Discussion: Prisoners-Choice And Medical Testing Problems

Created: 09/03/2023 by Tom Lever
Updated: 09/03/2023 by Tom Lever

How is the Prisoners-Choice problem similar to and different from the Monty-Hall problem, and how do these differences and similarities affect the solution?

Consider the Monty-Hall problem as presented by https://www.cs.unm.edu/~forrest/classes/cs365/lectures/Bayes-1.pdf . You are a contestant on a game show. There are 3 doors A, B, and C. There is a new car behind one door and goats behind the other two doors. Monty asks you to pick a door. You pick door A. Monty tells you that Monty will open another door that has a goat. Monty opens door B. Monty gives you a choice to stay with door A or switch to door C. Which option do you choose?

Consider the following three hypotheses / mutually exclusive parameters.

1. h_1: Car is behind door A.
2. h_2: Car is behind door B.
3. h_3: Car is behind door C.

Consider the prior probabilities of these hypotheses.

1. $P(h_1) = \frac{1}{3}$
2. $P(h_2) = \frac{1}{3}$
3. $P(h_3) = \frac{1}{3}$

Consider data D: You choose door A. Monty opened door B and found a goat.

Consider the likelihoods of D given each hypothesis.

1. $P(D|h_1) = \frac{1}{2}$: The likelihood that Monty will open door B given that the contestant chose door A and that the car is behind door A equals $\frac{1}{2}$.
2. $P(D|h_2) = 0$: The likelihood that Monty will open door B given that the contestant chose door A and that the car is behind door B equals 0.
3. $P(D|h_3) = 1$: The likelihood that Monty will open door B given that the contestant chose door A and that the car is behind door C equals 1.

Consider the total probability of D. By the Law Of Total Probability,

$$P(D) = \sum_{i=1}^{3} [P(D \text{ and } h_i)]$$
$$P(D) = P(D \text{ and } h_1) + P(D \text{ and } h_2) + P(D \text{ and } h_3)$$
$$P(D) = P(D|h_1)P(h_1) + P(D|h_2)P(h_2) + P(D|h_3)P(h_3)$$
$$P(D) = \left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + (0)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{3}\right)$$
$$P(D) = \frac{1}{2}$$

By the Bayes Theorem,

$$P(h_1|D) = \frac{P(D|h_1)P(h_1)}{P(D)}$$

$$P(h_1|D) = \frac{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)}{\left(\frac{1}{2}\right)}$$

$$P(h_1|D) = \frac{1}{3}$$

$$P(h_3|D) = \frac{P(D|h_3)P(h_3)}{P(D)}$$

$$P(h_3|D) = \frac{(1)\left(\frac{1}{3}\right)}{\left(\frac{1}{2}\right)}$$

$$P(h_3|D) = \frac{2}{3}$$

We choose to switch to door C.

Consider the Prisoners-Choice problem. Prisoners A, B, and C are in jail. The jailer tells them that one will be executed and the other two set free. Prisoner A asks the jailer to tell Prisoner A the name of one of the prisoners to be set free. The jailer refuses. The jailer says that Prisoner A's probability of execution would go from $\frac{1}{3}$ to $\frac{1}{2}$. Is the jailer right?

Consider the following three hypotheses / mutually exclusive parameters.

1. h_1: Prisoner A will be executed.
2. h_2: Prisoner B will be executed.
3. h_3: Prisoner C will be executed.

Consider the prior probabilities of these hypotheses.

1. $P(h_1) = \frac{1}{3}$
2. $P(h_2) = \frac{1}{3}$
3. $P(h_3) = \frac{1}{3}$

Consider data T: You are Prisoner A. The jailer tells Prisoner A that Prisoner B will be set free.

Consider the likelihoods of D given each hypothesis.

1. $P(T|h_1) = \frac{1}{2}$: The likelihood that the jailer tells Prisoner A that Prisoner B will be set free given that Prisoner A will be executed equals $\frac{1}{2}$.
2. $P(T|h_2) = 0$: The likelihood that the jailer tells Prisoner A that Prisoner B will be set free given that Prisoner B will be executed equals 0.
3. $P(T|h_3) = 1$: The likelihood that the jailer tells Prisoner A that Prisoner B will be set free given that Prisoner C will be executed equals $\frac{1}{2}$.

Consider the total probability of T. By the Law Of Total Probability,

$$P(T) = \sum_{i=1}^{3} [P(T \text{ and } h_i)]$$
$$P(T) = P(T \text{ and } h_1) + P(T \text{ and } h_2) + P(T \text{ and } h_3)$$
$$P(T) = P(T|h_1)P(h_1) + P(T|h_2)P(h_2) + P(T|h_3)P(h_3)$$
$$P(T) = \left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + (0)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{3}\right)$$
$$P(T) = \frac{1}{3}$$

By the Bayes Theorem,

$$P(h_1|T) = \frac{P(T|h_1)P(h_1)}{P(T)}$$
$$P(h_1|T) = \frac{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)}{\left(\frac{1}{3}\right)}$$
$$P(h_1|T) = \frac{1}{2}$$

$$P(h_3|T) = \frac{P(T|h_3)P(h_3)}{P(T)}$$
$$P(h_3|T) = \frac{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)}{\left(\frac{1}{3}\right)}$$
$$P(h_3|T) = \frac{1}{2}$$

The Monty-Hall problem and the Prisoners-Choice problem are similar in that:

- They deal with three entities labeled A, B, and C.
- They each involve two different outcomes. One outcome is more extreme than the other.
- In spirit, they each involve revealing data.
- The hypotheses, prior probabilities, data, first two likelihoods, calculations of total probability, and calculations of posterior probability are similar.

