Conditional Probability and Independence

Saravanan Vijayakumaran

Department of Electrical Engineering
Indian Institute of Technology Bombay

January 15, 2025

Conditional Probability

Conditional Probability

Definition

If $P(B) > 0$ then the conditional probability that A occurs given that B occurs is defined to be

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

Examples

- Two fair dice are thrown. Given that the first shows 3, what is the probability that the total exceeds 6?
- A box has three white balls w_1, w_2, and w_3 and two red balls r_1 and r_2. Two random balls are removed in succession. What is the probability that the first removed ball is white and the second is red?

Law of Total Probability

Theorem

For any events A and B such that $0 < P(B) < 1$*,*

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

More generally, let $B_1, B_2, \ldots, B_n$ *be a partition of* Ω *such that* $P(B_i) > 0$ *for all* i*. Then*

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

Examples

- Box 1 contains 3 white and 2 black balls. Box 2 contains 4 white and 6 black balls. If a box is selected at random and a ball is chosen at random from it, what is the probability that it is white?
- We have two coins; the first is fair and the second has heads on both sides. A coin is picked at random and tossed twice. What is the probability of heads showing up in both tosses?

Bayes' Theorem

Theorem

For any events A and B such that $P(A) > 0, P(B) > 0$*,*

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

If $A_1, \ldots, A_n$ *is a partition of* Ω *such that* $P(A_i) > 0$ *and* $P(B) > 0$*, then*

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}.$$

Examples

- Box 1 contains 3 white and 2 black balls. Box 2 contains 4 white and 6 black balls. A box is selected at random and a ball is chosen at random from it. If the chosen ball is white, what is the probability that box 1 was selected?
- We have two coins; the first is fair and the second has heads on both sides. A coin is picked at random and tossed twice. If heads showed up in both tosses, what is the probability that the coin is fair?

Independence

Independent Events

Definition

Events A and B are called independent if

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

More generally, a family $\{A_i : i \in I\}$ is called independent if

$$P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i)$$

for all finite subsets J of I.

Examples

- A fair coin is tossed twice. The first toss being Heads is independent of the second toss being Heads.
- A card is picked at random from a pack of 52 cards. The suit of the card being Spades is independent of its value being 5.
- Two fair dice are rolled. Is the the sum of the faces independent of the number shown by the first die?

Questions

- What is the relation between independence and conditional probability?
- Does pairwise independence imply independence?

  $\Omega = \{abc, acb, cab, cba, bca, bac, aaa, bbb, ccc\}$ with each outcome being equally likely.

  Let A_k be the event that the kth letter is a.

$$\begin{aligned} P(A_i) &= \frac{1}{3} \\ P(A_i \cap A_j) &= \frac{1}{9}, \quad i \neq j \\ P(A_1 \cap A_2 \cap A_3) &= \frac{1}{9} \end{aligned}$$

  $\{A_1, A_2, A_3\}$ are pairwise independent but not independent.

Conditional Independence

Definition

Let C be an event with $P(C) > 0$. Two events A and B are called conditionally independent given C if

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

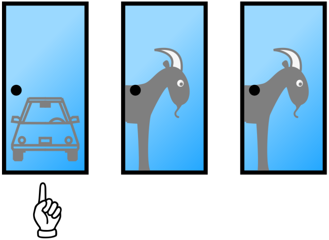
Example

- We have two coins; the first is fair and the second has heads on both sides. A coin is picked at random and tossed twice. Are the results of the two tosses independent? Are they independent if we know which coin was picked?

