# The Monty Hall 3-door Puzzle

You have a chance to win a large prize in a game show hosted by Monty.

The prize is behind one of 3 closed doors, and the other two doors are losers, with a goat behind each.

- You are asked to select (but not to open yet) one of the doors.
- Then Monty, who knows what is behind each door, does the following:

  Whether or not you selected the winning door, he opens one of the other two doors that he knows for sure is a losing door (selecting random if both are losing doors).
- Then he asks you whether you would like to switch doors.

**Which strategy should you follow to have a better chance?**

- Should you switch doors?
- Or stick to your original selection?
- Or does it not matter?

# Probability theory

An **experiment** is a procedure that yields one of a given set of possible outcomes.

The **sample space** of the experiment is the set of possible outcomes:

$$S = \{s_1, s_2, \ldots, s_n\}$$

A **probability distribution** in the sample space $S$ is a function $p$ assigning a number $p(s_i)$ to each possible outcome $s_i$ in $S$ (the **probability** of $s_i$) such the following two conditions hold:

- $0 \leq p(s_i) \leq 1$, for every $i = 1, 2, \ldots, n$,
- $p(s_1) + p(s_2) + \cdots + p(s_n) = 1$.

FOR EXAMPLE: the **uniform distribution** on $S$

$$p(s_i) = \frac{1}{n}, \quad \text{for every } i = 1, 2, \ldots, n.$$

If a *fair* dice is rolled, then there are six possible equally likely outcomes: $1, 2, 3, 4, 5, 6$

So $\boxed{p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = \frac{1}{6}}$

## Probability distribution: two further examples

(1) If a *fair* coin is flipped, then there are two equally likely outcomes:

$H$ (heads) and $T$ (tails).

So $\boxed{p(H) = p(T) = \frac{1}{2}}$.

(2) But what probabilities should we assign to these outcomes when the coin is *biased* so that heads comes up twice as often as tails?

SOLUTION: We have $p(H) = 2p(T)$. Since $p(H) + p(T) = 1$ should hold, it follows that $2p(T) + p(T) = 3p(T) = 1$.

So $\boxed{p(T) = \frac{1}{3}}$ and $\boxed{p(H) = \frac{2}{3}}$.

# Events and their probabilities

An **event** is a subset of the sample space, that is, a set of possible outcomes.

The **probability of an event $E$** is the sum of the probabilities of the outcomes in $E$.

FOR EXAMPLE: What is the probability that, when two fair dice are rolled, the sum of the two numbers on them is 7 ?

By the product rule, there are a total of $6 \cdot 6 = 36$ possible outcomes, each of them equally likely with a probability of $\frac{1}{36}$.

There are $6$ 'successful' outcomes:

$$(1,6), \quad (2,5), \quad (3,4), \quad (4,3), \quad (5,2), \quad (6,1).$$

Therefore, the probability of having $7$ as the sum is

$$\frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36} = 6\cdot\frac{1}{36} = \boxed{\frac{1}{6}}.$$

# Probability of events: another example

Suppose that a dice is loaded so that 3 appears twice as often as each of the other numbers, but the other five outcomes are equally likely.

What is the probability that an odd number appears when we roll this dice?

SOLUTION: We want to find the probability $p(E)$ of the event $E = \{1, 3, 5\}$.
We know that

- $p(1) + p(2) + p(3) + p(4) + p(5) + p(6) = 1$,
- $p(1) = p(2) = p(4) = p(5) = p(6)$, and
- $p(3) = 2p(1)$.

Therefore, $p(1) = p(2) = p(4) = p(5) = p(6) = \frac{1}{7}$ and $p(3) = \frac{2}{7}$.

It follows that $\boxed{p(E) = p(1) + p(3) + p(5) = \frac{4}{7}}$.

# Probability of events: yet another example

In a certain kind of lottery, people can win the top prize if they correctly choose a set of 6 numbers out of the first 59 positive natural numbers.

What is the probability of winning the top prize?

SOLUTION: The total number of ways to choose 6 numbers out of 59 is

$$\binom{59}{6} = \frac{59 \cdot 58 \cdot 57 \cdot 56 \cdot 55 \cdot 54}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} = 45\,057\,474\,.$$

Each of these outcomes is equally likely. The event $W$ of winning the top prize contains just one of these. Consequently,

$$\boxed{p(W) = \frac{1}{45\,057\,474} \approx 0.000000022}\,.$$

# Probability of the complementary event

Let $E$ be an event in a sample space $S$.

