UNIVERSITÀ DEGLI STUDI DI TRIESTE

Statistics

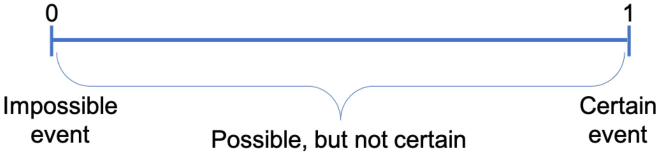
Probability

Sara Geremia

March 28th, 2024

Introduction to probability

Probability helps prevent randomness from being masked and perceived as non-randomness.



Probability is related to uncertain events:

- Ball landing on black on the roulette wheel
- The weather tomorrow in Trieste being sunny
- Inter soccer team winning the Italian league
- The guilt of a defendant

Introduction to probability: subjective interpretation

Bruno de Finetti in "Theory of probability" (1970)

1.3.1. Meanwhile, for those who are not aware of it, it is necessary to mention that in the conception we follow and sustain here only **subjective probabilities** exist – that is, the degree of belief in the occurrence of an event attributed by a given person at a given instant and with a given set of information. This is in contrast to other conceptions that limit themselves to special types of cases in which they attribute meaning to 'objective probabilities' (for instance, cases of symmetry as for dice etc., 'statistical' cases of 'repeatable' events, etc.).

The idea of probability is empirical: it is based on observation rather than theorizing.

Probability models are used to study the variation in observed data so that inferences about the underlying process can be developed.

Introduction to probability

- De Finetti's idea emphasizes that probability is a measure of individual's degree of belief about an event occurring.
- Psychologists have shown experimentally that heuristics, which are essentially mental shortcuts, often lead to errors when assessing uncertainty
- Individuals may rely on these mental shortcuts, which can be influenced by biases and cognitive errors, when assigning probabilities subjectively
- I toss a coin six times and record the outcome of Heads or Tails. Which result is more likely?

$$HHHHTTTT \text{ or } HTTHHTHH$$

Both sequences have an equal probability of occurring in a fair coin toss!

Introduction to probability

- By saying that probability is the 'degree of belief' we have given a good definition, but this is not of help when it comes to numerically determining this degree of belief.
- What is the probability that it will rain tomorrow?
  - By guessing?
  - Average monthly rainfall: 28%, meaning $P(\text{rain}) = 0.28$
  - OSMER: 5%, meaning $P(\text{rain}) = 0.05$

Introduction to probability

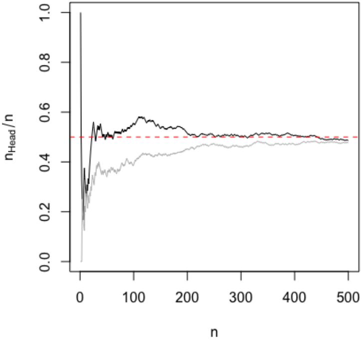
- All evaluations are valid: they are opinions, each legitimate, perhaps some more reasonable or more informed than others.
- For example, OSMER, for a relatively trivial event like 'rain tomorrow,' uses complex meteorological models and combines a lot of information.
- For now, we will look at some cases where the evaluation is simple and intuitive. In such situations, it will be easy to understand some rules for combining probabilities using the four operations and a bit of logic.

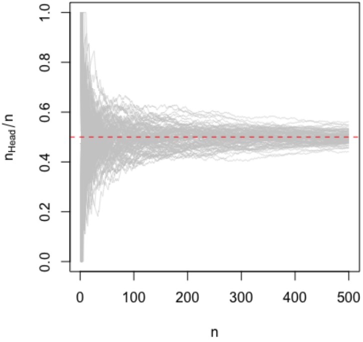
Frequentist interpretation

- Probability is not something directly observable
- To have a way of determining it, we can try to link it to something observable
- Let's consider an event E as E = 'getting Heads (H) in a coin toss'
- This event is repeatable, meaning we can toss a coin many times. Let's do it, or imagine doing it, and at each toss, calculate the percentage of heads observed up to that point.

Frequentist interpretation

Here are the outcomes from 500 tosses: chance behaviour is unpredictable in the short run but has a regular and predictable pattern in the long run





The **probability** of any outcome of a random phenomenon is the proportion of times the outcome would occur in a very long series of repetitions

Random experiment

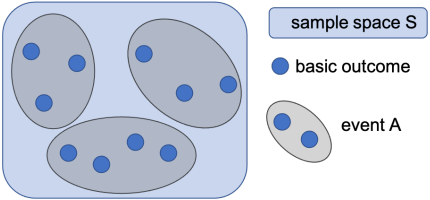
A **random experiment/phenomenon** is a process leading to two or more possible **outcomes**, without knowing exactly which one will occur

Sample space

The **sample space** S of a random experiment is the set of all the possible **basic outcomes**, that is, outcomes that can not occur together

Event

An **event** A is an outcome or a set of possible outcomes



Relative frequency probability

The **relative frequency probability** is the limit of the proportion of times that an event A occurs in a large number of trials, n,

$$P(A) = \frac{n_A}{n}$$

where n_A is the number of A outcomes and n is the total number of trials or outcomes.

Classical probability

The **classical probability** is the proportion of times that an event will occur, assuming that all outcomes in a sample space are equally likely to occur

$$P(A) = \frac{N_A}{N}$$

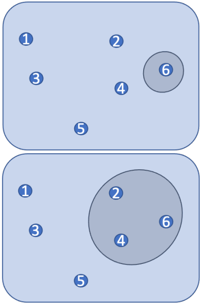
where N_A is the number of outcomes that satisfy A, and N is the total number of outcomes in the sample space.

Random experiment: rolling a dice

Sample space: $S = \{1, 2, 3, 4, 5, 6\}$

Event: $A =$ get a 6
$P(A) = \frac{N_A}{N} = \frac{1}{6}$

Event: $B =$ get an even number $= \{2, 4, 6\}$
$P(B) = \frac{N_B}{N} = \frac{3}{6} = \frac{1}{2}$

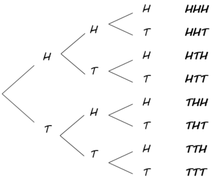


Random experiment: tossing a coin three times

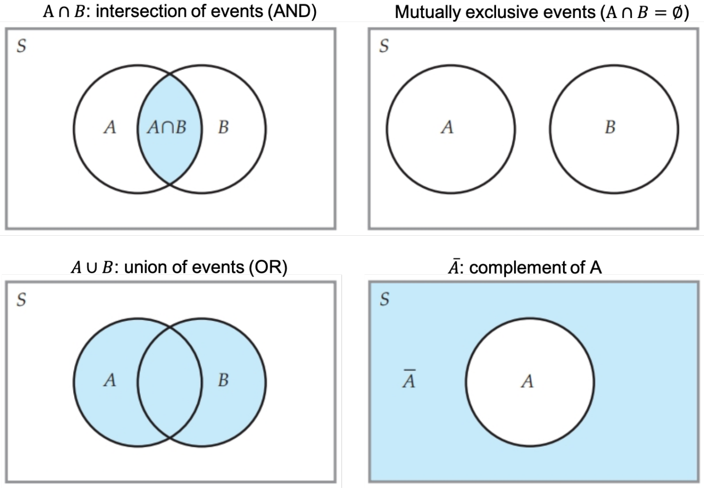
$S = \{HHH, HHT, HTT, TTT, TTH, THH, THT, HTH\}$

$A =$ two heads $= \{HHT, THH, HTH\}$
$P(A) = \frac{N_A}{N} = \frac{3}{8}$

$B =$ more than one head $= \{HHT, THH, HTH, HHH\}$
$P(B) = \frac{N_B}{N} = \frac{1}{2}$



Elementary events can be combined to define more complex events, that is, **compound events**. There are rules for combining the probabilities of elementary events to calculate the robability of compound events.