Monty Hall Problem

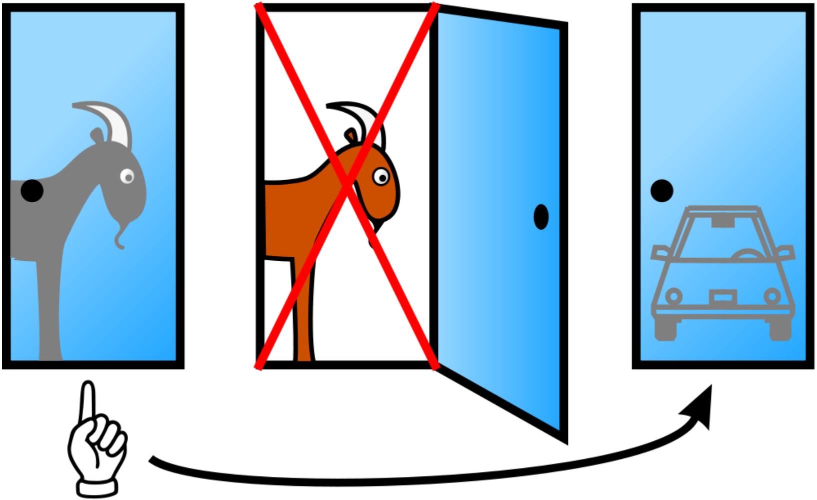
Monty Hall Problem

- Monty Hall was the host of an American game show *Let's Make a Deal*
- When game starts, contestant sees three closed doors

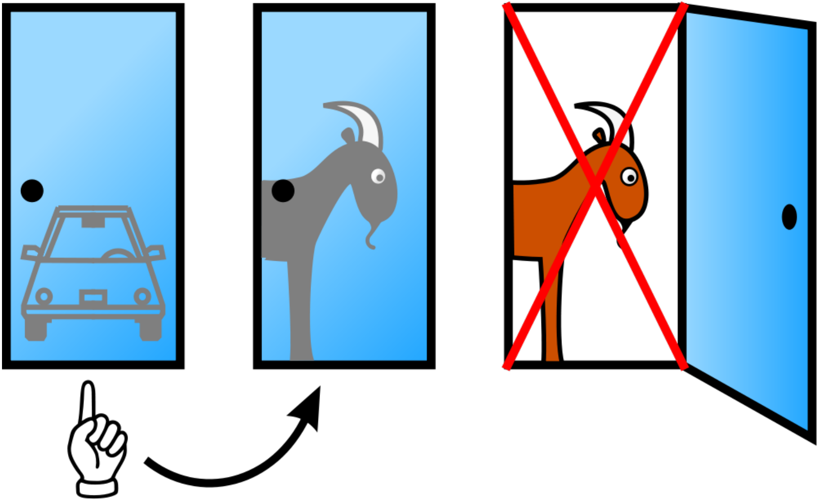


- One of the doors has a car behind it and the other two have goats
- The goal of the game is to pick the door which has the car behind it
- Rules of the game
  - Initially, contestant picks one of the doors, say door A
  - Monty Hall opens one of the other doors (B or C) which has a goat
  - The contestant is now given an option to change his choice
  - Should he switch from his current choice to the unopened door?

Switching May Win

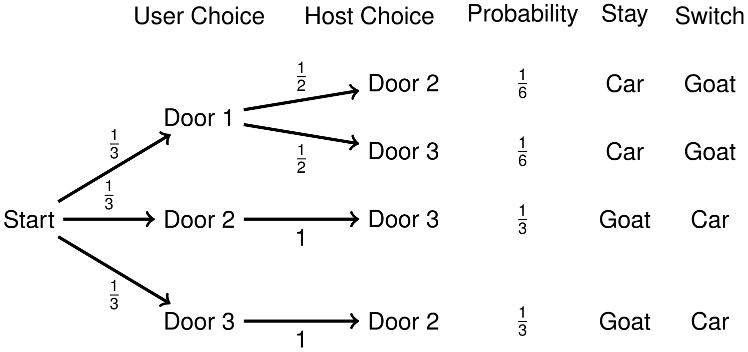


Switching May Lose



To switch or stay

- We will choose the strategy which has a higher probability of winning
- Suppose the car is behind Door 1
- What is the sample space?



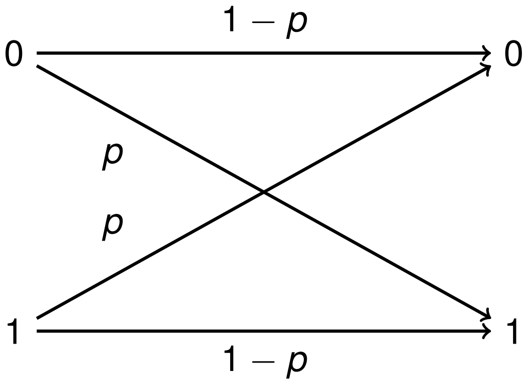
Probability of winning with staying = $\frac{1}{3}$