The probability of event $\bar{E}$, the **complement of $E$ in $S$**, is the sum of the probabilities of the outcomes *not* in $E$:

$$\boxed{p(\bar{E}) \;=\; 1 - p(E)}$$

FOR EXAMPLE: We flip a fair coin 10 times.
What is the probability that heads comes up at least once?

Let $E$ be the event that heads comes up at least once of the 10 flips.
Then $\bar{E}$ is the event that each of the 10 times tails comes up. So

$$p(E) \;=\; 1 - p(\bar{E}) \;=\; 1 - \frac{1}{2^{10}} \;=\; 1 - \frac{1}{1024} \;=\; \frac{1023}{1024} \;\approx\; 0.999\,.$$

## Probability of the union of events

$$\boxed{p(E_1 \cup E_2) \;=\; p(E_1) + p(E_2) - p(E_1 \cap E_2)}$$

FOR EXAMPLE: What is the probability that a number randomly selected from the set $X = \{n \in \mathbf{N}^+ \mid n \leq 100\}$ is divisible by either $2$ or $5$ or by both (by $10$)?

- let $E_1$ be the event that the selected number in $X$ is divisible by $2$
- let $E_2$ be the event that the selected number in $X$ is divisible by $5$
- then $E_1 \cap E_2$ is the event that it is divisible by both $2$ and $5$, that is, it is divisible by $10$
- as $|X| = 100$, $|E_1| = 50$, $|E_2| = 20$, and $|E_1 \cap E_2| = 10$, we obtain:

$$p(E_1 \cup E_2) \;=\; \frac{50}{100} + \frac{20}{100} - \frac{10}{100} \;=\; \frac{60}{100} \;=\; \boxed{\frac{3}{5}} \;=\; \boxed{0.6}$$

# Conditional probability

We flip a fair coin $3$ times. We already know how to calculate the probability of the event $E$ that *at least 2 heads come up in a row:*

The sample space $S$ consists of $2^3 = 8$ outcomes, each with probability $\frac{1}{8}$.

$3$ of them is in $E$: $HHT$, $THH$, $HHH$. So $p(E) = \boxed{\frac{3}{8}}$.

But what if the *first flip comes up heads* (event $F$ occurs), and we are asked about the probability of $E$, *knowing* that $F$ has already occurred?

Now the sample space is $F$: it consists of outcomes $HHT$, $HTH$, $HTT$, $HHH$

For an outcome in $E$ to occur, it must belong to $E \cap F$: either $HHT$ or $HHH$

The **conditional probability of $E$ given $F$** is

$$\boxed{p(E \mid F)} = \frac{|E \cap F|}{|F|} = \frac{\frac{|E \cap F|}{|S|}}{\frac{|F|}{|S|}} = \boxed{\frac{p(E \cap F)}{p(F)}}$$

IN THE EXAMPLE:

We have $p(F) = \frac{4}{8} = \frac{1}{2}$ and $p(E \cap F) = \frac{2}{8} = \frac{1}{4}$. So $p(E \mid F) = \frac{\frac{1}{4}}{\frac{1}{2}} = \boxed{\frac{1}{2}}$.

# Conditional probability: another example

We throw a pair of fair dice. What is the probability that at least one dice is a 3, given that the sum of the two dice is 5?

SOLUTION:

- let $A$ be the event that at least one dice is a 3
- let $B$ the event that the sum of the two dice is 5, that is, $B$ consists of the outcomes $(1,4)$, $(2,3)$, $(3,2)$, $(4,1)$.
- as the sample space consists of $6^2 = 36$ equally probable outcomes, $p(B) = \frac{4}{36} = \frac{1}{9}$
- the event $A \cap B$ consists of two outcomes: $(2,3)$ and $(3,2)$, and so $p(A \cap B) = \frac{2}{36} = \frac{1}{18}$

Therefore,

$$\boxed{p(A \mid B) = \frac{p(A \cap B)}{p(B)} = \frac{\frac{1}{18}}{\frac{1}{9}} = \frac{1}{2}}$$

Observe that the unconditional probability of $A$ is different: $p(A) = \frac{11}{36}$

# Conditional probability: yet another example

We flip a fair coin $3$ times. What is the probability that an odd number of tails appear, knowing that the first flip comes up tails?

