Elementary Statistics Course

Alfredo Sánchez Alberca (asalber@ceu.es)

Feb 2016

Department of Applied Math and Statistics
CEU San Pablo

License terms

Contents

Probability

Descriptive Statistics provide methods to describe de variables measured in the sample and their relations, but it doesn't allow to draw any conclusion about the population.

Now it's time to make the leap from the sample to the population and the bridge for that is the **probability theory**.

Remember that the sample has a limited information about the population, and in order to draw valid conclusions for the population the sample must be representative of it. For that reason, to guarantee the representativeness of the sample, this must be drawn randomly. This means that the choice of individuals in the sample is by chance.

The probability theory will give us the tools to control the random in the sampling and to determine the level of reliability of the conclusions drawn from the sample.

The study of a characteristic of the population is conducted through random experiments.

Definition (Random experiment)

A *random experiment* is an experiment that meets two conditions:

1. It is known the set of possible outcomes.
2. It is impossible to predict the outcome with absolute certainty.

Example. Gambling are typical examples of random experiments. The roll of a dice, for example, is a random experiment cause

- It is known the the set of possible outcomes: $\{1, 2, 3, 4, 5, 6\}$.
- Before rolling the dice, it's impossible to predict with absolute certainty the face of the dice that will occur.

Another non-gambling example is the random choice of an individual of a human population and the determination of its blood type. Generally, the draw of a sample by a random method is an random experiment.

Definition (Sample space)

The set Ω of the possible outcomes of a random experiment is known as *sample space*.

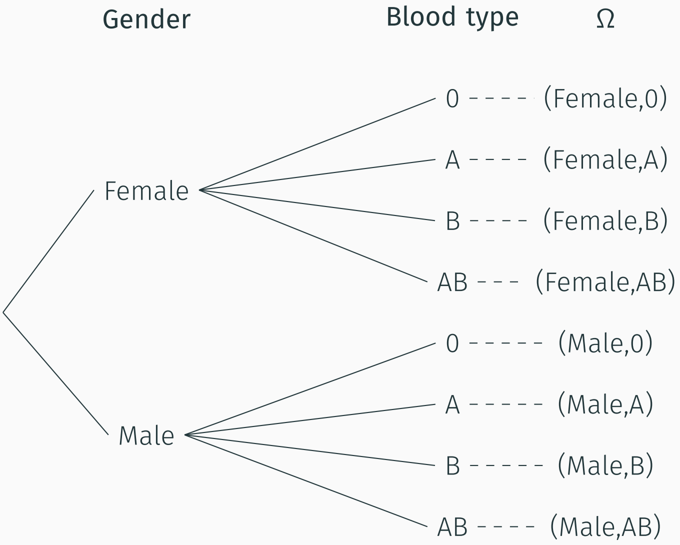
Example Some examples of sample spaces are:

- For the toss of a coin $\Omega = \{heads, tails\}$.
- For the roll of a dice $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- For the blood type of an individual drawn by chance $\Omega = \{\text{A}, \text{B}, \text{AB}, 0\}$.
- For the height of an individual drawn by chance $\Omega = \mathbb{R}^+$.

Sample space construction

In experiments where more than one variable is measured, the determination of the sample space can be difficult. In such a cases, it is advisable to use a tree diagram to construct the sample space.

In a tree diagram every variable is represented in a level of the tree and every possible outcome of the variable as a branch.



Gender
Blood type
Ω
Female
Male
0 - - - - (Female,0)
A - - - - (Female,A)
B - - - - (Female,B)
AB - - - (Female,AB)
0 - - - - - (Male,0)
A - - - - - (Male,A)
B - - - - - (Male,B)
AB - - - - (Male,AB)

Definition (Random event)

A *random event* is any subset of the sample space Ω of a random experiment.

There are different types of events:

- **Impossible event:** Is the event with no elements $\emptyset$. It has no chance of occurring.
- **Elemental events:** Are events with only one element, that is, a singleton.
- **Composed events:** Are events with two or more elements.
- **Sure event:** Is the event that contains the whole sample space. It always happens.

