivating problem

- A magician has two coins, one fair and one 2-headed coin. Consider the experiment where she picks one coin at random and flips it once, the output is a head, what is the probability that the coin is a fair coin?

- Initial belief $P(A_i)$ for possible cause of an observed event B.

  e.g. Both coins are equally likely —- a prior knowledge:

  = A_1: F. A_2: 2H

- probability of the observation under each A_i:

  e.g. Probability of a head under each coin.

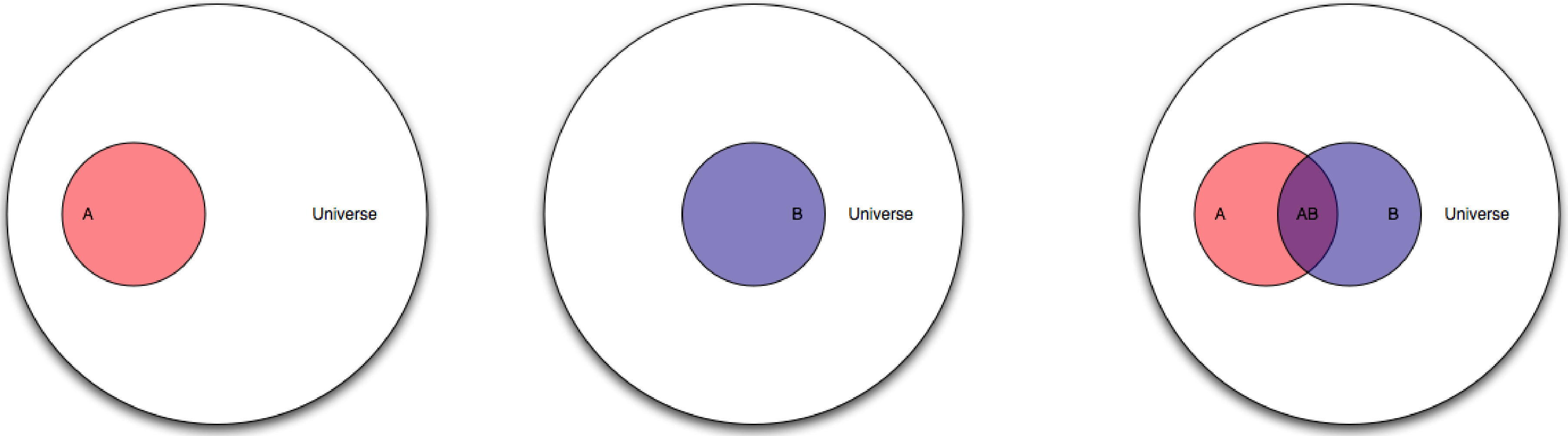
$$A_i \xrightarrow[\mathbf{P}(B \mid A_i)]{\text{model}} B \rightarrow \text{observed event "Heads"}$$

- Draw conclusion about the cause given the observed event.

  e.g. infer if the coin is fair or biased.

$$B \xrightarrow[\mathbf{P}(A_i \mid B)]{\text{inference}} A_i$$

Venn diagram visualization



$$A_i \xrightarrow[\mathbf{P}(B \mid A_i)]{\text{model}} B$$

Bayes Theorem

Consider two events A and B, by the chain rule:

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

and

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

Note that

observed

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

$$\underbrace{P(A|B)}_{\text{infer model form observation:}}\ \underbrace{P(B)}_{\text{event B occurs.}} = \underbrace{P(B|A)}_{\text{prediction based on model: A}}\ \underbrace{P(A)}_{\text{prior prob of selecting the model A.}}$$

Bayes Theorem

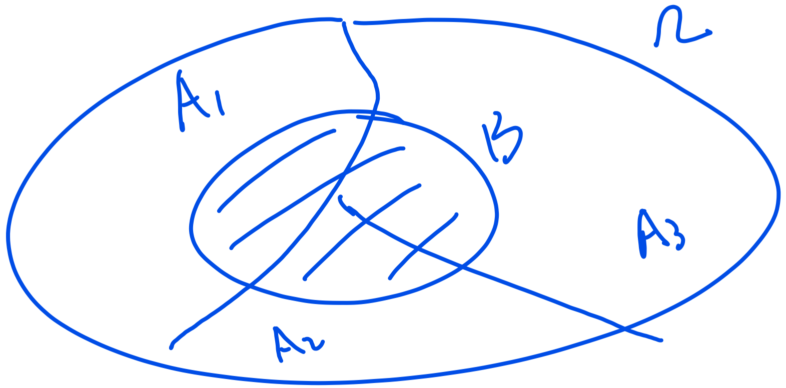
If the set of events $\{A_i\}_{i=1}^{n}$ partitions the sample space Ω, and assuming $P(A_i) > 0$, for all i. Then, for any event B such that $P(B) > 0$, we have