SOLUTION:

- let $X$ be the event that <u>an odd number of tails appear</u>
- let $Y$ be the event that <u>the first flip comes up tails</u>
- as the sample space consists of $2^3 = 8$ equally probable outcomes, $p(Y) = \frac{4}{8} = \frac{1}{2}$
- the event $X \cap Y$ consists of two outcomes: $THH$ and $TTT$, and so $p(X \cap Y) = \frac{2}{8} = \frac{1}{4}$

Therefore,

$$\boxed{p(X \mid Y) = \frac{p(X \cap Y)}{p(Y)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}}$$

Observe that this time the unconditional probability of $X$ is the same: $p(X) = \frac{4}{8} = \frac{1}{2}$

# Independence

When two events $E$ and $F$ are **independent,** the occurrence of one of the events gives no information about the probability of the other event:

$$\boxed{p(E \mid F) \ = \ p(E)} \qquad \text{or equivalently,} \qquad \boxed{p(F \mid E) \ = \ p(F)}$$

Using that $p(E \mid F) \ = \ \frac{p(E \cap F)}{p(F)}$, we obtain that $E$ and $F$ are independent events, whenever

$$\boxed{p(E \cap F) \ = \ p(E) \cdot p(F)}$$

FOR EXAMPLE:

- Events $E$ and $F$ on lecture slide 143 are not independent.
- Events $A$ and $B$ on lecture slide 144 are not independent.
- Events $X$ and $Y$ on lecture slide 145 are independent.

# Bayes' Theorem

This might be a very useful tool when

- $E$ and $F$ are events from the same sample space $S$
- we can easily compute $p(E \mid F)$
- but we want to know $\underline{p(F \mid E)}$

$$\boxed{p(F \mid E) \;=\; \frac{p(E \mid F)p(F)}{p(E \mid F)p(F) \;+\; p(E \mid \bar{F})p(\bar{F})}}$$

# Applying Bayes' Theorem: An example

A frog's climbing out of a well is affected by the weather. When it rains, he falls back down the well with a probability of $\frac{1}{10}$. In dry weather, he only falls back down with probability of $\frac{1}{25}$. The probability of rain is $\frac{1}{5}$.

If we know that the frog fell back, what is the probability that it was a rainy day?

SOLUTION: Let $\boxed{F}$ be the event that *the frog fell back,* and

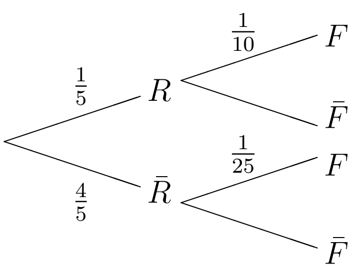
$\boxed{R}$ be the event that *it was a rainy day.*

We want to find $p(R \mid F)$.

- $\boxed{p(R) = \frac{1}{5}}$ $\boxed{p(\bar{R}) = 1 - \frac{1}{5} = \frac{4}{5}}$
- $\boxed{p(F \mid R) = \frac{1}{10}}$ $\boxed{p(F \mid \bar{R}) = \frac{1}{25}}$

$$p(R \mid F) = \frac{p(F \mid R)p(R)}{p(F \mid R)p(R) + p(F \mid \bar{R})p(\bar{R})} =$$

$$\frac{\frac{1}{10} \cdot \frac{1}{5}}{\frac{1}{10} \cdot \frac{1}{5} + \frac{1}{25} \cdot \frac{4}{5}} = \frac{\frac{1}{50}}{\frac{1}{50} + \frac{4}{125}} = \frac{\frac{25}{1250}}{\frac{25}{1250} + \frac{40}{1250}} = \frac{5}{13}$$



# Applying Bayes' Theorem: Example 2

There are two boxes, $Box_1$ and $Box_2$.

- $Box_1$ contains 2 green balls and 7 red balls
- $Box_2$ contains 4 green balls and 3 red balls

Bob is blindfolded, and selects a ball

- by randomly choosing one of the boxes first,
- then randomly choosing a ball from it.

If we know that Bob has selected a red ball, what is the probability that he selected a ball from $Box_1$?

## Applying Bayes' Theorem: Example 2 (cont.)

If we know that Bob has selected a red ball, what is the probability that he selected a ball from $Box_1$ ?

SOLUTION: Let $\boxed{R}$ be the event that *Bob has chosen a red ball*, and
$\boxed{F}$ be the event that *Bob has chosen a ball from $Box_1$*.

We want to find $p(F \mid R)$.