Definition (Event space)

Given a sample space Ω of a random experiment, the *event space* of Ω is the set of all possible events of Ω, and is noted $\mathcal{P}(\Omega)$.

Example. Given the sample space $\Omega = \{a, b, c\}$, its even space is

$$\mathcal{P}(\Omega) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$$

Event operations

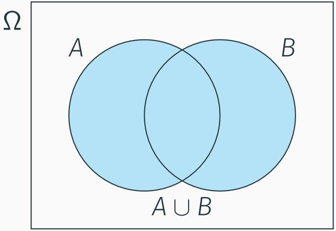
As events are subsets of the sample space, using the set theory we have the following operations on events:

- Union
- Intersection
- Complement
- Difference

Definition (Union event)

Given two events $A, B \in \mathcal{P}(\Omega)$, the *union* of A and B, denoted by $A \cup B$, is the event of all elements that are members of A or B or both.

$$A \cup B = \{x \,|\, x \in A \text{ or } x \in B\}.$$

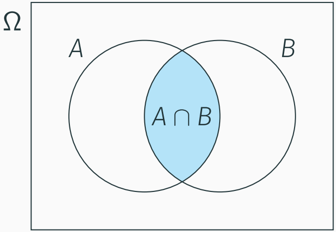


The union event $A \cup B$ happens when A or B happen.

Definition (Intersection event)

Given two events $A, B \in \mathcal{P}(\Omega)$, the *intersection* of A and B, denoted by $A \cap B$, is the event of all elements that are members of both A and B.

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}.$$



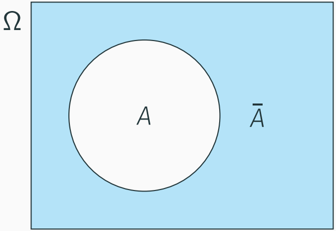
The intersection event $A \cap B$ happens when A and B happen.

Two events are **incompatible** if their intersection is empty.

Definition (Complementary event)

Given an event $A \in \mathcal{P}(\Omega)$, the *complementary or contrary event* of A, denoted by $\bar{A}$, is the event of all elements of Ω except the elements that are members of A.

$$\overline{A} = \{x \,|\, x \notin A\}.$$

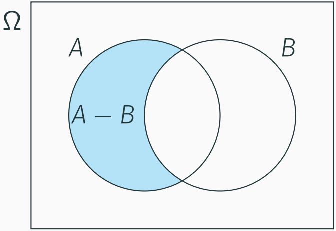


The complementary event $\bar{A}$ happens when A does not happen.

Definition (Difference event)

Given two events $A, B \in \mathcal{P}(\Omega)$, the *difference* of A and B, denoted by $A - B$, is the event of all elements that are members of A but not are members of B.

$$A - B = \{x \,|\, x \in A \text{ and } x \notin B\} = A \cap \bar{B}.$$



The difference event $A - B$ happens when A happens but B not.

Given the sample space of rolling a dice $\Omega = \{1, 2, 3, 4, 5, 6\}$ and the events $A = \{2, 4, 6\}$ and $B = \{1, 2, 3, 4\}$,

- The union of A and B is $A \cup B = \{1, 2, 3, 4, 6\}$.
- The intersection of A and B is $A \cap B = \{2, 4\}$.
- The complement of A is $\bar{A} = \{1, 3, 5\}$.
- The events A and $\bar{A}$ are incompatible.
- The difference of A and B is $A - B = \{6\}$, and the difference of B and A is $B - A = \{1, 3\}$.

Given the events $A, B, C \in \mathcal{P}(\Omega)$, the following properties are meet.

1. $A \cup A = A$, $A \cap A = A$ (idempotency).
2. $A \cup B = B \cup A$, $A \cap B = B \cap A$ (commutative).
3. $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$ (associative).
4. $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$, $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ (distributive).
5. $A \cup \emptyset = A$, $A \cap \Omega = A$ (neutral element).
6. $A \cup \Omega = \Omega$, $A \cap \emptyset = \emptyset$ (absorbing element).
7. $A \cup \overline{A} = \Omega$, $A \cap \overline{A} = \emptyset$ (complementary symmetric element).
8. $\overline{\overline{A}} = A$ (double contrary).
9. $\overline{A \cup B} = \overline{A} \cap \overline{B}$, $\overline{A \cap B} = \overline{A} \cup \overline{B}$ (Morgan's laws).