Collectively exhaustive: given K events $E_1, E_2, \ldots, E_K$ in the sample space S, if $E_1 \cup E_2 \cup \ldots \cup E_K = S$, the K events are said to be collectively exhaustive

Roulette: S= {0,1,. . . , 36}

A= first column= {1,4,7,. . . , 34}
B= red = {1,3,5,7,8,12,14,16,18,. . . }
C= {0}

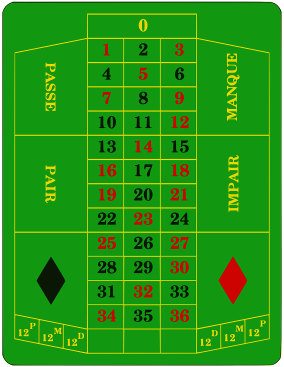
$$A \cap B = \text{first column AND red}$$
$$= \{1, 7, 16, 19, 25, 34\}$$

$$A \cup B = \text{first column OR red}$$
$$= \{1, 3, 4, 5, 7, 9, 10, 12, \ldots\}$$
$$= A + B - (A \cap B)$$

$$A \cap C = \text{first column AND } \{0\} = \varnothing$$

$$A \cup C = \text{first column OR } \{0\} = \{0, 1, 4, 7, \ldots, 34\}$$
$$= A + C$$

$$\bar{C} = \{1, 2, 3, \ldots, 36\} = S - C$$



The set of all basic outcomes contained in a sample space is mutually exclusive and collectively exhaustive

Postulates

- $0 \leq P(A) \leq 1$: any probability of an event is a number between 0 and 1
- $P(A) = \sum_A P(\text{basic outcome}_i) = P(O_1) + P(O_2) + \ldots$ basic outcomes are mutually exclusive
- $P(S) = 1$: all possible basic outcomes together must have probability 1, they are collectively exhaustive
- **Complement rule**: $P(\bar{A}) = 1 - P(A)$, note that A and $\bar{A}$ are mutually exclusive and collectively exhaustive
- **Addition rule of probability**: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Complement rule

$$P(\bar{A}) = 1 - P(A)$$

Note that A and $\bar{A}$ are collectively exhaustive

$$A \cup \bar{A} = S$$

and that A and $\bar{A}$ are mutually exclusive:

$$P(A \cup \bar{A}) = P(A) + P(\bar{A})$$

then

$$P(A) + P(\bar{A}) = P(A \cup \bar{A}) = P(S) = 1$$

Also follows that $P(\varnothing) = 0$

Addition rule of probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Note that $A \cup B = A \cup (\bar{A} \cap B)$, A and $\bar{A} \cap B$ are mutually exclusive, then

$$P(A \cup B) = P(A) + P(\bar{A} \cap B)$$

Moreover, we can write $B = (A \cap B) \cup (\bar{A} \cap B)$ with $(A \cap B)$ and $(\bar{A} \cap B)$ mutually exclusive

$$P(B) = P(A \cap B) + P(\bar{A} \cap B) \rightarrow P(\bar{A} \cap B) = P(B) - P(A \cap B)$$

Then,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Random experiment: rolling a dice

Sample space: $S = \{1, 2, 3, 4, 5, 6\}$

Event: $A =$ get a 6
Event: $B =$ get an even number $= \{2, 4, 6\}$

$P(A) = \frac{1}{6}$

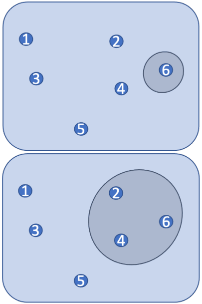
$P(B) = P(\text{get a } 2) + P(\text{get a } 4) + P(\text{get a } 6) = \frac{3}{6} = \frac{1}{2}$

$P(S) = P(\text{get a } 1) + P(\text{get a } 2) + \ldots + P(\text{get a } 6) = 1$

$P(\bar{A}) = 1 - \frac{1}{6} = \frac{5}{6}$

$P(A \cap B) = P(\text{get a } 6) = \frac{1}{6}$

$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{6} + \frac{1}{2} - \frac{1}{6} = \frac{1}{2}$



Note that A implies B ($A \subset B$), then $P(A) \leq P(B)$

Roulette: S= $\{0,1,\ldots,36\}$

A= first column= $\{1,4,7,\ldots,34\}$
B= red $=\{1,3,5,7,8,12,14,16,18,\ldots\}$
C= $\{0\}$

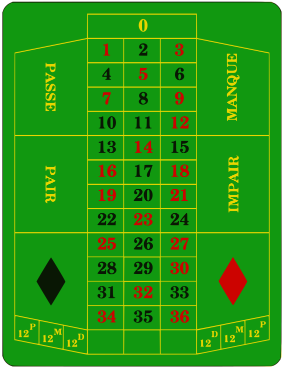
$P(A\cap B)=P(\text{first column AND red})=\frac{6}{37}$

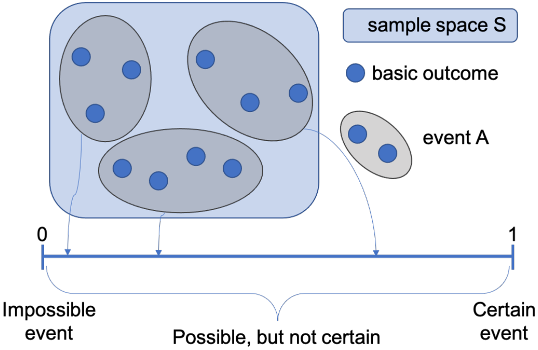
$P(A\cup B)=P(\text{first column OR red})=\frac{24}{37}$

$P(A\cap C)=P(\text{first column AND}\{0\})=1$

$P(A\cup C)=P(\text{first column OR}\{0\})=\frac{13}{37}$

$P(\bar{C})=P(S)-P(C)=1-\frac{1}{37}=\frac{36}{37}$





Probability model

A **probability model** is a mathematical description of a random experiment consisting of a sample space and a way of assigning probabilities to events

It is a convenient way to describe the distribution of an experiment's outcomes and involves listing all possible outcomes and their associated probabilities.

Examples

Canada has two official languages, English and French. Choose a Canadian randomly and ask, "What is your mother tongue?". Here is the distribution of responses, combining many separate languages from the broad Asian/Pacific region:

Language	English	French	Asian/Pacific	Other
Probability	0.59	0.23	0.07	?

a) Complete the table
b) What is the probability that a Canadian's mother tongue is not English?

$$S = \{English, French, Asian/Pacific, Other\}$$

Examples

Canada has two official languages, English and French. Choose a Canadian randomly and ask, "What is your mother tongue?". Here is the distribution of responses, combining many separate languages from the broad Asian/Pacific region:

Language	English	French	Asian/Pacific	Other
Probability	0.59	0.23	0.07	?

a) Complete the table
b) What is the probability that a Canadian's mother tongue is not English?

$$S = \{English, French, Asian/Pacific, Other\}$$

$$P(\overline{English}) = P(French) + P(Asian/Pacific) + P(Other)$$

equivalently

$$P(\overline{English}) = 1 - P(English) = .41$$

Examples

The sales manager wants to estimate the probability that a car will be returned for service during the warranty period. The following table shows a manager's probability assessment for the number of returns

Number of returns	0	1	2	3	4
Probability	0.28	0.36	0.23	0.09	0.04

Let A be the event "the number of returns will be more than two", and let B be the event "the number of returns will be less than four".

a) Find the probability of event A
b) Find the probability of event B
c) Find the probability of the complement of event A
d) Find the probability of the intersection of events A and B
e) Find the probability of the union of events A and B

Examples

Number of returns	0	1	2	3	4
Probability	0.28	0.36	0.23	0.09	0.04

Let A be the event "the number of returns will be more than two", and let B be the event "the number of returns will be less than four".

a) Find the probability of event A

Examples

Number of returns	0	1	2	3	4
Probability	0.28	0.36	0.23	0.09	0.04

Let A be the event "the number of returns will be more than two", and let B be the event "the number of returns will be less than four".

a) Find the probability of event A

$$\begin{aligned} P(A) &= P(\text{returns more than 2}) = P(returns > 2) \\ &= P(returns = 3) + P(returns = 4) \end{aligned}$$

b) Find the probability of event B

Examples

Number of returns	0	1	2	3	4
Probability	0.28	0.36	0.23	0.09	0.04

Let A be the event "the number of returns will be more than two", and let B be the event "the number of returns will be less than four".