infer

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

where $P(B)$ can be computed using the Law of Total Probability,

$$P(B) = P(B|A_1)P(A_1) + \cdots + P(B|A_n)P(A_n)$$



$$P(A_i|B) \propto P(B|A_i)P(A_2)$$

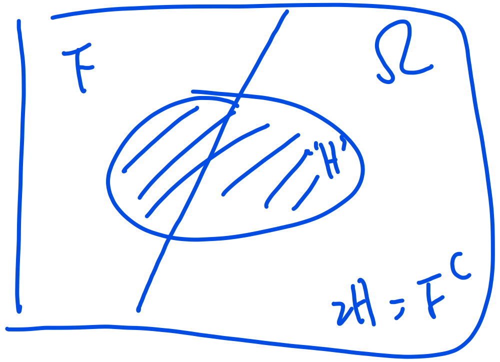
"proportional to"

Example

$P(F) = \frac{1}{2}$ $\xrightarrow{\text{observed } H}$ $P(F|H) = \frac{1}{3}$

initial belief current belief.

- A magician has two coins, one fair and one 2-headed coin. Consider the experiment where she picks one coin at random and flips it once, the output is a head, what is the probability that the coin is a fair coin?



$\{A_1 \dots A_n\}$ partition

observed event B :

$$P(F|H) = \frac{P(H|F)\,P(F)}{P(H)} = \frac{\frac{1}{2}\times\frac{1}{2}}{\frac{3}{4}} = \frac{1}{3}$$

A_1

inference

$B \longrightarrow A_i$

$P(A_i|B)$

$$P(H) = P(H|F)\,P(F) + P(H|2H)\,P(2H)$$

$$= \frac{3}{4}$$

? if observed another head: $P(F\,|\,\text{2nd head}) = \dfrac{P(H|F)\cdot\hat{P}(F)}{\hat{P}(H)} = \dfrac{\frac{1}{2}\times\hat{p}(F)}{\hat{P}(H)}$

Example

$$P(H_2|H_1) = P(H_2|F)\,P_1(F) + P(H_2|2H)\,P_1(2H)$$

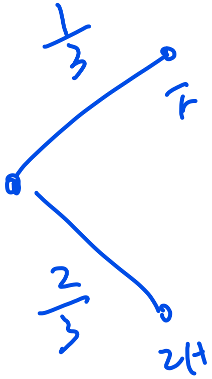
belief after 1 heads.

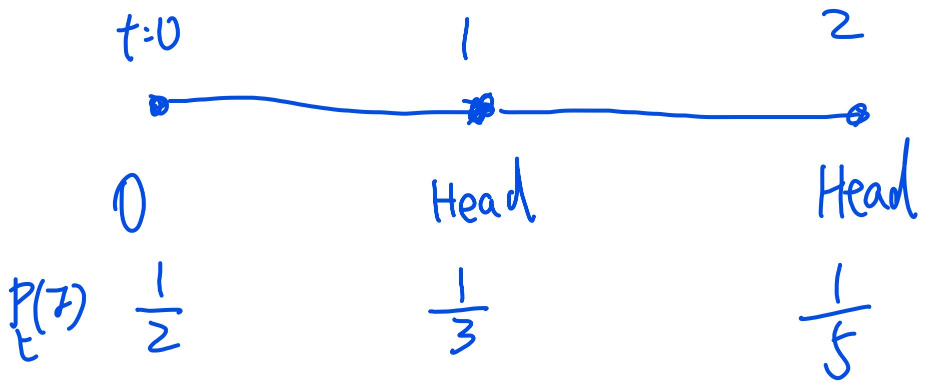
$$= \frac{1}{2}\times\frac{1}{3} + 1\times\frac{2}{3} = \frac{5}{6}.$$

UF UNIVERSITY of FLORIDA

- A magician has two coins, one fair and one 2-headed coin. Consider the experiment where she picks one coin at random and flips it once, the output is a head, what is the probability that the next second flip is a head?

