KTH Computer Science
and Communication

Monty Hall Problem

Problem definition

In a TV programme you are given the chance of winning a sport car if you pick the right door to open among three options (see Figure 1). The car is said to have an equal probability of being hidden behind each door. You are first asked to choose a door, say door A, as indicated by the white arrow in the figure. Subsequently the programme host (who knows the position of the car) opens one of the remaining doors, (door B in the example). The host will always open a door that does not hide a car, to allow the game to continue. Finally the host asks you to decide if you want to hold on your previously picked door, or if you want to change to the other unopened door (door C in the example).

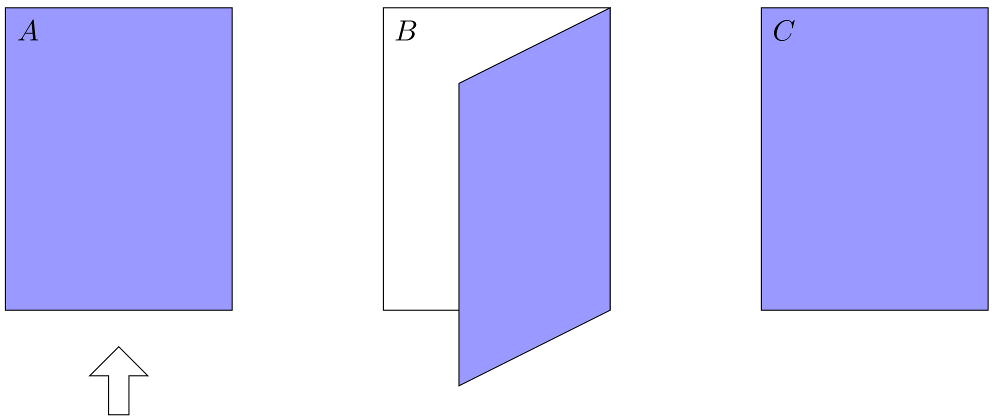

Figure 1. The TV show scenario

You can find more information, including historical aspects, of this problem at `https://en.wikipedia.org/wiki/Monty_Hall_problem`

Solution

This problem often generates a lot of confusion and discussions. The source of misunderstanding is the assumption that the event *the host chooses to open door* X and the event *the car is not behind door* X are equivalent. In reality the first event brings more information to the player than the second, as we will see.

Calling the doors A, B and C, we will indicate the event *the car is behind door* A as C_A, (C_B and C_C are respectively for doors B and C). The event *the host chooses to open door* A is indicated by H_A and, respectively H_B and H_C for doors B and C.

According to the scenario, the initial (prior) probability of the car being behind each of the doors is equal: $P(C_A) = P(C_B) = P(C_C) = \frac{1}{3}$.

We will now estimate the conditional probabilities that the host will open a certain door given the position of the car and your initial choice:

[I] The host cannot open the door you have chosen (door A), regardless the position of the car:

$$P(H_A|C_i) = 0,\ \forall i \in \{A, B, C\}$$

This also means that

$$P(H_A) = P(H_A|C_A)P(C_A) + P(H_A|C_B)P(C_B) + P(H_A|C_C)P(C_C) = 0$$

[II] If the car is behind door i then the host cannot open that door:

$$P(H_i|C_i) = 0,\ \forall\ i \in \{A, B, C\}$$

[III] The host chooses randomly among the doors that are allowed (not chosen by you, and not hiding the car). The consequence of this is that the probabilities of opening a certain door have to be equally distributed among the available doors.

If we use these observations in the three possible cases we obtain:

1) if the car is behind door A:

$$\begin{aligned} P(H_A|C_A) &= 0 \ \text{[either I, or II]} \\ P(H_B|C_A) &= \frac{1}{2} \ \text{[III]} \\ P(H_C|C_A) &= \frac{1}{2} \ \text{[III]} \end{aligned}$$

2) if the car is behind door B:

$$\begin{aligned} P(H_A|C_B) &= 0 \ \text{[I]} \\ P(H_B|C_B) &= 0 \ \text{[II]} \\ P(H_C|C_B) &= 1 \ \text{[probabilities sum to 1]} \end{aligned}$$

3) if the car is behind door C

$$\begin{aligned} P(H_A|C_C) &= 0 \ \text{[I]} \\ P(H_B|C_C) &= 1 \ \text{[probabilities sum to 1]} \\ P(H_C|C_C) &= 0 \ \text{[II]} \end{aligned}$$

Note that in the first case the host can open indifferently the second or third door, while in the other two cases he is forced to open the only door that doesn't hide the car. This should already make you realize that the host's choice might bring you (on average) some useful information about the location of the car.

We can calculate the probability of the car being behind a specific door i, given the host's choice of door j, by applying Bayes rule:

$$P(C_i|H_j) = \frac{P(H_j|C_i)P(C_i)}{P(H_j)} = \frac{P(H_j|C_i)P(C_i)}{P(H_j|C_A)P(C_A) + P(H_j|C_B)P(C_B) + P(H_j|C_C)P(C_C)}$$

where all the terms in the formula are now known.

Because of the example scenario, we are only interested in two probabilities: $P(C_A|H_B)$ and $P(C_C|H_B)$. Substituting the values obtained above in those two cases we obtain:

$$\begin{aligned} P(C_A|H_B) &= \frac{\frac{1}{2}\frac{1}{3}}{\frac{1}{3}\left(\frac{1}{2}+0+1\right)} = \frac{1}{3} \\ P(C_C|H_B) &= \frac{1\frac{1}{3}}{\frac{1}{3}\left(\frac{1}{2}+0+1\right)} = \frac{2}{3}, \end{aligned}$$

showing that, after door B has been opened, you should always decide to change your selection from door A to door C if you want to maximize the probability to win the car.