a) Find the probability of event A

$$\begin{aligned} P(A) &= P(\text{returns more than 2}) = P(returns > 2) \\ &= P(returns = 3) + P(returns = 4) \end{aligned}$$

b) Find the probability of event B

$$\begin{aligned} P(B) &= P(\text{returns less than 4}) = P(returns < 4) \\ &= P(returns = 0) + P(returns = 1) + P(returns = 2) + P(returns = 3) \end{aligned}$$

or

Examples

Number of returns	0	1	2	3	4
Probability	0.28	0.36	0.23	0.09	0.04

Let A be the event "the number of returns will be more than two", and let B be the event "the number of returns will be less than four".

a) Find the probability of event A

$$\begin{aligned} P(A) &= P(\text{returns more than 2}) = P(returns > 2) \\ &= P(returns = 3) + P(returns = 4) \end{aligned}$$

b) Find the probability of event B

$$\begin{aligned} P(B) &= P(\text{returns less than 4}) = P(returns < 4) \\ &= P(returns = 0) + P(returns = 1) + P(returns = 2) + P(returns = 3) \end{aligned}$$

or

$$P(B) = P(\text{returns less than 4}) = 1 - P(returns = 4)$$

Examples

Number of returns	0	1	2	3	4
Probability	0.28	0.36	0.23	0.09	0.04

Let A be the event "the number of returns will be more than two", and let B be the event "the number of returns will be less than four".

c) Find the probability of the complement of event A

Examples

Number of returns	0	1	2	3	4
Probability	0.28	0.36	0.23	0.09	0.04

Let A be the event "the number of returns will be more than two", and let B be the event "the number of returns will be less than four".

c) Find the probability of the complement of event A

$$P(\bar{A}) = 1 - P(A) = 1 - P(\text{returns more than 2})$$

d) Find the probability of the intersection of events A and B

Examples

Number of returns	0	1	2	3	4
Probability	0.28	0.36	0.23	0.09	0.04

Let A be the event "the number of returns will be more than two", and let B be the event "the number of returns will be less than four".

c) Find the probability of the complement of event A

$$P(\bar{A}) = 1 - P(A) = 1 - P(\text{returns more than 2})$$

d) Find the probability of the intersection of events A and B

$$P(A \cap B) = P(\text{returns more than 2 AND returns less than 4}) = P(\text{returns} = 3)$$

d) Find the probability of the union of events A and B

Examples

Number of returns	0	1	2	3	4
Probability	0.28	0.36	0.23	0.09	0.04

Let A be the event "the number of returns will be more than two", and let B be the event "the number of returns will be less than four".

c) Find the probability of the complement of event A

$$P(\bar{A}) = 1 - P(A) = 1 - P(\text{returns more than 2})$$

d) Find the probability of the intersection of events A and B

$$P(A \cap B) = P(\text{returns more than 2 AND returns less than 4}) = P(\text{returns} = 3)$$

d) Find the probability of the union of events A and B

$$P(A \cup B) = P(\text{returns more than 2 OR returns less than 4}) = 1$$

Examples

A cell phone company found that 75% of customers want text messaging on their phone, 80% want photo capability, and 65% want both.
What is the probability that a customer will want at least one of these?

Define the events:

$A =$ costumer wants text messaging

$B =$ costumer wants photo capability

$A \cap B =$ customer wants text messaging AND photo capability

We know that $P(A) = .75$, $P(B) = .8$, and $P(A \cap B) = .65$

Examples

A cell phone company found that 75% of customers want text messaging on their phone, 80% want photo capability, and 65% want both.
What is the probability that a customer will want at least one of these?

Define the events:

$A =$ costumer wants text messaging

$B =$ costumer wants photo capability

$A \cap B =$ customer wants text messaging AND photo capability

We know that $P(A) = .75$, $P(B) = .8$, and $P(A \cap B) = .65$

We need $P(A \cup B) =$ probability that a customer wants text messaging OR photo capability

Examples

A cell phone company found that 75% of customers want text messaging on their phone, 80% want photo capability, and 65% want both.
What is the probability that a customer will want at least one of these?

Define the events:

$A =$ costumer wants text messaging

$B =$ costumer wants photo capability

$A \cap B =$ customer wants text messaging AND photo capability

We know that $P(A) = .75$, $P(B) = .8$, and $P(A \cap B) = .65$

We need $P(A \cup B) =$ probability that a customer wants text messaging OR photo capability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = .75 + .8 - .65$$

Motor vehicles sold in the US are classified as either cars or light trucks and as either domestic or imported
Last year, 80% of the new vehicles sold to individuals were domestic, 54% were light trucks, and 47% were domestic light trucks

Choose a vehicle sale at random and compute:

$$P(\text{domestic} \cup \text{light truck})$$

Motor vehicles sold in the US are classified as either cars or light trucks and as either domestic or imported

Last year, 80% of the new vehicles sold to individuals were domestic, 54% were light trucks, and 47% were domestic light trucks

Choose a vehicle sale at random and compute:

$$P(\text{domestic} \cup \text{light truck})$$

$P(\text{domestic}) = 0.80$
$P(\text{truck}) = 0.54$
$P(\text{domestic} \cap \text{light truck}) = 0.47$

Then,

$$\begin{aligned} P(\text{domestic} \cup \text{light truck}) &= P(\text{domestic}) + P(\text{light truck}) \\ &\quad - P(\text{domestic} \cap \text{light truck}) \\ &= 0.8 + 0.54 - 0.47 \end{aligned}$$

$P(\text{domestic}) = 0.80$
$P(\text{truck}) = 0.54$
$P(\text{domestic} \cap \text{truck}) = 0.47$

$P(\text{domestic} \cup \text{truck}) = 0.87$

	Domestic	Imported	
Light truck			
Car			

$P(\text{domestic}) = 0.80$
$P(\text{truck}) = 0.54$
$P(\text{domestic} \cap \text{truck}) = 0.47$

$P(\text{domestic} \cup \text{truck}) = 0.87$

	Domestic	Imported	
Light truck			
Car			

	Domestic	Imported	
Light truck	0.47		0.54
Car			
	0.80		

The total row/column can be obtained from the joint probabilities by the addition rule $P(\text{truck}) = P(\text{truck} \cap \text{domestic}) + P(\text{truck} \cap \text{imported})$

	Domestic	Imported	
Light truck	0.47	0.07	0.54
Car	0.33	0.13	0.46
	0.80	0.20	

The probability that a vehicle is a truck is

$$P(\text{truck}) = P(\text{truck} \cap \text{domestic}) + P(\text{truck} \cap \text{imported}) = 0.47 + 0.07 = 0.54$$

Does knowing that the chosen vehicle is imported, change the probability that it is a truck?

	Domestic	Imported	
Light truck	0.47	0.07	0.54
Car	0.33	0.13	0.46
	0.80	0.20	

The probability that a vehicle is a truck is

$P(\text{truck}) = P(\text{truck}\cap\text{domestic}) + P(\text{truck}\cap\text{imported}) = 0.47 + 0.07 = 0.54$

Does knowing that the chosen vehicle is imported, change the probability that it is a truck?

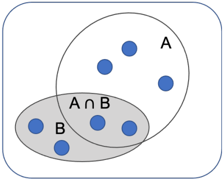
$$P(\text{truck}|\text{imported}) = \frac{P(\text{truck} \cap \text{imported})}{P(\text{imported})} = \frac{0.07}{0.20} = .35$$

The probability of an event can change if we know that some other event has occurred

Conditional probability

Let A and B be two events, the **conditional probability** of event A, given that event B has occurred is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \text{provided that } P(B) > 0$$

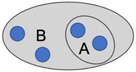


Similarly

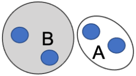
$$P(B|A) = \frac{P(B \cap A)}{P(A)}, \quad \text{provided that } P(A) > 0$$

Note that:

if $A \subset B$, then $P(B|A) = 1$



if A and B are mutually exclusive, then $P(A|B) = 0$



	Domestic	Imported	
Light truck	0.47	0.07	0.54
Car	0.33	0.13	0.46
	0.80	0.20	

What proportion of imports are trucks?

$$P(\text{truck}|\text{imported}) \quad \text{or} \quad P(\text{imported}|\text{truck})$$

What proportion of trucks are imported?

$$P(\text{truck}|\text{imported}) \quad \text{or} \quad P(\text{imported}|\text{truck})$$

Multiplication Rule

From the definition of conditional probability follows the **multiplication rule**

Given two events A and B, the probability of their intersection is

$$\begin{aligned} P(A \cap B) &= P(B)P(A|B) \\ &= P(A)P(B|A) \end{aligned}$$

Namely, for both events to occur, the first one must occur, i.e., $P(A)$, and then, given that the first occurred, the second must occur, i.e. $P(B|A)$.