$$P(F \mid 2\text{nd Head}) = \frac{\frac{1}{2}\times\frac{1}{3}}{\hat{P}(H)} = \frac{\frac{1}{2}\times\frac{1}{3}}{\frac{1}{3}\times\frac{1}{2}+\frac{2}{3}\times 1} = \frac{\frac{1}{6}}{\frac{1}{6}+\frac{4}{6}} = \frac{1}{5}$$

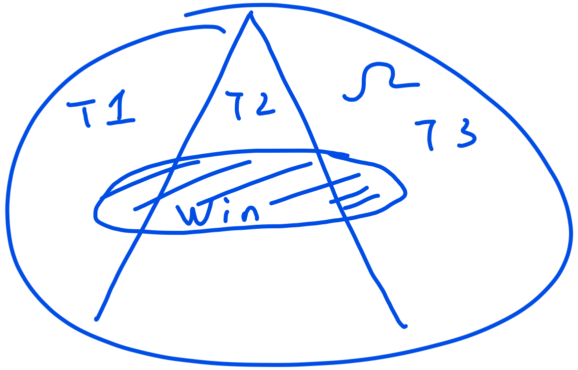




Some terminologies

- $P(A_i|B)$ as the **posterior probability** of event A_i given the information
- $P(A_i)$ as the **prior probability**
- $P(B|A_i)$ as the **likelihood**
- $P(B)$ as the **evidence/effect probability**

Example



- Three types of players.
  - Type 1: 50%
  - Type 2: 25%
  - Type 3: 25%
- You winning probability with these players:
  - Against type 1: 0.3.
  - Against type 2: 0.4.
  - Against type 3: 0.5.
- Now you play a game with a randomly chosen player.
- *Question*: What's your winning probability?

$$P(W) = P(W|T_1)P(T_1) + P(W|T_2)P(T_2) + P(W|T_3)P(T_3)$$

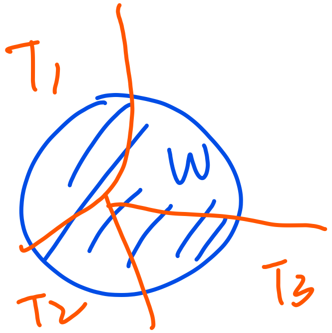
$$= 0.3 \times 0.5 + 0.4 \times 0.25 + 0.5 \times 0.25.$$

observed event. $W \xrightarrow{\text{infer}}$ Type 1.

prior: Types. Win. against type. $\xrightarrow[\text{win}]{\text{infer}}$? $P(T_1|W)$

$$P(T_1|W) = \frac{P(W|T_1)\cdot P(T_1)}{P(W)} = \frac{0.3\times 0.5}{P(W)}$$

↑ posterior



?

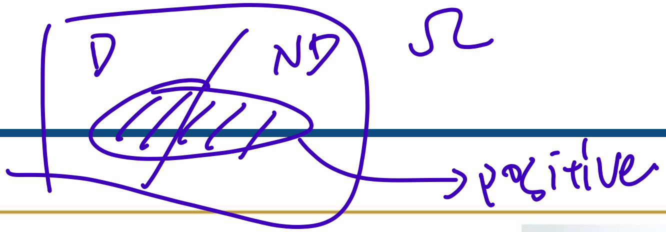
$$P(T_1|W) + P(T_2|W) + P(T_3|W) = \begin{cases} 1 & \checkmark \\ P(W) & \end{cases}$$

$$= \frac{P(W|T_1)P(T_1) + P(W|T_2)P(T_2) + P(W|T_3)P(T_3)}{P(W)} = 1$$

Example

- Suppose that you win. What is the probability that you had an opponent of type 1?

Example



Example: Diagnosis

- A random person drawn from a certain population has probability 0.001 of having a certain disease.