Definition (Probability — Laplace)

Given a sample space Ω of a random experiment where all elements of Ω are equally likely, the *probability* of an event $A \subseteq \Omega$ is the quotient between the number of elements of A and the number of elements of Ω

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{number of favorable outcomes}}{\text{number of possible outcomes}}$$

This definition is widespread, but it has important restrictions:

- It is required that all the elements of the sample space are equally likely (*equiprobability*).
- It can't be used with infinite sample spaces.

Watch out! These conditions are not meet in many real experiments.

Given the sample space of rolling a dice $\Omega = \{1, 2, 3, 4, 5, 6\}$ and the event $A = \{2, 4, 6\}$, the probability of A is

$$P(A) = \frac{|A|}{|\Omega|} = \frac{3}{6} = 0.5.$$

However, given the sample space of the blood type of a random individual $\Omega\{O, A, B, AB\}$, it's not possible to use the classical definition to compute the probability of having group A,

$$P(A) \neq \frac{|A|}{|\Omega|} = \frac{1}{4} = 0.25,$$

cause the blood types are not equally likely in human populations.

Theorem (Law of large numbers)

When a random experiment is repeated a large number of times, the relative frequency of an event tends to a number that is the real probability of the event.

The following definition of probability uses this theorem.

Definition (Frequency probability)

Given a sample space Ω of a replicable random experiment, the *probability* of an event $A \subseteq \Omega$ is the relative frequency of the event A in an infinite number of repetitions of the experiment

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

Although frequency probability avoid the restrictions of classical definition, it also have some drawbacks

- It computes an estimation of the real probability (more accurate the

Given the sample space of tossing a coin $\Omega = \{H, T\}$, if after tossing the coin 100 times we got 54 heads, then the probability of H is

$$P(H) = \frac{n_H}{n} = \frac{54}{100} = 0.54.$$

Given the sample space of the blood type of a random individual $\Omega\{O, A, B, AB\}$, if after drawing a random sample of 1000 persons we got 412 with blood type A, then the probability of A is

$$P(A) \neq \frac{n_A}{n} = \frac{412}{1000} = 0.412.$$

Axiomatic definition of probability

Definition (Probability — Kolmogórov)

Given a sample space Ω of a random experiment, a *probability* function is a function that maps every event $A \subseteq \Omega$ a real number $P(A)$, known as the probability of A, that meets the following axioms:

1. The probability of any event is nonnegative,
$$P(A) \geq 0.$$
2. The probability of the sure event is 1,
$$P(\Omega) = 1$$
3. The probability of the union of two incompatible events ($A \cap B = \emptyset$) is the sum of their probabilities
$$P(A \cup B) = P(A) + P(B).$$

From the previous axioms is possible to deduce some important properties of a probability function.

Given a sample space Ω of a random experiment and the events $A, B \subseteq \Omega$, the following properties are meet:

From the previous axioms is possible to deduce some important properties of a probability function.

Given a sample space Ω of a random experiment and the events $A, B \subseteq \Omega$, the following properties are meet:

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

From the previous axioms is possible to deduce some important properties of a probability function.

Given a sample space Ω of a random experiment and the events $A, B \subseteq \Omega$, the following properties are meet:

1. $P(\bar{A}) = 1 - P(A)$.
2. $P(\emptyset) = 0$.

Properties of the axiomatic probability

From the previous axioms is possible to deduce some important properties of a probability function.

Given a sample space Ω of a random experiment and the events $A, B \subseteq \Omega$, the following properties are meet:

1. $P(\bar{A}) = 1 - P(A)$.
2. $P(\emptyset) = 0$.
3. If $A \subseteq B$ then $P(A) \leq P(B)$.

Properties of the axiomatic probability

From the previous axioms is possible to deduce some important properties of a probability function.