Statistical Independence

Two events A and B are said to be **statistically independent** if and only if

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

From the multiplication rule, we can also derive

$$P(A|B) = P(A)$$

and

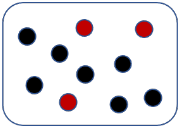
$$P(B|A) = P(B)$$

that is, the information about the occurrence of B is of no value in determining $P(A)$ (same for A and B)

A jar contains 3 red balls and 7 black balls

$A =$ draw a red ball
$P(A) = \frac{3}{10} = 0.3$

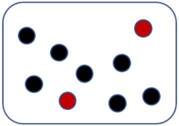
$B =$ the second ball drawn is red (replacing the first ball)

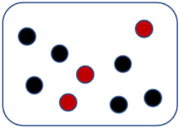


$$P(B|A) = P(B|\bar{A}) = \frac{3}{10}$$

A and B are independent

$B =$ the second ball drawn is red (not replacing the first ball)





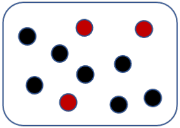
$$P(B|A) = \frac{2}{9} \text{ and } P(B|\bar{A}) = \frac{3}{9}$$

A and B are not independent

A jar contains 3 red balls and 7 black balls

$A =$ draw a red ball
$P(A) = \frac{3}{10} = 0.3$

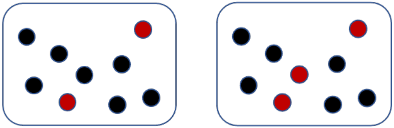
$B =$ the second ball drawn is red (replacing the first ball)



$$P(B|A) = P(B|\bar{A}) = \frac{3}{10}$$

A and B are independent

$B =$ the second ball drawn is red (not replacing the first ball)



$$P(B|A) = \frac{2}{9} \text{ and } P(B|\bar{A}) = \frac{3}{9}$$

A and B are not independent

We have to prove that $P(B|A) = P(B)$ in the first scenario and $P(B|A) \neq P(B)$ in the second one

Law of total probability

Given the events A and B with $0 < P(B) < 1$, then

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

- B and $\bar{B}$ are mutually exclusive and collectively exhaustive, $P(B \cup \bar{B}) = 1$
- $P(B)$ and $P(\bar{B})$ act as weights in considering the conditional probabilities

Law of total probability

Given the events A and B with $0 < P(B) < 1$, then

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

- B and $\bar{B}$ are mutually exclusive and collectively exhaustive, $P(B \cup \bar{B}) = 1$
- $P(B)$ and $P(\bar{B})$ act as weights in considering the conditional probabilities

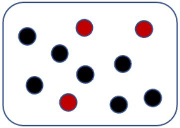
In general,
Given the events A and $B_1, B_2, \ldots$ with $B_i \cap B_j = \varnothing$, $i \neq j$ and collectively exhaustive then

$$P(A) = \sum_j P(A \cap Bj) = P(B_j)P(A|B_j)$$

A jar contains 3 red balls and 7 black balls

$A =$ draw a red ball
$P(A) = \frac{3}{10} = 0.3$

$B =$ the second ball drawn is red (replacing the first ball)

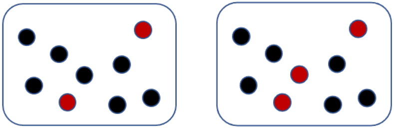


$$P(B|A) = P(B|\bar{A}) = \frac{3}{10}$$

A and B are independent

$$P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A})$$
$$= 0.3\frac{3}{10} + 0.7\frac{3}{10} = \frac{3}{10}$$

$B =$ the second ball drawn is red (not replacing the first ball)



$$P(B|A) = \frac{2}{9} \text{ and } P(B|\bar{A}) = \frac{3}{9}$$

A and B are not independent

$$P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A})$$
$$= 0.3\frac{2}{9} + 0.7\frac{3}{9} = \frac{3}{10}$$

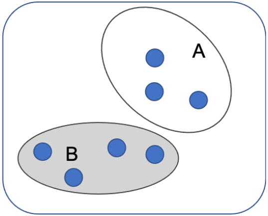
Multiplication rule

If A and B are independent then $P(A \cap B) = P(A)P(B)$, not otherwise

Addition rule

If A and B are mutually exclusive then $P(A \cup B) = P(A) + P(B)$, not otherwise

Mutually exclusive events $\neq$ independent events



$$P(A|B) = 0$$

A and B are not independent

Independence cannot be depicted in the Venn diagram because it involves the probabilities of the events rather than the outcomes

Complement rule

$$P(\bar{A}) = 1 - P(A)$$

Addition rule

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

if the events are mutually exclusive $P(A \cup B) = P(A) + P(B)$

Conditional probability

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

Multiplication Rule

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

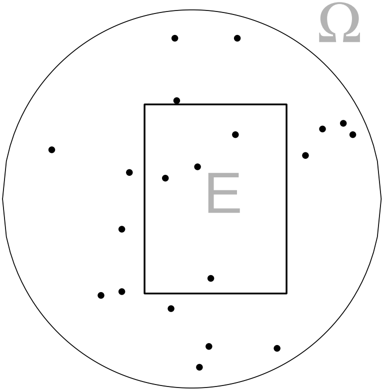
if the events are independent: $P(A|B) = P(A)$ and
$P(A \cap B) = P(B)P(A) = P(A)P(B)$

Law of total probability

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

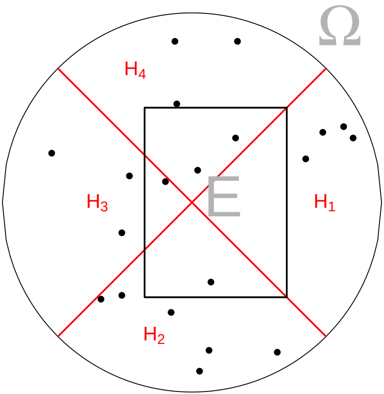
Law of total probability

$$P(E) \;=\; P\left(E \cap S\right)$$



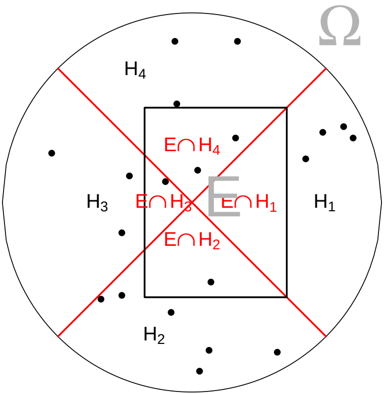
Law of total probability

$$
\begin{aligned}
P(E) &= P(E \cap S) \\
&= P\left(E \cap \bigcup_{i=1}^{k} H_i\right)
\end{aligned}
$$



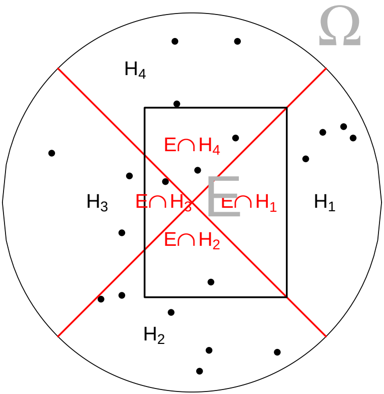
Law of total probability

$$
\begin{aligned}
P(E) &= P(E \cap S) \\
&= P\left(E \cap \bigcup_{i=1}^{k} H_i\right) \\
&= P\left(\bigcup_{i=1}^{k} (E \cap H_i)\right)
\end{aligned}
$$