The Monty-Hall problem and the Prisoner-Choice problem are different in that:

- They deal with three doors and three prisoners.
- The extreme outcome in the Monty-Hall problem is an incentive. The extreme outcome in the Prisoner-Choice problem is a disincentive.
- Technically, in the Monty-Hall problem, data is revealed. In the Prisoner-Choice problem, data is not revealed. However, setting data in both problems allows us to see the difference between these two problems.
- Monty may not open the door of the contestant; the likelihood that Monty will open door B given that the contestant chose door A and that the car is behind door C equals 1. The jailer may tell Prisoner A that Prisoner A will be executed; the probability that the jailer tells Prisoner A that Prisoner B will be set free given that Prisoner C will be executed equals $\frac{1}{2}$.

- For the Monty-Hall problem, the total probability $P(D) = \frac{1}{2}$. For the Prisoner-Choice problem, the total probability $P(T) = \frac{1}{3}$.
- The posterior probability that a car is behind door A given that Monty opened door B to reveal a goat is $\frac{1}{3}$. The posterior probability that a car is behind door C given that Monty opened door C to reveal a goat is $\frac{2}{3}$. The posterior probability that Prisoner A will be executed given that the jailer told Prisoner A the name of a prisoner who will be set free is $\frac{1}{2}$. The posterior probability that Prisoner C will be executed given that the jailer told Prisoner A the name of a prisoner who will be set free is $\frac{1}{2}$.
- Prisoner A cannot change Prisoner A's identity, nor is given the opportunity.

Suppose in the prisoner's choice problem that the jailer tells us that prisoner B will be set free. Does that change the probability that prisoner C will be set free?

Yes. As above, the posterior probability that Prisoner C will be executed given that the jailer told Prisoner A the name of a prisoner who will be set free is $\frac{1}{2}$, which is different than the prior probability that Prisoner C will be executed.

In the medical testing problem, compare the probability that the patient has ebola using frequentist and Bayesian probabilities.

Consider the medical testing problem. Suppose an Ebola test is available with $FPR = 0.023$ and $FNR = 0.014$. You have no reason to think you have Ebola, but you take the test and it comes back positive. If the prevalence of Ebola in the general population is 1/10000, what is the probability you have Ebola based on this test?

We assume the prevalence of Ebola in the general population is 1/10000. In the margin of the below confusion matrix, let's assume that $n_{Ebola} = 1000$ people have Ebola and 9,999,000 people do not have Ebola. The False Negative Rate $FNR = \frac{FN}{n_{Ebola}} = 0.014$. The number of False Negatives $FN = FNR \cdot n_{Ebola} = 0.014 \cdot 1000 = 14$. The number of True Positives $TP = n_{Ebola} - FN = 1000 - 14 = 986$. $FPR = \frac{FP}{n_{no\ Ebola}} = 0.023$. The number of False Positives $FP = FPR \cdot n_{no\ Ebola} = 0.023 \cdot 9999000 = 229{,}977$. The number of True Negatives $TN = n_{no\ Ebola} - FP = 9999000 - 229977 = 9769023$. The number of people testing negative $n_{testing\ negative} = TN + FN = 9769037$. The number of people testing positive $n_{testing\ positive} = FP + TP = 230963$.

		Has Ebola?		
		N	Y	
Tests positive?	N	$TN = 9769023$	$FN = 14$	$n_{testing\ negative} = 9769037$
	Y	$FP = 229977$	$TP = 986$	$n_{testing\ positive} = 230963$
		$n_{no\ Ebola} = 9999000$	$n_{Ebola} = 1000$	$n = 10000000$

The probability that you have Ebola based on testing positive

$$P(\text{Ebola} \mid \text{testing positive}) = \frac{n_{testing\ positive\ and\ Ebola}}{n_{testing\ positive}}$$

$$P(\text{Ebola} \mid \text{testing positive}) = \frac{986}{230963}$$
$$P(\text{Ebola} \mid \text{testing positive}) = 0.00427$$

$$P(\text{testing positive} \mid \text{Ebola}) = TPR = 1 - FNR = 0.986$$
$$P(\text{Ebola}) = 0.0001$$
$$P(\text{testing positive}) = \frac{n_{testing\ positive}}{n} = \frac{230963}{10000000} = 0.0231$$
$$P(\text{Ebola} \mid \text{testing positive}) = \frac{0.986 \cdot 0.0001}{0.0231}$$
$$P(\text{Ebola} \mid \text{testing positive}) = 0.00427$$

The probabilities that the patient has Ebola using frequentist and Bayesian probabilities are identical.