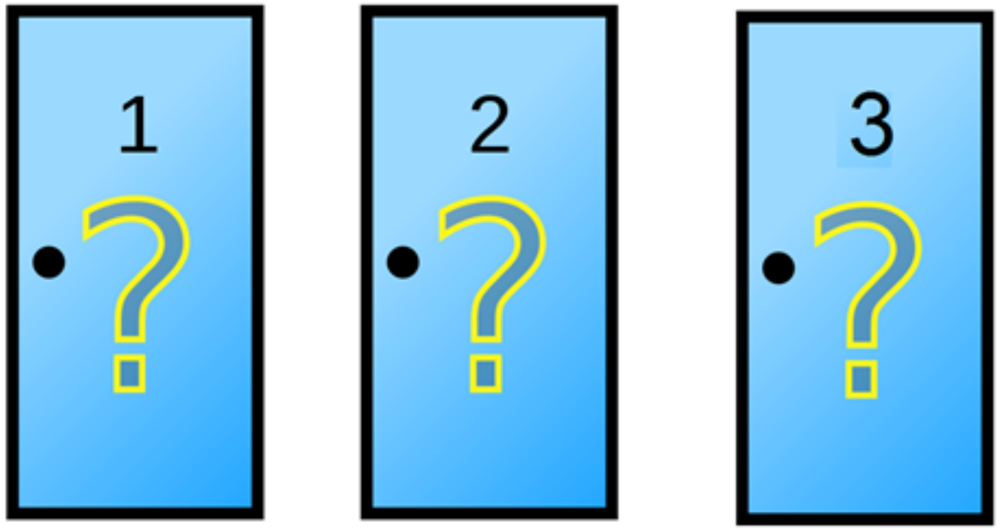
1179: Probability
Lecture 4 — Conditional Probability

Ping-Chun Hsieh (謝秉均)