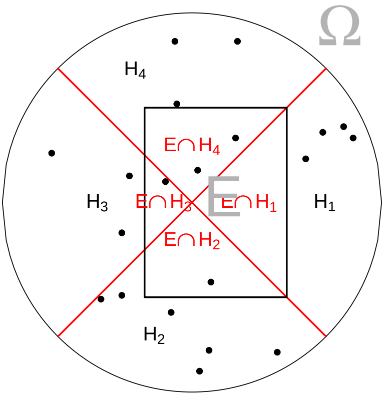
Law of total probability

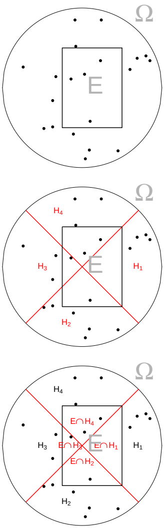
$$
\begin{aligned}
P(E) &= P(E \cap S) \\
&= P\left(E \cap \bigcup_{i=1}^{k} H_i\right) \\
&= P\left(\bigcup_{i=1}^{k} (E \cap H_i)\right) \\
&= \sum_{i=1}^{k} P(E \cap H_i)
\end{aligned}
$$



Law of total probability

$$
\begin{aligned}
P(E) &= P(E \cap S) \\
&= P\left(E \cap \bigcup_{i=1}^{k} H_i\right) \\
&= P\left(\bigcup_{i=1}^{k} (E \cap H_i)\right) \\
&= \sum_{i=1}^{k} P(E \cap H_i) \\
&= \sum_{i=1}^{k} P(H_i) P(E|H_i)
\end{aligned}
$$





Assuming that

- the basic outcomes (20) have the same probability to occur
- the four events $H_i, i = 1, \ldots, 4$ are mutually exclusive and collectively exhaustive

$$P(E) = \frac{4}{20} = \frac{1}{5}$$

Compute $P(E)$ using the law of total probability:

$$\begin{aligned}
P(E) &= \sum_i P(H_i)P(E|H_i) = \\
&= P(H_1)P(E|H_1) + P(H_2)P(E|H_2) + \\
&+ P(H_3)P(E|H_3) + P(H_4)P(E|H_4) = \\
&= \frac{1}{5}\frac{0}{4} + \frac{7}{20}\frac{1}{7} + \frac{1}{5}\frac{1}{4} + \frac{1}{4}\frac{2}{5} = \\
&= \frac{1}{5}
\end{aligned}$$

Smallpox virus in Boston (1721)

The data concern 6,224 people exposed to the smallpox virus in Boston in 1721 (OpenIntroStat).

Of these, 244 were exposed - on a voluntary basis - to the disease in a controlled manner by doctors (a kind of ante litteram vaccination).

It is then known who survived the epidemic and who did not.

		Exposed		
		Yes	No	Total
Survived	Yes	238	5136	5374
	No	6	844	850
	Total	244	5980	6224

Smallpox virus in Boston (1721)

The data concern 6,224 people exposed to the smallpox virus in Boston in 1721 (OpenIntroStat).

Of these, 244 were exposed - on a voluntary basis - to the disease in a controlled manner by doctors (a kind of ante litteram vaccination).

It is then known who survived the epidemic and who did not.

		Exposed		
		Yes	No	Total
Survived	Yes	238	5136	5374
	No	6	844	850
	Total	244	5980	6224

Let E be the event "Exposed", and let S be the event "Survived"

		Exposed		
		Yes (E)	No ($\bar{E}$)	Total
Survived	Yes (S)	238 ($S \cap E$)	5136 ($S \cap \bar{E}$)	5374
	No ($\bar{S}$)	6 ($\bar{S} \cap E$)	844 ($\bar{S} \cap \bar{E}$)	850
	Total	244	5980	6224

Let E be the event "Exposed", and let S be the event "Survived"

		Exposed		
		Yes (E)	No ($\bar{E}$)	Total
Survived	Yes (S)	238 ($S \cap E$)	5136 ($S \cap \bar{E}$)	5374
	No ($\bar{S}$)	6 ($\bar{S} \cap E$)	844 ($\bar{S} \cap \bar{E}$)	850
	Total	244	5980	6224

a) For a randomly selected individual, what is the probability of surviving the epidemic?

Let E be the event "Exposed", and let S be the event "Survived"

		Exposed		
		Yes (E)	No ($\bar{E}$)	Total
Survived	Yes (S)	238 ($S \cap E$)	5136 ($S \cap \bar{E}$)	5374
	No ($\bar{S}$)	6 ($\bar{S} \cap E$)	844 ($\bar{S} \cap \bar{E}$)	850
	Total	244	5980	6224

a) For a randomly selected individual, what is the probability of surviving the epidemic?

$$P(S) = \frac{\#S}{\#Total} = \frac{5374}{6224} = 0.86$$

b) For a randomly selected individual, what is the probability of being exposed?

Let E be the event "Exposed", and let S be the event "Survived"

		Exposed		
		Yes (E)	No ($\bar{E}$)	Total
Survived	Yes (S)	238 ($S \cap E$)	5136 ($S \cap \bar{E}$)	5374
	No ($\bar{S}$)	6 ($\bar{S} \cap E$)	844 ($\bar{S} \cap \bar{E}$)	850
	Total	244	5980	6224

a) For a randomly selected individual, what is the probability of surviving the epidemic?

$$P(S) = \frac{\#S}{\#Total} = \frac{5374}{6224} = 0.86$$

b) For a randomly selected individual, what is the probability of being exposed?

$$P(E) = \frac{\#E}{\#Total} = \frac{244}{6224} = 0.039$$

Let E be the event "Exposed", and let S be the event "Survived"

		Exposed		
		Yes (E)	No ($\bar{E}$)	Total
Survived	Yes (S)	238 ($S \cap E$)	5136 ($S \cap \bar{E}$)	5374
	No ($\bar{S}$)	6 ($\bar{S} \cap E$)	844 ($\bar{S} \cap \bar{E}$)	850
	Total	244	5980	6224

c) For a random individual, what is the probability that s/he was exposed and survived?

Let E be the event "Exposed", and let S be the event "Survived"

		Exposed		
		Yes (E)	No ($\bar{E}$)	Total
Survived	Yes (S)	238 ($S \cap E$)	5136 ($S \cap \bar{E}$)	5374
	No ($\bar{S}$)	6 ($\bar{S} \cap E$)	844 ($\bar{S} \cap \bar{E}$)	850
	Total	244	5980	6224

c) For a random individual, what is the probability that s/he was exposed and survived?

$$P(E \cap S) = \frac{\# E \cap S}{\# Total} = \frac{238}{6224} = 0.038$$

d) For a randomly selected individual who has been exposed, what is the probability of surviving?

Let E be the event "Exposed", and let S be the event "Survived"

		Exposed		
		Yes (E)	No ($\bar{E}$)	Total
Survived	Yes (S)	238 ($S \cap E$)	5136 ($S \cap \bar{E}$)	5374
	No ($\bar{S}$)	6 ($\bar{S} \cap E$)	844 ($\bar{S} \cap \bar{E}$)	850
	Total	244	5980	6224

c) For a random individual, what is the probability that s/he was exposed and survived?

$$P(E \cap S) = \frac{\#E \cap S}{\#Total} = \frac{238}{6224} = 0.038$$

d) For a randomly selected individual who has been exposed, what is the probability of surviving?

$$P(S|E) = \frac{\#E \cap S}{\#E} = \frac{238}{244} = 0.974$$

or

$$P(S|E) = \frac{P(E \cap S)}{P(E)} = \frac{0.038}{0.039} = 0.974$$

Let E be the event "Exposed", and let S be the event "Survived"

		Exposed		
		Yes (E)	No ($\bar{E}$)	Total
Survived	Yes (S)	238 ($S \cap E$)	5136 ($S \cap \bar{E}$)	5374
	No ($\bar{S}$)	6 ($\bar{S} \cap E$)	844 ($\bar{S} \cap \bar{E}$)	850
	Total	244	5980	6224

e) For a randomly selected individual who has not been exposed, what is the probability of surviving?