Given a sample space Ω of a random experiment and the events $A, B \subseteq \Omega$, the following properties are meet:

1. $P(\bar{A}) = 1 - P(A)$.
2. $P(\emptyset) = 0$.
3. If $A \subseteq B$ then $P(A) \leq P(B)$.
4. $P(A) \leq 1$. This means that $P(A) \in [0, 1]$.

From the previous axioms is possible to deduce some important properties of a probability function.

Given a sample space Ω of a random experiment and the events $A, B \subseteq \Omega$, the following properties are meet:

1. $P(\bar{A}) = 1 - P(A)$.
2. $P(\emptyset) = 0$.
3. If $A \subseteq B$ then $P(A) \leq P(B)$.
4. $P(A) \leq 1$. This means that $P(A) \in [0, 1]$.
5. $P(A - B) = P(A) - P(A \cap B)$.

Properties of the axiomatic probability

From the previous axioms is possible to deduce some important properties of a probability function.

Given a sample space Ω of a random experiment and the events $A, B \subseteq \Omega$, the following properties are meet:

1. $P(\bar{A}) = 1 - P(A)$.
2. $P(\emptyset) = 0$.
3. If $A \subseteq B$ then $P(A) \leq P(B)$.
4. $P(A) \leq 1$. This means that $P(A) \in [0, 1]$.
5. $P(A - B) = P(A) - P(A \cap B)$.
6. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

From the previous axioms is possible to deduce some important properties of a probability function.

Given a sample space Ω of a random experiment and the events $A, B \subseteq \Omega$, the following properties are meet:

1. $P(\bar{A}) = 1 - P(A)$.
2. $P(\emptyset) = 0$.
3. If $A \subseteq B$ then $P(A) \leq P(B)$.
4. $P(A) \leq 1$. This means that $P(A) \in [0, 1]$.
5. $P(A - B) = P(A) - P(A \cap B)$.
6. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
7. If $A = \{e_1, \ldots, e_n\}$, where e_i $i = 1, \ldots, n$ are elemental events, then

$$P(A) = \sum_{i=1}^{n} P(e_i).$$

As set by the previous axioms, the probability of an event A, is a real number $P(A)$ that always ranges from 0 to 1.

In a certain way, this number expresses the plausibility of the event, that is, the chances that the event A occurs in the experiment. Therefore, it also gives a measure of the uncertainty about the event.

- The maximum uncertainty correspond to probability $P(A) = 0.5$ (A and $\bar{A}$ have the same chances of happening.)
- The minimum uncertainty correspond to probability $P(A) = 1$ (A will happen with absolute certainty) and $P(A) = 0$ (A won't happen with absolute certainty)

When $P(A)$ is closer to 0 than to 1, the chances of not happening A are grater than the chances of happening A. On the contrary, when $P(A)$ is closer to 1 than to 0, the chances of happening A are grater than the chances of not happening A.

Occasionally, we can get some information about the experiment before its realization. Usually that information is given as an event B of the same sample space that we know that is true before to conduct the experiment.

In such a case, we will say that B is a *conditioning* event and the probability of another event A known as a **conditional probability** and expressed

$$P(A|B).$$

This must be read as *probability of event A conditional on event B occurring*.

Usually, conditioning events change the sample space and therefore the probabilities of events.

Assume that we have a sample of 100 women and 100 men with the following frequencies

	Non-smokers	Smokers
Females	80	20
Males	60	40

Then, using the frequency definition of probability, the probability of being smoker from the whole sample is

$$P(\text{Smoker}) = \frac{60}{200} = 0.3.$$

Usually, conditioning events change the sample space and therefore the probabilities of events.

Assume that we have a sample of 100 women and 100 men with the following frequencies

	Non-smokers	Smokers
Females	80	20
Males	60	40

Then, using the frequency definition of probability, the probability of being smoker from the whole sample is

$$P(\text{Smoker}) = \frac{60}{200} = 0.3.$$

However, if we know that the person is a woman, then the sample is reduced to the first row, and the probability of being smoker is

$$P(\text{Smoker}|\text{Female}) = \frac{20}{100} = 0.2.$$

Definition (Conditional probability)

Given a sample space Ω of a random experiment, and two events $A, B \subseteq \Omega$, the probability of A conditional on B occurring is

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

as long as, $P(B) \neq 0$.