  0.1

- The test satisfies
  - Pr[test positive | disease] = 0.95
  - Pr[test negative | no disease] = 0.95

false negative = P(test negative | disease) = 0.05

false positive = P(test positive | no disease) = 0.05

- *Question*: Given that the person just tested positive, what is the probability of having the disease?

observed event

$$P(\underbrace{\text{disease}}_{\text{infer}} \mid \underbrace{\text{test positive}}_{\text{observed}}) = \frac{P(\text{test positive} \mid \text{disease})\, P(\text{disease})}{P(\text{test positive})}$$

1

$$P(\text{test positive}) = P(\text{positive} \mid \text{disease})\,P(\text{disease}) + P(\text{positive} \mid \text{No disease}) \cdot P(\text{no disease})$$

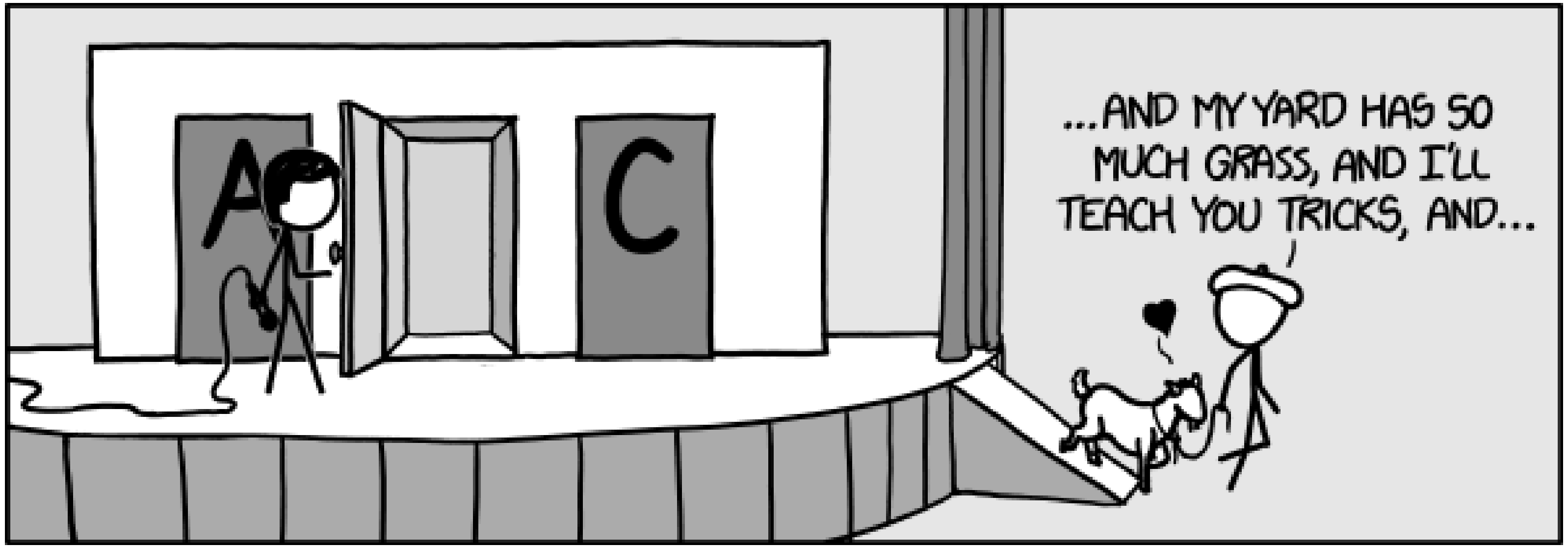
$$= 0.95 \times 0.001 + 0.05 \times 0.999$$

$$① = \frac{0.95 \times 0.001}{0.95 \times 0.001 + 0.05 \times 0.999} \doteq 0.01866\ldots$$

Monty Hall Problem

- Suppose you're on a game show, and you're given the choice of three doors:
- behind one door is a car
- behind the other doors are goats

You pick a door, and the host, who knows what's behind the doors, opens another door, which he knows has a goat. The host then offers you the option to switch doors. Does it matter if you switch?



Monty hall problem

- Let W_i be the event of winning a car on the i-th choice.
- Consider three strategies:
  - → Never switch.
  - Always switch.
  - Flip a coin, if heads, switch, if tail, no switch.