Let E be the event "Exposed", and let S be the event "Survived"

		Exposed		
		Yes (E)	No ($\bar{E}$)	Total
Survived	Yes (S)	238 ($S \cap E$)	5136 ($S \cap \bar{E}$)	5374
	No ($\bar{S}$)	6 ($\bar{S} \cap E$)	844 ($\bar{S} \cap \bar{E}$)	850
	Total	244	5980	6224

e) For a randomly selected individual who has not been exposed, what is the probability of surviving?

$$P(S|\bar{E}) = \frac{\#\bar{E} \cap S}{\#\bar{E}} = \frac{5136}{5980} = 0.859$$

or

$$P(S|\bar{E}) = \frac{P(\bar{E} \cap S)}{P(\bar{E})} = \frac{5136/6224}{5980/6224} = 0.859$$

or

$$P(S|\bar{E}) = 1 - P(\bar{S}|\bar{E})$$

f) Verify that you obtain the same result in point a) $P(S) =$ using the law of total probability

Let E be the event "Exposed", and let S be the event "Survived"

		Exposed		
		Yes (E)	No ($\bar{E}$)	Total
Survived	Yes (S)	238 ($S \cap E$)	5136 ($S \cap \bar{E}$)	5374
	No ($\bar{S}$)	6 ($\bar{S} \cap E$)	844 ($\bar{S} \cap \bar{E}$)	850
	Total	244	5980	6224

e) For a randomly selected individual who has not been exposed, what is the probability of surviving?

$$P(S|\bar{E}) = \frac{\#\bar{E} \cap S}{\#\bar{E}} = \frac{5136}{5980} = 0.859$$

or

$$P(S|\bar{E}) = \frac{P(\bar{E} \cap S)}{P(\bar{E})} = \frac{5136/6224}{5980/6224} = 0.859$$

or

$$P(S|\bar{E}) = 1 - P(\bar{S}|\bar{E})$$

f) Verify that you obtain the same result in point a) $P(S) =$ using the law of total probability

U.S. presidential elections 2020

The table below shows the results of the popular vote in the last U.S. presidential elections reported by the Federal Election Commission.

Candidates are listed individually if they received more than 0.1% of the popular vote.
Americans eligible to vote in 2020 were 239,661,805.

Presidential candidate	Party	Popular vote
Joseph Robinette Biden Jr.	Democratic	81,268,924
Donald John Trump	Republican	74,216,154
Jo Jorgensen	Libertarian	1,865,724
Howard Gresham Hawkins	Green	405,035
Other		627,566
Total		158,383,403

a) What is the probability that an eligible voter selected at random voted?

U.S. presidential elections 2020

The table below shows the results of the popular vote in the last U.S. presidential elections reported by the Federal Election Commission.

Candidates are listed individually if they received more than 0.1% of the popular vote.
Americans eligible to vote in 2020 were 239,661,805.

Presidential candidate	Party	Popular vote
Joseph Robinette Biden Jr.	Democratic	81,268,924
Donald John Trump	Republican	74,216,154
Jo Jorgensen	Libertarian	1,865,724
Howard Gresham Hawkins	Green	405,035
Other		627,566
Total		158,383,403

a) What is the probability that an eligible voter selected at random voted?

$$P(Vote) = \frac{158,383,403}{239,661,805} = 0.661$$

Americans eligible to vote in 2020 were 239,661,805.

Presidential candidate	Party	Popular vote
Joseph Robinette Biden Jr.	Democratic	81,268,924
Donald John Trump	Republican	74,216,154
Jo Jorgensen	Libertarian	1,865,724
Howard Gresham Hawkins	Green	405,035
Other		627,566
Total		158,383,403

a) What is the probability that an eligible voter selected at random actually voted?

$$P(Vote) = \frac{158,383,403}{239,661,805} = 0.661$$

b) What is the probability that an eligible voter selected at random voted for the Democratic party?

Americans eligible to vote in 2020 were 239,661,805.

Presidential candidate	Party	Popular vote
Joseph Robinette Biden Jr.	Democratic	81,268,924
Donald John Trump	Republican	74,216,154
Jo Jorgensen	Libertarian	1,865,724
Howard Gresham Hawkins	Green	405,035
Other		627,566
Total		158,383,403

a) What is the probability that an eligible voter selected at random actually voted?

$$P(Vote) = \frac{158,383,403}{239,661,805} = 0.661$$

b) What is the probability that an eligible voter selected at random voted for the Democratic party?

$$P(Democratic) = \frac{81,268,924}{239,661,805} = 0.339$$

Americans eligible to vote in 2020 were 239,661,805.

Presidential candidate	Party	Popular vote
Joseph Robinette Biden Jr.	Democratic	81,268,924
Donald John Trump	Republican	74,216,154
Jo Jorgensen	Libertarian	1,865,724
Howard Gresham Hawkins	Green	405,035
Other		627,566
Total		158,383,403

a) What is the probability that an eligible voter selected at random actually voted?

$$P(Vote) = \frac{158,383,403}{239,661,805} = 0.661$$

b) What is the probability that an eligible voter selected at random voted for the Democratic party?

$$P(Democratic) = \frac{81,268,924}{239,661,805} = 0.339$$

c) What is the probability that an eligible voter selected at random, went to vote and voted for the Democratic party?

Americans eligible to vote in 2020 were 239,661,805.

Presidential candidate	Party	Popular vote
Joseph Robinette Biden Jr.	Democratic	81,268,924
Donald John Trump	Republican	74,216,154
Jo Jorgensen	Libertarian	1,865,724
Howard Gresham Hawkins	Green	405,035
Other		627,566
Total		158,383,403

a) What is the probability that an eligible voter selected at random actually voted?

$$P(Vote) = \frac{158,383,403}{239,661,805} = 0.661$$

b) What is the probability that an eligible voter selected at random voted for the Democratic party?

$$P(Democratic) = \frac{81,268,924}{239,661,805} = 0.339$$

c) What is the probability that an eligible voter selected at random, went to vote and voted for the Democratic party?

$$P(Vote \cap Democratic) = P(Democratic) = 0.339$$

Presidential candidate	Party	Popular vote
Joseph Robinette Biden Jr.	Democratic	81,268,924
Donald John Trump	Republican	74,216,154
Jo Jorgensen	Libertarian	1,865,724
Howard Gresham Hawkins	Green	405,035
Other		627,566
Total		158,383,403

a) $P(Vote) = \frac{158{,}383{,}403}{239{,}661{,}805} = 0.661$

b) $P(Democratic) = \frac{81{,}268{,}924}{239{,}661{,}805} = 0.339$

c) $P(Vote \cap Democratic) = P(Democratic) = 0.339$

d) Knowing that a voter selected randomly went to vote, what is the probability that s/he voted for the Democratic party?

Presidential candidate	Party	Popular vote
Joseph Robinette Biden Jr.	Democratic	81,268,924
Donald John Trump	Republican	74,216,154
Jo Jorgensen	Libertarian	1,865,724
Howard Gresham Hawkins	Green	405,035
Other		627,566
Total		158,383,403

a) $P(Vote) = \frac{158{,}383{,}403}{239{,}661{,}805} = 0.661$

b) $P(Democratic) = \frac{81{,}268{,}924}{239{,}661{,}805} = 0.339$

c) $P(Vote \cap Democratic) = P(Democratic) = 0.339$

d) Knowing that a voter selected randomly went to vote, what is the probability that s/he voted for the Democratic party?

$$P(Democratic|Vote) = \frac{P(Vote \cap Democratic)}{P(Vote)} = \frac{P(Democratic)}{P(Vote)} = \frac{0.339}{0.661} = 0.51$$

Bayes' island

Due to a storm, Bayes' Island has become inaccessible from the mainland.

The 100,001 residents on the island cannot leave, and no one can enter.

In the morning, a man's body was discovered on the beach by a pensioner walking his dog.

After conducting an investigation, the police determined that the man was killed during the night with a kitchen knife by a single individual whose DNA is known

Bayes' island

Due to a storm, Bayes' Island has become inaccessible from the mainland.