- $p(R \mid F) = \frac{7}{9}$. as *$Box_1$* contains 2 green balls and 7 red balls
- $\boxed{\bar{F}}$ : is the event that *Bob has chosen a ball from $Box_2$*.
  Then $p(F) = p(\bar{F}) = \frac{1}{2}$.
- $p(R \mid \bar{F}) = \frac{3}{7}$. as *$Box_2$* contains 4 green balls and 3 red balls

Then by Bayes' Theorem:

$$p(F \mid R) = \frac{\frac{7}{9} \cdot \frac{1}{2}}{\frac{7}{9} \cdot \frac{1}{2} + \frac{3}{7} \cdot \frac{1}{2}} = \frac{\frac{7}{9}}{\frac{49+27}{63}} = \frac{\frac{7}{9}}{\frac{76}{63}} = \frac{49}{76} \approx 0.645$$

# Applying Bayes' Theorem: Example 3

Suppose that 1 person in 100 000 has a particular rare disease for which there is a quite accurate diagnostic test:

- It is correct 99% of the time when given to someone with the disease.
- It is correct 99.5% of the time when given to someone who does not have the disease.

What is the probability that <u>someone who tests positive for the disease actually has the disease?</u>

SOLUTION: Let $\boxed{D}$ be the event that a person has the disease, and $\boxed{T}$ be the event that a person tests positive for the disease.

- $p(T \mid D) = 0.99$
- $p(\bar{T} \mid \bar{D}) = 0.995$, so $p(T \mid \bar{D}) \overset{Ex.5.5}{=} 1 - p(\bar{T} \mid \bar{D}) = 1 - 0.995 = 0.005$
- $p(D) = \frac{1}{100\,000} = 0.00001$, so $p(\bar{D}) = 1 - 0.00001 = 0.99999$
- $\underline{p(D \mid T)} = \frac{0.99 \cdot 0.00001}{0.99 \cdot 0.00001 + 0.005 \cdot 0.99999} \approx \underline{0.002}$

# Bayesian spam filters

A Bayesian spam filter uses information about previously seen e-mail messages to guess whether an incoming e-mail is spam. It looks for occurrences of particular words in messages. For a word $w$, the probability that $w$ appears in a spam message is estimated by determining

- the number of times $w$ appears in a message from a large set of messages *known to be spam*, and
- the number of times $w$ appears in a large set of messages that are *known not to be spam*.

Unfortunately, spam filters sometimes fail to identify a spam message as spam: this is called a **false negative**. And they sometimes identify a message that is not spam as spam: this is called a **false positive**.

When testing for spam, it is important to minimise false positives, because filtering out wanted e-mail is much worse than letting some spam through.

## Bayesian spam filters: an example

Suppose we have found that the word "*Rolex*" occurs in 250 of 2000 messages known to be spam, and in 5 of 1000 messages known not to be spam.

Estimate the probability that <u>an incoming message containing the word "*Rolex*" is spam,</u> assuming that it is equally likely that an incoming message is spam or not spam. If our threshold for rejecting a message as spam is 0.9, will we reject this message?

SOLUTION: Let $S$ be the event that an incoming message is spam, and $R$ be the event that the message contains the word "*Rolex*". Then:

- $p(R \mid S) = \frac{250}{2000} = 0.125$
- $p(R \mid \bar{S}) = \frac{5}{1000} = 0.005$
- $p(S) = p(\bar{S}) = 0.5$
- $\boxed{p(S \mid R) = \frac{0.125 \cdot 0.5}{0.125 \cdot 0.5 + 0.005 \cdot 0.5} = \frac{0.125}{0.13} \approx 0.962}$

As $0.962 > 0.9$, we would reject the message as spam.

# The 3-door puzzle revisited: conditional probability directly

You selected a door: let's call it **door**$_1$. Then Monty opened another door:

| Prize location: | | Monty opens: | Overall probability: | What you get if you stick: | switch: |
|---|---|---|---|---|---|
| $\frac{1}{3}$ **door**$_1$ | $\frac{1}{2}$ | **door**$_2$ | $\frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$ | prize | goat |
| | $\frac{1}{2}$ | **door**$_3$ | $\frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$ | prize | goat |
| $\frac{1}{3}$ **door**$_2$ | $1$ | **door**$_3$ | $\frac{1}{3} \cdot 1 = \frac{1}{3}$ | goat | prize |
| $\frac{1}{3}$ **door**$_3$ | $1$ | **door**$_2$ | $\frac{1}{3} \cdot 1 = \frac{1}{3}$ | goat | prize |

As $\frac{1}{3} + \frac{1}{3} = \frac{2}{3} > \frac{1}{3} = \frac{1}{6} + \frac{1}{6}$, you'll have a better chance if you switch from **door**$_1$ to the door Monty did not open.