This definition allows to calculate conditional probabilities without changing the original sample space.

Example. In the previous example

$$P(\text{Smoker}|\text{Female}) = \frac{P(\text{Smoker} \cap \text{Female})}{P(\text{Female})} = \frac{20/200}{100/200} = \frac{80}{100} = 0.8.$$

From the definition of conditional probability it's possible to derive the formula for the probability of the intersection of two events.

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

Example. In a population there are a 30% of smokers and we know that there are a 40% of smokers with breast cancer. The probability of a random person being smoker and having breast cancer is

$$P(\text{Smoker} \cap \text{Cancer}) = P(\text{Smoker})P(\text{Cancer}|\text{Smoker}) = 0.3 \times 0.4 = 0.12.$$

Sometimes, the probability the conditioning event doesn't change the original probability of the main event.

Definition (Independent events)

Given a sample space Ω of a random experiment, two events $A, B \subseteq \Omega$ are *independents* if the probability of A doesn't change when conditioning on B, and vice-versa, that is,

$$P(A|B) = P(A) \quad \text{and} \quad P(B|A) = P(B),$$

if $P(A) \neq 0$ and $P(B) \neq 0$.

This means that the occurrence of one event doesn't give relevant information to change the uncertainty of the other.

When two events are independent, the probability of the intersection of them is the product of their probabilities,

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

The sample space of tossing twice a coin is $\Omega = \{(H,H),(H,T),(T,H),(T,T)\}$ and all the elements are equiprobable if the coin is fair. Thus, applying the classical definition of probability we have

$$P((H,H)) = \frac{1}{4} = 0.25.$$

If we name $H_1 = \{(H,H),(H,T)\}$, that is, having heads in the first toss, and $H_2 = \{(H,H),(T,H)\}$, that is, having heads in the second toss, we can get the same result assuming that these events are independent,

$$P(H,H) = P(H_1 \cap H_2) = P(H_1)P(H_2) = \frac{2}{4}\frac{2}{4} = \frac{1}{4} = 0.25.$$

Definition (Probabilistic space)

A *probabilistic space* of a random experiment is a triplet $(\Omega, \mathcal{F}, P)$ where

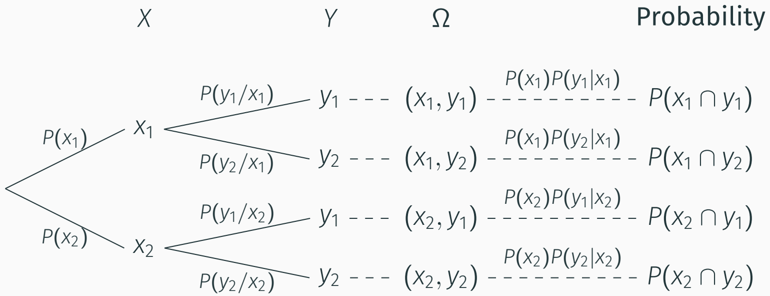
- Ω is the sample space of the experiment.
- $\mathcal{F}$ is a set of events of the experiment.
- P is a probability function.

If we know the probabilities of all the elements of Ω, then we can calculate the probability of every event in $\mathcal{F}$ and we can construct easily the probability space.

Probabilistic space construction

In order to determine the probability of every elemental event we can use a tree diagram, using the following rules:

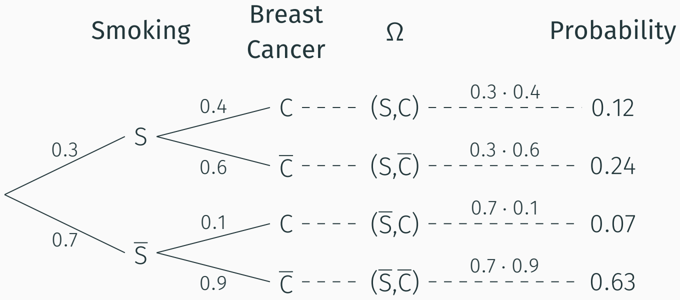
1. For every node of the tree label the incoming edge with the probability of the variable in that level having the value of the node, conditioned by events corresponding to its ancestor nodes in the tree.
2. The probability of every elemental event in the leaves is the product of the probabilities on edges that go form the root tho the leave.