The 100,001 residents on the island cannot leave, and no one can enter.

In the morning, a man's body was discovered on the beach by a pensioner walking his dog.

After conducting an investigation, the police determined that the man was killed during the night with a kitchen knife by a single individual whose DNA is known

The inspector decides to proceed with the comparing the DNA found at the crime scene with that of the 100,000 residents

It is assumed that only one DNA is compatible (no twins)

A Positive test result indicates that the DNA is compatible with that of the murderer.
However...

However...

The test is not foolproof!

In one out of 10,000 cases, incompatible DNA can yield a positive result:
$\rightarrow$ False Positive: The test result is positive, but the subject is innocent.

In one out of 100,000 cases, a matched DNA can yield a negative result:
$\rightarrow$ False Negative: The test result is negative, but the subject is guilty.

However...

The test is not foolproof!

In one out of 10,000 cases, incompatible DNA can yield a positive result:
$\rightarrow$ False Positive: The test result is positive, but the subject is innocent.

In one out of 100,000 cases, a matched DNA can yield a negative result:
$\rightarrow$ False Negative: The test result is negative, but the subject is guilty.

The first test gives a positive result, indicating the subject's DNA is compatible.

The subject with compatible DNA is then arrested

No further tests are carried out

At the trial, the DNA evidence is the only piece of evidence brought by the prosecution, which argues:

"The probability of a False Positive ($FP = Positive|Innocent$) is so low ($1/10{,}000$) that the defendant must be convicted!"

What would happen if we ran all the 100,000 tests?

The 99,999 innocents all have incompatible DNA, but due to the false positive rate, on average 0.1% will test positive, which is 10 individuals

	G	I	
P	1	10	11
N	0	99,989	99,989
	1	99,999	100,000

The criminal will test positive with a high probability

We would find 11 positives, of which only one is guilty.
What is the probability that an individual with a positive DNA match is guilty?

Let G and T be the events:

- G : The individual is guilty (i.e., the DNA is compatible);
- T : The test is positive.

The question is: "Knowing that the test is positive, what is the probability that the subject is guilty?"

$$P(G|T)$$

Let G and T be the events:

- G : The individual is guilty (i.e., the DNA is compatible);
- T : The test is positive.

The question is: "Knowing that the test is positive, what is the probability that the subject is guilty?"

$$P(G|T)$$

We know that

$$P(T \cap G) = P(T)P(G|T) = P(G)P(T|G)$$

So, we have

$$P(G|T) = \frac{P(G)P(T|G)}{P(T)}$$

$$P(G|T) = \frac{P(G)P(T|G)}{P(T)}$$

No instigator has been found, so all the subjects are potentially guilty

$$P(G) = \frac{1}{100,000} = 0.00001$$

Applying the law of total probability:

$$P(T) = P(G)P(T|G) + P(\bar{G})P(T|\bar{G})$$

where:

$P(T|G) = 1 - \frac{1}{100{,}000} = \frac{99{,}999}{100{,}000}$ because of the FN probability

$P(T|\bar{G}) = \frac{1}{10{,}000}$ because of the FP probability

Thus,

$$P(T) = \frac{1}{100,000}\frac{99,999}{100,000} + \frac{99,999}{100,000}\frac{1}{10,000} \approx \frac{11}{100,000}$$

$$P(G|T) = \frac{P(G)P(T|G)}{P(T)}$$

No instigator has been found, thus all the subjects potentially are guilty

$$P(G) = \frac{1}{100,000} = 0.00001$$

Applying the law of total probability:

$$P(T) = P(G)P(T|G) + P(\bar{G})P(T|\bar{G})$$

where:

$P(T|G) = 1 - \frac{1}{100{,}000} = \frac{99{,}999}{100{,}000}$ because of the FN probability

$P(T|\bar{G}) = \frac{1}{10{,}000}$ because of the FP probability

$P(T) = \frac{1}{100{,}000} \times \frac{99{,}999}{100{,}000} + \frac{99{,}999}{100{,}000} \times \frac{1}{10{,}000} \approx \frac{11}{100{,}000}$

$$P(G|T) = \frac{P(G)P(T|G)}{P(T)} = \frac{\frac{1}{100{,}000}\frac{99{,}999}{100{,}000}}{\frac{11}{100{,}000}} \approx \frac{1}{11}$$

Bayes' Theorem

Given two events E e H, such that $P(E) \neq 0$ e $P(H) \neq 0$,

$$P(H|E) = \frac{P(H)P(E|H)}{P(E)}$$

or for a generic set of mutually exclusive and collectively exhaustive events $H_i, i = 1, \ldots, k$, and E, with $P(E) \neq 0$ and $P(H_i) \neq 0$

$$P(H_i|E) = \frac{P(H_i)P(E|H_i)}{P(E)} = \frac{P(H_i)P(E|H_i)}{\sum_{i=1}^{k} P(H_i)P(E|H_i)}$$

Twins

About 30% of twins are homozygous (identical twins), and the rest are heterozygous (fraternal twins).

Identical twins are necessarily of the same sex, in 50% of cases males, in 50% females. On the other hand, the pairs of fraternal twins are 25% both males, 25% both females, and the remaining 50% mixed.

Knowing that two twins are two girls, what is the probability that these are identical twins?

Twins

About 30% of twins are homozygous (identical twins), and the rest are heterozygous (fraternal twins).

Identical twins are necessarily of the same sex, in 50% of cases males, in 50% females. On the other hand, the pairs of fraternal twins are 25% both males, 25% both females, and the remaining 50% mixed.

Knowing that two twins are two girls, what is the probability that these are identical twins?

We define two events: $F =$ two females and $I =$ identical twins

We have to find

$$P(I|F) = \frac{P(I)P(F|I)}{P(F)}$$

We define two events: $F =$ two females and $I =$ identical twins

We have to find

$$P(I|F) = \frac{P(I)P(F|I)}{P(F)}$$

We know:

$P(I) = 0.3$

$P(F|I) = 0.5$

And by the law of total probability:

$$P(F) = P(F \cap I) + P(F \cap \bar{I}) = P(F|I)P(I) + P(F|\bar{I})P(\bar{I})$$

$P(F|\bar{I}) = 0.25$

$P(\bar{I}) = 1 - P(I) = 1 - 0.3 = 0.7$

Thus,

$$P(F) = P(F|I)P(I) + P(F|\bar{I})P(\bar{I}) = 0.5 \cdot 0.3 + 0.25 \cdot 0.7 = 0.325$$

We define two events: F = two females and I = identical twins

We have to find

$$P(I|F) = \frac{P(I)P(F|I)}{P(F)}$$

We know:

$P(I) = 0.3$

$P(F|I) = 0.5$

$P(F) = 0.325$

$$P(I|F) = \frac{P(I)P(F|I)}{P(F)} = \frac{0.3 \cdot 0.5}{0.325} = 0.462$$

Automobile sales incentive

A car dealership knows from past experience that 10% of the people who come into the showroom and talk to a salesperson will eventually purchase a car.

To increase the chances of success, you propose to offer a free dinner with a salesperson for all people who agree to listen to a complete sales presentation.

You know that some people will do anything for a free dinner, even if they do not intend to purchase a car. However, some people would rather not spend a dinner with a car salesperson. Thus, you wish to test the effectiveness of this sales promotion incentive.

The project is conducted for 6 months, and 40% of the people who purchased cars had a free dinner. In addition, 10% of the people who did not purchase cars had a free dinner.

Automobile sales incentive

A car dealership knows from past experience that 10% of the people who come into the showroom and talk to a salesperson will eventually purchase a car.

To increase the chances of success, you propose to offer a free dinner with a salesperson for all people who agree to listen to a complete sales presentation.

You know that some people will do anything for a free dinner, even if they do not intend to purchase a car. However, some people would rather not spend a dinner with a car salesperson. Thus, you wish to test the effectiveness of this sales promotion incentive.

The project is conducted for 6 months, and 40% of the people who purchased cars had a free dinner. In addition, 10% of the people who did not purchase cars had a free dinner.

a. Do people who accept the dinner have a higher probability of purchasing a new car?