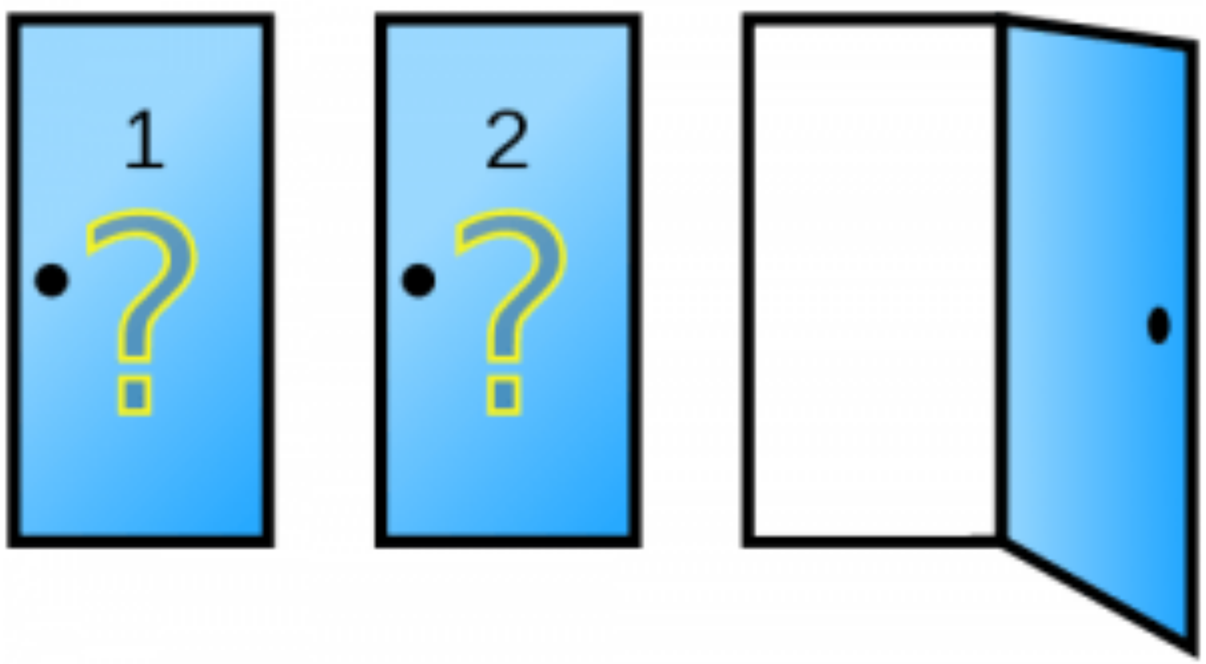
September 24, 2021

Monty Hall Problem



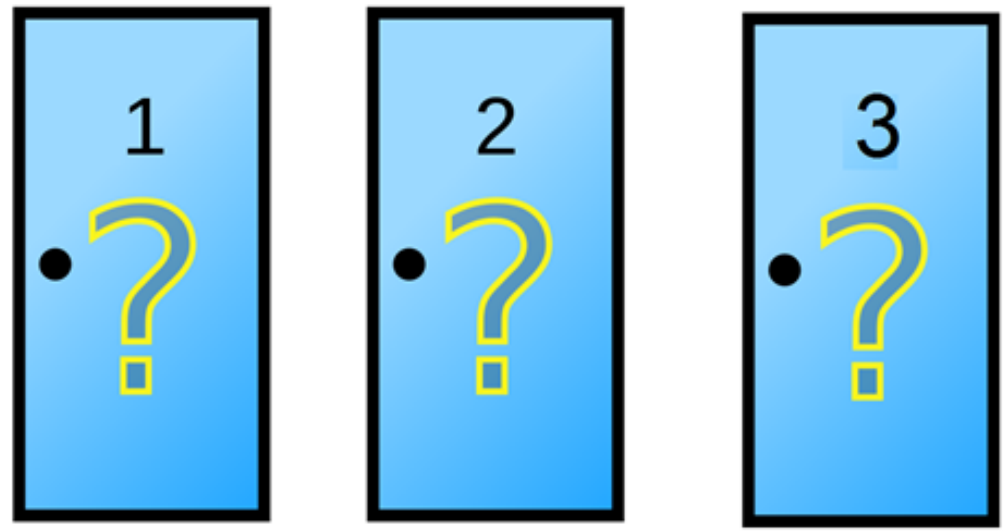
- Suppose Bill is given 3 options
- 2 of them are empty, and the remaining one has the prize

- Then, Bill picks a door (say Door #1).
- Next, the moderator opens an empty door.



- Bill is then asked by the moderator: “Do you want to switch or stay?”

Monty Hall Problem (Cont.)



- What is random (from Bill's perspective)?
- Sample space? $\Omega = \{1, 2, 3\}$
- What's the probability of winning the prize if Bill stays? $\frac{1}{3}$
- What's the probability of winning the prize if Bill switches? $\frac{2}{3}$

Quick Review

if E_n is either an increasing or a decreasing event

- What is "continuity of probability"? $\lim_{n\to\infty} P(E_n) = P(\lim_{n\to\infty} E_n)$
- What is conditional probability? $P(A|B) = \frac{P(A\cap B)}{P(B)}$

called P A given B

This Lecture

1. Conditional Probability and 3 Useful Tools

- Reading material: Chapter 3.1~3.4

Conditional Probability Defines a New Probability Assignment

Theorem (Reduction of Sample Space):

Let Ω be the sample space and let B be an event with $P(B) > 0$. Then, we have:

by axiom 1

1. $P(A \mid B) \geq 0$, for any event A $\quad P(A|B) = \dfrac{P(A\cap B)}{P(B) > 0} \geq 0$ #

2. $P(\Omega \mid B) = 1 \quad P(\Omega|B) = \dfrac{P(\Omega\cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$ #