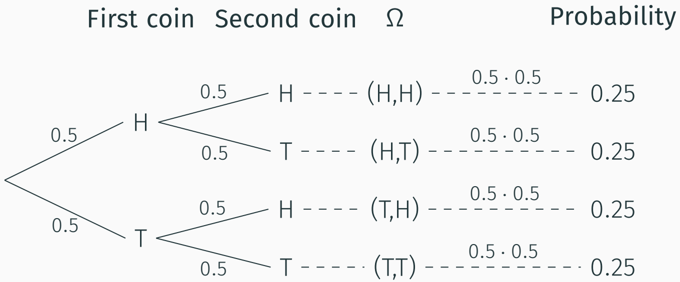
Probability tree with dependent variables

Example of smoking and cancer

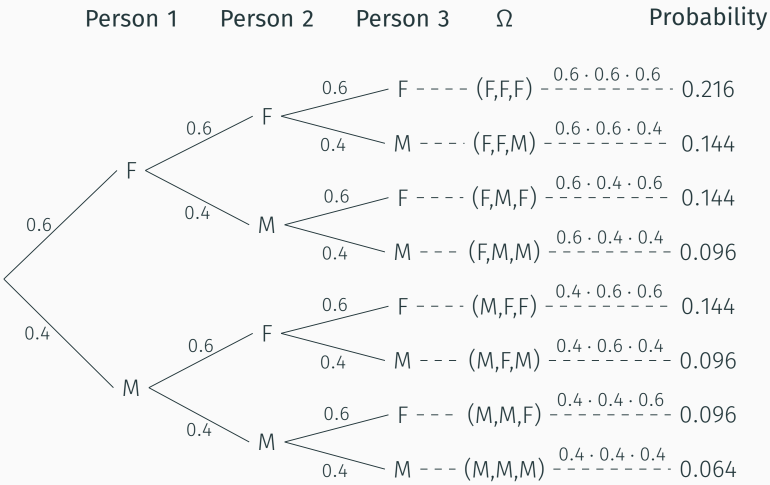
In a population there are a 30% of smokers and we know that there are a 40% of smokers with breast cancer, while only 10% of non-smokers have breast cancer. The probability tree of the probabilistic space of the random experiment consisting in picking a random person an measuring the variables smoking and breast cancer is below.



The probability tree of the random experiment of tossing two coins is below.



In a population there are 40% of males and 60% of females, the probability tree of drawing a random sample of three persons is below.



Partition of the sample space

Definition (Partition of the sample space)

A collection of events $A_1, A_2, \ldots, A_n$ of the same sample space Ω is a *partition* of the sample space if it meets the following conditions

1. The union of the events is the sample space, that is, $A_1 \cup \cdots \cup A_n = \Omega$.
2. All the events are mutually incompatible, that is, $A_i \cap A_j = \emptyset\ \forall i \neq j$.

Usually it's easy to get a partition of the sample space splitting a population according to some categorical variable, like for example gender, blood type, etc.

Total probability theorem

If we have a partition of a sample space, we can use it to calculate the probabilities of other events in the same sample space.