$t=0$ choose one of A, B, C

$t=1$ make a choice

① Not switch:

$$P(W_2) = P(W_2 \mid W_1)\,P(W_1) + P(W_2 \mid W_1^c)\,P(W_1^c)$$

$$= 1 \times \frac{1}{3} + 0 \times \frac{2}{3} = \frac{1}{3}.$$

② Always switch:

$$P(W_2) = P(W_2 \mid W_1)\,P(W_1) + P(W_2 \mid W_1^c)\,P(W_1^c)$$

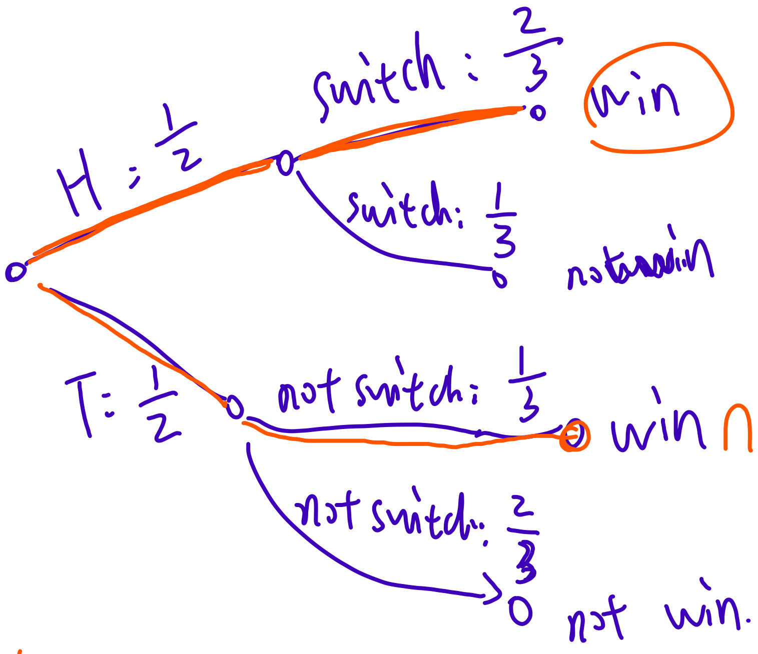
$$= 0 \times \frac{1}{3} + 1 \times \frac{2}{3} = \frac{2}{3}$$

A 1*	B	C

opened.

loaded: p $(1-p)$

③ Flip a loaded coin:



$P(\text{win} \cap \text{Heads}) = \frac{1}{2} \times \frac{2}{3}$

win ∩ Tails: $P(\text{win} \cap \text{Tails}) = \frac{1}{2} \times \frac{1}{3}$

$$P(\text{win}) = \frac{1}{2} \times \frac{2}{3} + \frac{1}{2} \times \frac{1}{3} < \frac{2}{3}$$

$$p \times \frac{2}{3} + (1-p) \times \frac{1}{3} < \frac{2}{3} \quad \text{for any } p \in (0,1)$$

Independence

In general, for two events A and B, when $P(A|B) = P(A)$, we say that A is statistically independent (s.i.) of B, since the probabilities are not affected by knowledge of B having occurred.

* By the chain rule, if A is independent of B:

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

Events A and B are **statistically independent (s.i.)** if and only if (iff)

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

Independence

In general, for two events A and B, when $P(A|B) = P(A)$, we say that A is statistically independent (s.i.) of B, since the probabilities are not affected by knowledge of B having occurred.

* By the chain rule, if A is independent of B:

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

Events A and B are **statistically independent (s.i.)** if and only if (iff)

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

Independence

- If A is independent of B, then B is also independent of A.
- Why? $P(B \cap A) = P(B|A)\,P(A) = P(B)\,P(A)$

independence.

If A and B are s.i. events, then the following pairs of events are also s.i.:
* A and $\overline{B}$
* $\overline{A}$ and B
* $\overline{A}$ and $\overline{B}$

$P(A \cap \overline{B}) = P(A)P(\overline{B})$ given $P(A \cap B) = P(A)P(B)$

Conditional independence

given an event C, the events A and B are called *conditionally independent* if

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

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

Consider two independent coin tosses. Let

? $P(H_1 \cap H_2) = P(H_1) \cdot P(H_2)$

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

$$H_1 := \{\text{ 1st toss is a head}\}$$

$$H_2 := \{\text{ 2nd toss is a head}\}$$

$$D := \{\text{ two tosses have different results}\}$$

Compare $\underset{0}{\underset{\parallel}{P(H_1 \cap H_2|D)}}$ and $\underbrace{P(H_1 \cap H_2)}_{}$ $= \frac{1}{4}$ indep.

Example

$P(H_1 \cap H_2 | D) = 0$

$\parallel$ ✗

? $P(H_1 | D) \cdot P(H_2 | D) > 0$

>0 >0

$\{H, T\}$ $\{T, H\}$

not independent conditioned on D