3. $A_1, A_2, \cdots$ is an infinite sequence of <u>mutually exclusive</u> events, then

$$P(\bigcup_{i=1}^{\infty} A_i \mid B) = \sum_{i=1}^{\infty} P(A_i \mid B)$$

By axiom 3

$$P(\bigcup_{i=1}^{\infty} A_i | B) = \frac{P((\bigcup_{i=1}^{\infty} A_i)\cap B)}{P(B)} = \frac{P(\bigcup_{i=1}^{\infty} (A_i\cap B))}{P(B)} = \frac{\sum_{i=1}^{\infty} P(A_i\cap B)}{P(B)}$$

$$= \sum_{i=1}^{\infty} P(A_i|B) \ \#$$

- Conditional universe!

Tool #1: Multiplication Rule

Assuming that all of the conditioning events have positive probability, we have:

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2 \mid A_1)\cdots P(A_n \mid A_1 \cap A_2 \cap \cdots A_{n-1})$$

- How to intuitively interpret this?
- How to prove this? mathematical induction

1. $n=2$: $P(A_1 \cap A_2) = P(A_1) \cdot P(A_2 \mid A_1)$ (by the definition of conditional property)

2. Assume that $n=k$, $P(\bigcap_{i=1}^{k} A_i) = P(A_1) \cdot P(A_2 \mid A_1) \cdot \cdots P(A_k \mid A_1 \cap A_2 \cap \cdots A_{k-1})$ is true

then if $n=k+1$. $P(\bigcap_{i=1}^{k+1} A_i) = P((\bigcap_{i=1}^{k} A_i) \cap A_{k+1}) = P(\bigcap_{i=1}^{k} A_i) \cap P(A_{k+1} \mid A_1 \cap A_2 \cap \cdots A_k)$

by mathetical nduction #

Example: Find the Defective Fuses

- Example: Suppose that 7 good and 2 defective fuses are mixed up. To find the defective ones, we test them one by one. P(we find both defective fuses in exactly 3 tests) = ?

$\begin{cases} GDD \\ DGD \end{cases}$

$P(GDD) = P(\text{1st is } G) \cdot P(\text{2nd is } D \mid \text{1st is } G) \cdot P(\text{3rd is } D \mid GD)$

$= \frac{7}{9} \times \frac{2}{8} \times \frac{1}{7} = \frac{1}{36}$

$P(DGD) = P(\text{1st is } D) \cdot P(\text{2nd is } G \mid \text{1st is } D) \cdot P(\text{3rd is } D \mid DG)$

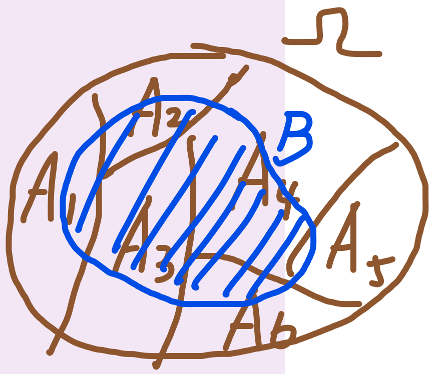
$= \frac{2}{9} \cdot \frac{7}{8} \cdot \frac{1}{7} = \frac{1}{36}$

$\frac{1}{36} + \frac{1}{36} = \frac{1}{18}$ #

Tool #2: Total Probability Theorem

Theorem: Let $A_1, A_2, \cdots, A_n$ be mutually exclusive events that form a partition of Ω, and $P(A_i) > 0$, for all i. Then, for any event B, we have

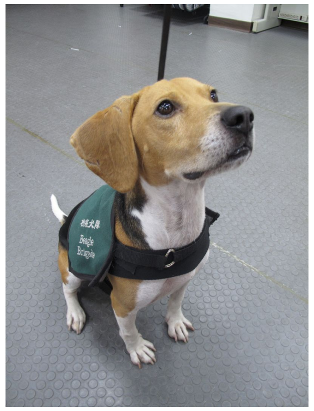
$$
\begin{aligned}
P(B) &= P(A_1 \cap B) + \cdots + P(A_n \cap B) \\
&= P(A_1)P(B \mid A_1) + \cdots + P(A_n)P(B \mid A_n)
\end{aligned}
$$