Theorem (Total probability)

Given a partition $A_1, \ldots, A_n$ of a sample space Ω, the probability of any other event B of the same sample space can be calculated with the formula

$$P(B) = \sum_{i=1}^{n} P(A_i \cap B) = \sum_{i=1}^{n} P(A_i)P(B|A_i).$$

Total probability theorem
Proof

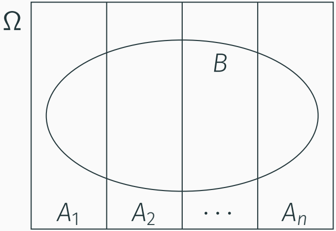
The theorem proof is quite simple. As $A_1, \ldots, A_n$ is a partition of Ω, we have

$$B = B \cap \Omega = B \cap (A_1 \cup \cdots \cup A_n) = (B \cap A_1) \cup \cdots \cup (B \cap A_n).$$

And all the events of this union are mutually incompatible as $A_1, \ldots, A_n$ are, thus

$$P(B) = P((B \cap A_1) \cup \cdots \cup (B \cap A_n)) = P(B \cap A_1) + \cdots + P(B \cap A_n) =$$

$$= P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n) = \sum_{i=1}^{n} P(A_i)P(B|A_i).$$



Total probability theorem
Proof

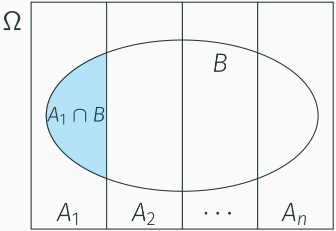
The theorem proof is quite simple. As $A_1, \ldots, A_n$ is a partition of Ω, we have

$$B = B \cap \Omega = B \cap (A_1 \cup \cdots \cup A_n) = (B \cap A_1) \cup \cdots \cup (B \cap A_n).$$

And all the events of this union are mutually incompatible as $A_1, \ldots, A_n$ are, thus

$$P(B) = P((B \cap A_1) \cup \cdots \cup (B \cap A_n)) = P(B \cap A_1) + \cdots + P(B \cap A_n) =$$

$$= P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n) = \sum_{i=1}^{n} P(A_i)P(B|A_i).$$



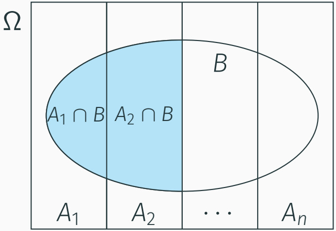
The theorem proof is quite simple. As $A_1, \dots, A_n$ is a partition of Ω, we have

$$B = B \cap \Omega = B \cap (A_1 \cup \cdots \cup A_n) = (B \cap A_1) \cup \cdots \cup (B \cap A_n).$$

And all the events of this union are mutually incompatible as $A_1, \dots, A_n$ are, thus

$$P(B) = P((B \cap A_1) \cup \cdots \cup (B \cap A_n)) = P(B \cap A_1) + \cdots + P(B \cap A_n) =$$

$$= P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n) = \sum_{i=1}^{n} P(A_i)P(B|A_i).$$



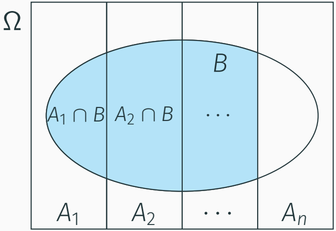
The theorem proof is quite simple. As $A_1, \ldots, A_n$ is a partition of Ω, we have

$$B = B \cap \Omega = B \cap (A_1 \cup \cdots \cup A_n) = (B \cap A_1) \cup \cdots \cup (B \cap A_n).$$

And all the events of this union are mutually incompatible as $A_1, \ldots, A_n$ are, thus

$$P(B) = P((B \cap A_1) \cup \cdots \cup (B \cap A_n)) = P(B \cap A_1) + \cdots + P(B \cap A_n) =$$

$$= P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n) = \sum_{i=1}^{n} P(A_i)P(B|A_i).$$



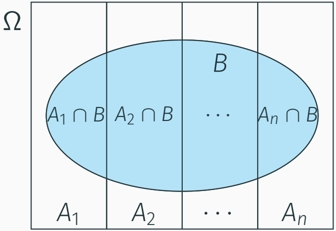
The theorem proof is quite simple. As $A_1, \dots, A_n$ is a partition of Ω, we have

$$B = B \cap \Omega = B \cap (A_1 \cup \dots \cup A_n) = (B \cap A_1) \cup \dots \cup (B \cap A_n).$$

And all the events of this union are mutually incompatible as $A_1, \dots, A_n$ are, thus

$$P(B) = P((B \cap A_1) \cup \dots \cup (B \cap A_n)) = P(B \cap A_1) + \dots + P(B \cap A_n) =$$

$$= P(A_1)P(B|A_1) + \dots + P(A_n)P(B|A_n) = \sum_{i=1}^{n} P(A_i)P(B|A_i).$$



A symptom S can be caused by a disease D, but can also be present in persons without the disease. In a population, the rate of people with the disease is 0.2. We know also that 90% of persons with the disease present the symptom, while only 40% of persons without the disease present it.

