01b - Probability - 02

Master in Foundations of Data Science
Bayesian Statistics and Probabilistic Programming
Fall 2018-2019

Josep Fortiana

Wednesday, September 26, 2018

Universitat de Barcelona

The Monty Hall problem

The double dice problem

01b - Probability 02

The Monty Hall problem

The double dice problem

Setting

A TV game show. Choose one of three closed doors.

One of them hides a hefty prize (a car).

If you choose it you win the car.

Behind each of the other two doors there is a goat.

If you choose either of them you win nothing.

Procedure

You choose a door. For the time being, it remains closed.

The game show host (Monty Hall), who knows where the car is hidden, opens one of the unselected doors, showing a goat.

Then you have the opportunity to maintain your initial choice or to switch to the remaining closed door.

Which is the best strategy?

Stick to the first choice or change?

Or, perhaps, it is indifferent?

Once a door has been selected, for sure at least one of the other two hides a goat, hence opening one of them supplies no new information.

Thus we can conclude that switching will not affect chances of winning.

Which is the best strategy?

Stick to the first choice or change?

Or, perhaps, it is indifferent?

Once a door has been selected, for sure at least one of the other two hides a goat, hence opening one of them supplies no new information.

Thus we can conclude that switching will not affect chances of winning.

Or does it?

Experiment

Perform an experiment, simulating many repetitions of the game with each strategy.

We observe results are VERY different.

Switching doors, the prize is won
DOUBLE number of times
as in the conservative strategy.

A Bayesian solution: 1. Hypotheses, prior probabilities

$\Longrightarrow$ Notebooks for *Think Bayes* by Roger Labbe

Three hypotheses H:

$$H = A, \quad H = B, \quad \text{and} \quad H = C,$$

respectively representing the (unknown but possible) fact that the car is behind the namesake door.

Initially, the *prior probabilities* $\mathsf{P}(H)$ are:

$$\mathsf{P}(A) = \mathsf{P}(B) = \mathsf{P}(C) = 1/3.$$

2. Observed data

AFTER picking your door, relabel the three doors so that A is your choice.

Assume the observed data is:

$$D = \boxed{\begin{array}{c}\text{Monty has opened door } B\\ \text{AND}\\ \text{inside door } B \text{ there is a goat.}\end{array}}$$

3. Likelihood

Conditional probabilities $\mathrm{P}(D|H)$ of the observed data D, assuming each possible hypothesis H as being true:

	Prior $\mathrm{P}(H)$	Likelihood $\mathrm{P}(D\|H)$
A	$1/3$	$1/2$
B	$1/3$	0
C	$1/3$	1

4. Joint probabilities

Multiply each $\mathrm{P}(D|H)$ times the prior $\mathrm{P}(H)$, giving the Joint or intersection probabilities $\mathrm{P}(D \cap H) \equiv \mathrm{P}(D, H)$:

	Prior $\mathrm{P}(H)$	Likelihood $\mathrm{P}(D\|H)$	Joint $\mathrm{P}(D, H)$
A	$1/3$	$1/2$	$1/6$
B	$1/3$	0	0
C	$1/3$	1	$1/3$

5. Bayes' formula $\Longrightarrow$ Posterior probabilities

Divide $\mathrm{P}(D, H)$ by their sum $\mathrm{P}(D) = 1/2$, giving the posterior probabilities $\mathrm{P}(H|D) = \mathrm{P}(D, H)\,/\mathrm{P}(D)$:

	Prior $\mathrm{P}(H)$	Likelihood $\mathrm{P}(D\|H)$	Joint $\mathrm{P}(D, H)$	Posterior $\mathrm{P}(H\|D)$
A	$1/3$	$1/2$	$1/6$	$1/3$
B	$1/3$	0	0	0
C	$1/3$	1	$1/3$	$2/3$

01b - Probability 02

The Monty Hall problem

The double dice problem

Setting

Suppose I have a box that contains one each of 4-sided, 6-sided, 8-sided, and 12-sided dice.

I choose a die at random, and roll it twice without letting you see the die or the outcome.

1. What is the probability of obtaining the same outcome of both rolls?

Observed data and its effects

Now, I report that I got the same outcome on both rolls.

2. What is the *(posterior)* probability that I rolled each of the dice?

Observed data and its effects

Now, I report that I got the same outcome on both rolls.

2. What is the *(posterior)* probability that I rolled each of the dice?

3. If I roll the same die again, what is the probability that I get the same outcome a third time?

Obtaining the likelihood

See: Probably Overthinking It. A blog by Allen Downey.

1. Considering the results of both rolls, denoting by $S2$ the event:

$$S2 = \text{“Same outcome on both rolls”},$$

the probability of $S2$, conditional to having chosen an m-sided die, is:

$$\mathrm{P}(S2|m) = \frac{\text{“Favorable outcomes”}}{\text{“Total outcomes”}} = \frac{m}{m^2} = \frac{1}{m}.$$

Prior probabilities

There are $k = 4$ dice, randomly chosen with prior probabilities $\mathrm{P}(m) = \dfrac{1}{4}$, $m = 4, 6, 8, 12$, i.e., the prior distribution is:

$$\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right).$$

Total probability

The total probability of $S2$ is:

$$\mathrm{P}(S2) = \frac{1}{4} \times \left(\frac{1}{4} + \frac{1}{6} + \frac{1}{8} + \frac{1}{12} \right) = \frac{5}{32}.$$

Posterior probabilities

2. From Bayes' formula we obtain the posterior probabilities:

$$P(m|S2) = \frac{P(S2|m) \cdot P(m)}{P(S2)} = \frac{1/m}{5/8} = \frac{8}{5\,m},$$

for $m = 4, 6, 8, 12.$

Posterior probability mass function (pmf)

Substituting each $m = 4, 6, 8, 12$ we get the posterior distribution:

$$\begin{cases} m = 4, & \to \ 8/20 \ = \ 6/15, \\ m = 6, & \to \ 8/30 \ = \ 4/15, \\ m = 8, & \to \ 8/40 \ = \ 3/15, \\ m = 12, & \to \ 8/60 \ = \ 2/15. \end{cases}$$

Posterior prediction

3. With these posterior probabilities we *predict* the total probability of:

$$S3 = \text{"Same outcome a third time"},$$

we proceed as above.

Likelihood for the third roll

The probability of $S3$, conditional to having chosen an m-sided die, is:

$$\mathrm{P}(S3|m) = \frac{\text{``Favorable outcomes''}}{\text{``Total outcomes''}} = \frac{1}{m},$$

since now there is only a single favorable outcome, namely the same result obtained in both previous rolls.

Total (posterior) probability

The total probability of $S3$ is:

$$
\begin{aligned}
\mathrm{P}(S3) &= \mathrm{P}(S3|4)\cdot\mathrm{P}(4)+\mathrm{P}(S3|6)\cdot\mathrm{P}(6)\\
&\qquad+\mathrm{P}(S3|8)\cdot\mathrm{P}(8)+\mathrm{P}(S3|12)\cdot\mathrm{P}(12)\\
&=\frac{1}{4}\times\frac{6}{15}+\frac{1}{6}\times\frac{4}{15}+\frac{1}{8}\times\frac{3}{15}+\frac{1}{12}\times\frac{2}{15}\\
&=\frac{13}{72}.
\end{aligned}
$$