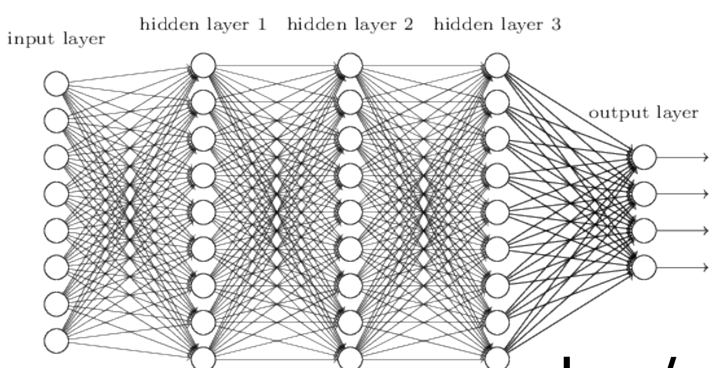


- Idea: Divide and conquer

Example: Image Classifier

- Example: Suppose we have a well-trained image classifier
  - Each input image is either a dog/cat/bird with $P(\text{dog}) = P(\text{cat}) = 2 \cdot P(\text{bird})$ $\longrightarrow$ $\begin{cases} P(\text{dog}) = \frac{2}{5} \\ P(\text{cat}) = \frac{2}{5} \\ P(\text{bird}) = \frac{1}{5} \end{cases}$
  - The probability that a dog is misclassified is 0.1
  - The probability that a cat is misclassified is 0.05
  - The probability that a bird is misclassified is 0.15
  - P(an image is correctly classified) = ?





dog/cat/bird

$$\frac{2}{5} \times 0.9 + \frac{2}{5} \times 0.95 + \frac{1}{5} \times 0.85$$

$$= 0.36 + 0.38 + 0.17$$

$$= 0.91$$ #

Example: Gambler's Ruin

- Example: Two gamblers A and B keep tossing a fair coin
  - If ''head'' occurs, A pays $1 to B; otherwise, B pays $1 to A
  - Initially, A has 2 dollars, and B has 3 dollars
  - The game ends when either A or B has zero dollar
  - What is the probability that A wins the game?

Tool #3: Bayes' Rule

Theorem: Let $A_1, A_2, \cdots, A_n$ be mutually exclusive events that form a partition of Ω, and $P(A_i) > 0$, for all i. Then, for any event B, we have

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(B)} = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + \cdots + P(A_n)P(B \mid A_n)}$$

- Why is Bayes' rule useful? → Inference

Bayesian Inference: Crush and Dates

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(B)} = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + \cdots + P(A_n)P(B \mid A_n)}$$

- Example: Bill has a crush on Amy, and Bill wants to ask Amy out to see whether Amy likes him or not.
  - $A_1 = \{$Amy likes Bill$\}$, $A_2 = \{$Amy does not like Bill$\}$
  - $B = \{$Amy looks happy during the date$\}$
  - $P(B \mid A_1) = 0.9$, and $P(B \mid A_2) = 0.3$
  - What are $P(A_1 \mid B)$ and $P(A_1 \mid B^c)$?

Example: Answer an Exam Question

- Example: Bill answers a question with 4 choices (A, B, C, D)
  - Bill either knows the correct answer or makes a random guess
  - $P(\text{Bill knows the correct answer}) = 2/3$
  - $P(\text{Bill does not make a random guess} \mid \text{answer is correct}) = ?$

1-Minute Summary

1. Conditional Probability

- Multiplication rule / Total probability theorem / Bayes' rule
- Bayesian inference