What is the probability that a random person of the population presents the symptom?

To answer the question we can apply the total probability theorem using the partition $\{A, \bar{A}\}$:

$$P(S) = P(D)P(S|D) + P(\bar{D})P(S/\bar{D}) = 0.2 \cdot 0.9 + 0.8 \cdot 0.4 = 0.5.$$

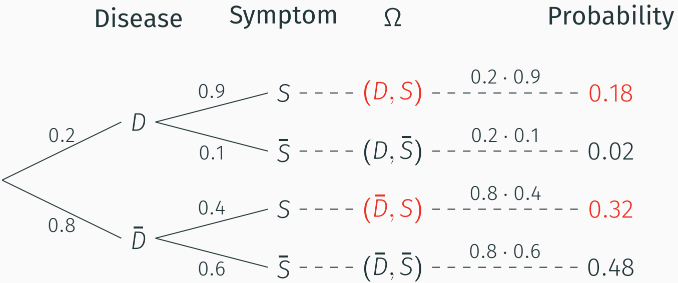
That is, half of the population have the symptom.

Indeed, it is a weighted mean of probabilities!

Total probability theorem

Example of diagnosis with a tree diagram

The answer to the previous question is even clearer with the tree diagram of the probabilistic space.



$$P(S) = P(D,S) + P(\bar{D},S) = P(D)P(S|D) + P(\bar{D})P(S|\bar{D})$$
$$= 0.2 \cdot 0.9 + 0.8 \cdot 0.4 = 0.18 + 0.32 = 0.5.$$

A partition of a sample space $A_1, \cdots, A_n$ may also be interpreted as a set of feasible hypothesis to a fact B.

In such cases may be helpful to calculate the posterior probability $P(A_i|B)$ of every hypothesis.

Theorem (Bayes)

Given a partition $A_1, \ldots, A_n$ of a sample space Ω and another event B of the same sample space, the conditional probability of every even A_i $i = 1, \ldots, n$ on B can be calculated with the following formula

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}.$$

In the previous example, a more interesting question is about the diagnosis for a person with the symptom.

In this case we can interpret D and $\bar{D}$ as the two feasible hypothesis for the symptom S. The prior probabilities for them are $P(D) = 0.2$ and $P(D) = 0.8$. That means that if we don't have information about the symptom, the diagnosis would be that the person doesn't have the disease.

However, if after examining the person we observe the symptom, that information changes the uncertainty about the hypothesis, an we need calculate the posterior probabilities to diagnose, that is,

$$P(D|S) \text{ y } P(\bar{D}|S)$$

To calculate the posterior probabilities we can use the Bayes theorem.

$$P(D/S) = \frac{P(D)P(S/D)}{P(D)P(S/D)+P(\bar{D})P(S/\bar{D})} = \frac{0.2\cdot 0.9}{0.2\cdot 0.9+0.8\cdot 0.4} = \frac{0.18}{0.5} = 0.36,$$
$$P(\bar{D}/S) = \frac{P(\bar{D})P(S/\bar{D})}{P(D)P(S/D)+P(\bar{D})P(S/\bar{D})} = \frac{0.8\cdot 0.4}{0.2\cdot 0.9+0.8\cdot 0.4} = \frac{0.32}{0.5} = 0.64.$$

As we can see the probability of having the disease has increased. Nevertheless, the probability of not having the disease is still greater than the probability of having it, and for that reason, the diagnosis is not having the disease.

In this case it is said the the symptom S is *not decisive* in order to diagnose the disease.