A car dealership knows from past experience that 10% of the people who come into the showroom and talk to a salesperson will eventually purchase a car.

You wish to test the effectiveness of the sales promotion incentive: free dinner with a salesperson.

The project is conducted for 6 months, and 40% of the people who purchased cars had a free dinner. In addition, 10% of the people who did not purchase cars had a free dinner.

a. Do people who accept the dinner have a higher probability of purchasing a new car?

We define the events:

$D =$ dinner with the salesperson and $C =$ purchase a new car

A car dealership knows from past experience that 10% of the people who come into the showroom and talk to a salesperson will eventually purchase a car.

You wish to test the effectiveness of the sales promotion incentive: free dinner with a salesperson.

The project is conducted for 6 months, and 40% of the people who purchased cars had a free dinner. In addition, 10% of the people who did not purchase cars had a free dinner.

a. Do people who accept the dinner have a higher probability of purchasing a new car?

We define the events:

$D =$ dinner with the salesperson and $C =$ purchase a new car

We have to find:

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

We know:

A car dealership knows from past experience that 10% of the people who come into the showroom and talk to a salesperson will eventually purchase a car.

You wish to test the effectiveness of the sales promotion incentive: free dinner with a salesperson.

The project is conducted for 6 months, and 40% of the people who purchased cars had a free dinner. In addition, 10% of the people who did not purchase cars had a free dinner.

a. Do people who accept the dinner have a higher probability of purchasing a new car?

We define the events:

$D =$ dinner with the salesperson and $C =$ purchase a new car

We have to find:

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

We know:

$$P(C) = 0.1, \qquad \text{and} \qquad P(D|C) = 0.4, \qquad \text{and} \qquad P(D|\bar{C}) = 0.1$$

a. Do people who accept the dinner have a higher probability of purchasing a new car?

We define the events:

$D =$ dinner with the salesperson and $C =$ purchase a new car

We have to find:

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

We know:

$$P(C) = 0.1, \qquad \text{and} \qquad P(D|C) = 0.4, \qquad \text{and} \qquad P(D|\bar{C}) = 0.1$$

By the law of total probability

$$\begin{aligned} P(D) = P(D \cap C) + P(D \cap \bar{C}) &= P(D|C)P(C) + P(D|\bar{C})P(\bar{C}) \\ &= 0.4 \cdot 0.1 + 0.1 \cdot 0.9 = 0.13 \end{aligned}$$

a. Do people who accept the dinner have a higher probability of purchasing a new car?

We define the events:

$D =$ dinner with the salesperson and $C =$ purchase a new car

We have to find:

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

We know:

$$P(C) = 0.1, \qquad \text{and} \qquad P(D|C) = 0.4, \qquad \text{and} \qquad P(D|\bar{C}) = 0.1$$

By the law of total probability

$$\begin{aligned} P(D) = P(D \cap C) + P(D \cap \bar{C}) &= P(D|C)P(C) + P(D|\bar{C})P(\bar{C}) \\ &= 0.4 \cdot 0.1 + 0.1 \cdot 0.9 = 0.13 \end{aligned}$$

Therefore,

$$P(C|D) = \frac{0.1 \cdot 0.4}{0.13} = 0.308 > P(C) = 0.1$$

b. What is the probability that a person who does not accept a free dinner will purchase a car?

We define the events:

$D =$ dinner with the salesperson and $C =$ purchase a new car

We have to find:

$$P(C|\bar{D}) = \frac{P(C)P(\bar{D}|C)}{P(\bar{D})}$$

We know:

$P(C) = 0.1, \qquad P(D|C) = 0.4, \qquad P(D|\bar{C}) = 0.1, \qquad \text{and} \qquad P(D) = 0$

By the complement rule

$P(\bar{D}) = 1 - P(D) = 0.87 \qquad$ and also $\qquad P(\bar{D}|C) = 1 - P(D|C) = 0.6$

b. What is the probability that a person who does not accept a free dinner will purchase a car?

We define the events:

$D =$ dinner with the salesperson and $C =$ purchase a new car

We have to find:

$$P(C|\bar{D}) = \frac{P(C)P(\bar{D}|C)}{P(\bar{D})}$$

We know:

$P(C) = 0.1, \qquad P(D|C) = 0.4, \qquad P(D|\bar{C}) = 0.1, \qquad \text{and} \qquad P(D) = 0$

By the complement rule

$P(\bar{D}) = 1 - P(D) = 0.87 \qquad$ and also $\qquad P(\bar{D}|C) = 1 - P(D|C) = 0.6$

Therefore,

$$P(C|\bar{D}) = \frac{0.1 \cdot 0.6}{0.87} = 0.069 < P(C) = 0.1$$

Who uses the Internet?

Internet usage varies among millennials (adults born between 1981 and 1986), gen X-ers (adults born between 1965 and 1980), boomers (adults born between 1946 and 1964), and the silent generation (adults born before 1946).

Among adults born between 1946 and 1986, millennials comprise 30%, gen X-ers 27%, boomers 32%, and silent generation 11%.

Also, 100% of millennials uses the Internet, 91% of gen X-ers use the Internet, 85% of boomers use the Internet, and 78% of the silent generation uses the Internet.

What percentage of Internet users from these four groups are millennials? We define the events: $I =$ use Internet and $M =$ millennial

We have to find

$$P(M|I) = \frac{P(M)P(I|M)}{P(I)}$$

Who uses the Internet?

Internet usage varies among millennials (adults born between 1981 and 1986), gen X-ers (adults born between 1965 and 1980), boomers (adults born between 1946 and 1964), and the silent generation (adults born before 1946).

Among adults born between 1946 and 1986, millennials comprise 30%, gen X-ers 27%, boomers 32%, and silent generation 11%.

Also, 100% of millennials uses the Internet, 91% of gen X-ers use the Internet, 85% of boomers use the Internet, and 78% of the silent generation uses the Internet.

What percentage of Internet users from these four groups are millennials? We define the events: $I =$ use Internet and $M =$ millennial

We have to find

$$P(M|I) = \frac{P(M)P(I|M)}{P(I)}$$

What percentage of Internet users from these four groups are millennials?

We define the events: I = use Internet and M = millennial

We have to find

$$P(M|I) = \frac{P(M)P(I|M)}{P(I)}$$

We know:

$$\begin{array}{cccc} P(M) = 0.3 & P(X) = 0.27 & P(B) = 0.32 & P(S) = 0.11 \\ P(I|M) = 1 & P(I|X) = 0.91 & P(I|B) = 0.85 & P(I|S) = 0.78 \end{array}$$

By the law of total probability

$$\begin{aligned} P(I) &= P(I \cap M) + P(I \cap X) + P(I \cap B) + P(I \cap S) \\ &= P(I|M)P(M) + P(I|X)P(X) + P(I|B)P(B) + P(I|S)P(S) \\ &= 1 \cdot 0.3 + 0.27 \cdot 0.91 + 0.32 \cdot 0.85 + 0.11 \cdot 0.78 = 0.9035 \end{aligned}$$

What percentage of Internet users from these four groups are millennials?

We define the events: I = use Internet and M = millennial

We have to find

$$P(M|I) = \frac{P(M)P(I|M)}{P(I)}$$

We know:

$$\begin{array}{cccc} P(M) = 0.3 & P(X) = 0.27 & P(B) = 0.32 & P(S) = 0.11 \\ P(I|M) = 1 & P(I|X) = 0.91 & P(I|B) = 0.85 & P(I|S) = 0.78 \end{array}$$

By the law of total probability

$$\begin{aligned} P(I) &= P(I \cap M) + P(I \cap X) + P(I \cap B) + P(I \cap S) \\ &= P(I|M)P(M) + P(I|X)P(X) + P(I|B)P(B) + P(I|S)P(S) \\ &= 1 \cdot 0.3 + 0.27 \cdot 0.91 + 0.32 \cdot 0.85 + 0.11 \cdot 0.78 = 0.9035 \end{aligned}$$

Therefore,

$$P(M|I) = \frac{0.3 \cdot 1}{0.9035} = 0.332$$

that is, about 33% of Internet users are millennials