Probability of winning with switching = $\frac{2}{3}$

Repetition Code over a Binary Symmetric Channel

Binary Symmetric Channel

- Channel with binary input and output

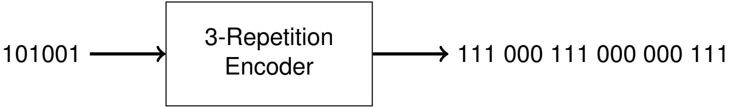


- The parameter p is called the crossover probability
- p is assumed to be less than $\frac{1}{2}$
- Errors introduced on different input bits are independent

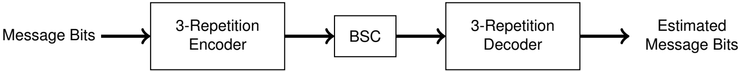
The 3-Repetition Code

- Given a block of message bits, each 0 is replaced with three 0's and each 1 is replaced with three 1's

$$0 \to 000, 1 \to 111$$



- Suppose we transmit encoded bits over a BSC



- How should we design the decoder?

Decoding the 3-Repetition Code

- Suppose we observe $\mathbf{y} = (y_1, y_2, y_3)$ as the output corresponding to the 3-repetition of a single bit b

$$b \to bbb \to (y_1, y_2, y_3)$$

- What values can $\mathbf{y}$ take? Can we deduce the value of b from $\mathbf{y}$?
- Suppose we use the following decoding rule:
  Decide $b = 0$ if $P(0 \text{ sent}|\mathbf{y} \text{ received}) > P(1 \text{ sent}|\mathbf{y} \text{ received})$
  Decide $b = 1$ if $P(0 \text{ sent}|\mathbf{y} \text{ received}) \leq P(1 \text{ sent}|\mathbf{y} \text{ received})$
- Assume $P(0 \text{ sent}) = P(1 \text{ sent}) = \frac{1}{2}$

$$P(0 \text{ sent}|\mathbf{y} \text{ received}) \underset{1}{\overset{0}{\gtrless}} P(1 \text{ sent}|\mathbf{y} \text{ received})$$

$$\iff \frac{P(\mathbf{y} \text{ received}|0 \text{ sent})P(0 \text{ sent})}{P(\mathbf{y} \text{ received})} \underset{1}{\overset{0}{\gtrless}} \frac{P(\mathbf{y} \text{ received}|1 \text{ sent})P(1 \text{ sent})}{P(\mathbf{y} \text{ received})}$$

$$\iff P(\mathbf{y} \text{ received}|0 \text{ sent}) \underset{1}{\overset{0}{\gtrless}} P(\mathbf{y} \text{ received}|1 \text{ sent})$$

Decoding the 3-Repetition Code

- $P(111 \text{ received} | 1 \text{ sent}) = (1-p)^3$, $P(101 \text{ received} | 1 \text{ sent}) = p(1-p)^2$
- Let $d(\mathbf{y}, 111)$ be the Hamming distance between $\mathbf{y}$ and 111
  Let $d(\mathbf{y}, 000)$ be the Hamming distance between $\mathbf{y}$ and 000

$$
\begin{aligned}
P(\mathbf{y} \text{ received} | 1 \text{ sent}) &= p^{d(\mathbf{y},111)}(1-p)^{3-d(\mathbf{y},111)} \\
P(\mathbf{y} \text{ received} | 0 \text{ sent}) &= p^{d(\mathbf{y},000)}(1-p)^{3-d(\mathbf{y},000)}
\end{aligned}
$$

- If $p < \frac{1}{2}$, then

$$
\begin{aligned}
P(\mathbf{y} \text{ received} | 0 \text{ sent}) &\underset{1}{\overset{0}{\gtrless}} P(\mathbf{y} \text{ received} | 1 \text{ sent}) \\
\iff d(\mathbf{y}, 000) &\underset{0}{\overset{1}{\gtrless}} d(\mathbf{y}, 111)
\end{aligned}
$$

- This is called the minimum distance decoder

References

- Sections 1.4, 1.5 from *Probability and Random Processes*, G. Grimmett and D. R. Stirzaker, 2001 (3rd Edition)
- Chapter 1 from *The Pleasures of Probability*, Richard Isaac, 1995