# The 3-door puzzle revisited: using Bayes theorem

- You selected a door: let's call it **door**$_1$
- Then Monty opened another door: let's call it **door**$_2$
- Let's call the remaining 3rd door **door**$_3$

Consider the following events:

| | |
|---|---|
| $W_1$: **door**$_1$ is the winning door | $W_2$: **door**$_2$ is the winning door |
| $W_3$: **door**$_3$ is the winning door | $M_2$: Monty opened **door**$_2$ |

We are interested in $\boxed{p(\overline{W_1} \mid M_2)}$:

- if $p(\overline{W_1} \mid M_2) > \frac{1}{2}$ then you should switch from **door**$_1$ to **door**$_3$
- if $p(\overline{W_1} \mid M_2) < \frac{1}{2}$ then you should stick to **door**$_1$
- and if $p(\overline{W_1} \mid M_2) = \frac{1}{2}$ then it doesn't matter

# The 3-door puzzle revisited: using Bayes theorem (cont.)

We want to use Bayes' theorem to compute $p(\overline{W_1} \mid M_2)$ :

$$p(\overline{W_1} \mid M_2) \;=\; \frac{p(M_2 \mid \overline{W_1}) \cdot p(\overline{W_1})}{p(M_2 \mid \overline{W_1}) \cdot p(\overline{W_1}) + p(M_2 \mid W_1) \cdot p(W_1)}$$

So we need to compute:

- $p(W_1)$
- $p(\overline{W_1})$
- $p(M_2 \mid W_1)$
- $p(M_2 \mid \overline{W_1})$

# The 3-door puzzle revisited: using Bayes theorem (cont.)

- $\underline{p(W_1)} = p(W_2) = p(W_3) = \frac{1}{3}$

  These are the 3 different outcomes and they are equally likely.

- $\underline{p(\overline{W_1})} = 1 - p(W_1) = 1 - \frac{1}{3} = \frac{2}{3}$

- $\underline{p(M_2 \mid W_1)} = \frac{1}{2}$

  If your selected **door**$_1$ is the winning one, then Monty opens randomly one of the other two doors.

- $\underline{p(M_2 \mid \overline{W_1})} = \dfrac{p(M_2 \cap \overline{W_1})}{p(\overline{W_1})}$

  $p(M_2 \cap \overline{W_1}) = p\big(M_2 \cap (W_2 \cup W_3)\big) = p\big((M_2 \cap W_2) \cup (M_2 \cap W_3)\big) = p(M_2 \cap W_3)$

  $p(M_2 \cap W_3) = p(M_2 \mid W_3) \cdot p(W_3) = 1 \cdot \frac{1}{3} = \frac{1}{3}$

  $p(M_2 \mid W_3) = 1$ because if **door**$_3$ is the winning door and you selected **door**$_1$, then Monty surely opens **door**$_2$ (as he never opens the winning door).

- $p(M_2 \mid \overline{W_1}) = \dfrac{p(M_2 \cap \overline{W_1})}{p(\overline{W_1})} = \dfrac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}$

## The 3-door puzzle revisited: the conclusion by Bayes' theorem

So we have: $p(W_1) = \frac{1}{3}$ $\quad p(\overline{W_1}) = \frac{2}{3}$ $\quad p(M_2 \mid W_1) = \frac{1}{2}$ $\quad p(M_2 \mid \overline{W_1}) = \frac{1}{2}$

Now we can apply Bayes' theorem:

$$\boxed{p(\overline{W_1} \mid M_2)} = \frac{p(M_2 \mid \overline{W_1}) \cdot p(\overline{W_1})}{p(M_2 \mid \overline{W_1}) \cdot p(\overline{W_1}) + p(M_2 \mid W_1) \cdot p(W_1)}$$

$$= \frac{\frac{1}{2} \cdot \frac{2}{3}}{\frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3}} = \frac{\frac{2}{6}}{\frac{3}{6}} = \boxed{\frac{2}{3}}$$

Therefore:

$$\boxed{p(\overline{W}_1 \mid M_2) = \frac{2}{3} > \frac{1}{2}}$$

So you'll have a <u>better chance if you switch</u> from **door**$_1$ to **door**$_3$.