Class 9: Probability

Objective

Define the basic laws of probability and their interpretation.
Introduce the concept of a random variable.

- Random experiments, sample space, elementary and composite events.
- Definition of probability.
- Properties of probability.
- Conditional probability and independence.
- Random variables.
- Supplementary notes:The law of total probability and Bayes theorem.

Recommended reading:

- Have a look at this video.

Random experiments etc.

Suppose that we are going to carry out a random experiment and that we are interested in the probability that a particular event occurs.

Experiment: Ask a Spanish adult who they voted for at the last election

Sample space S: The set of all the basic results of the experiment. For example {Didn't vote, Cs, PP, PSOE, Podemos, …}

Elementary event: Any of the basic results of the experiment. Cs

Composite event: They voted for a left wing party {PSOE, Podemos, ERC, …}

Definition of probability

Probability is based on a mathematical theory (Kolmogorov axioms) and has various interpretations.

1. Classical probability
2. Frequentist probability
3. Subjective probability
4. Philosophical interpretations

Classical probability

Consider an experiment where all of the elementary events are equally likely. If there are K elementary events in the sample space, the probability of an event A is

Probability of A = P(A) = number of elementary events in A x 1/K

The table shows the list of presidents since Spain became a Democracy. If a person is selected at random from the list, what is the probability they are from PSOE?

Calculate the chance that they governed for at least 2 legislatures.

	Presidente	Partido	Mandato
1.	Adolfo Suárez González	UCD	2 de abril de 1979 - 26 de febrero de 1981 (I Legislatura)
2.	Leopoldo Calvo-Sotelo	UCD	26 de febrero de 1981 - 2 de diciembre de 1982 (I Legislatura)
3.	Felipe González Márquez	PSOE	2 de diciembre de 1982 - 5 de mayo de 1996 (II, III, IV y V Legislaturas)
4.	José María Aznar	PP	5 de mayo de 1996 - 17 de abril de 2004 (VI y VII Legislaturas)
5.	José Luis Rodríguez Zapatero	PSOE	17 de abril de 2004 - 21 de diciembre de 2011 (VIII y IX Legislaturas)
6.	Mariano Rajoy Brey	PP	21 de diciembre de 2011 - 1 de junio de 2018 (X, XI y XII Legislaturas)
7.	Pedro Sánchez Pérez-Castejón	PSOE	2 de junio de 2018 - actualidad (XII Legislatura)

Classical probability

The concept of classical probability can be applied in many of the examples in our course, when, for example, we are choosing a person at random from a survey.

In other cases, the concept of elementary events with equal probabilities is not reasonable:

- Getting the answer right on a multiple choice question. (If we haven't studied, classical probability will apply, but if we have studied, we should be more likely to get the correct answer).
- The chance Rayo Vallecano will win the league.
- The chance Cataluña will be independent within 10 years.

In these cases, we need to use different interpretations of probability.

Frequentist probability

If we repeat an experiment many times, the (relative) frequency of times that a particular event occurs will approximate the probability.

Probability = the limit of the relative frequency.

Subjective probability

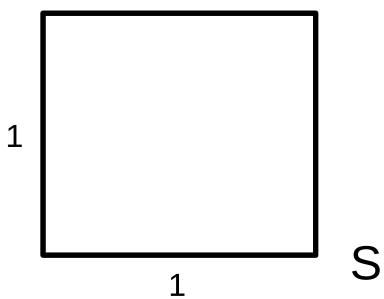
Each individual has their own probabilities which depend on their own personal knowledge, experience and uncertainty.

Philosophical interpretations

Philosophy is bunk!

Properties of probability

Areas in Venn diagrams obey the same rules as probability.

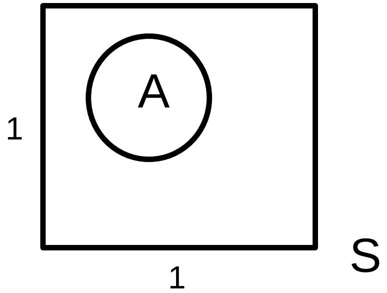


P(S) = 1.

When we do an experiment, the probability that one of the possible results occurs is 1.

Properties of probability

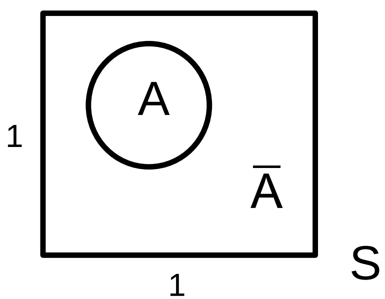
Areas in Venn diagrams obey the same rules as probability.



For any event A, $0 \leq P(A) \leq 1$.

Properties of probability

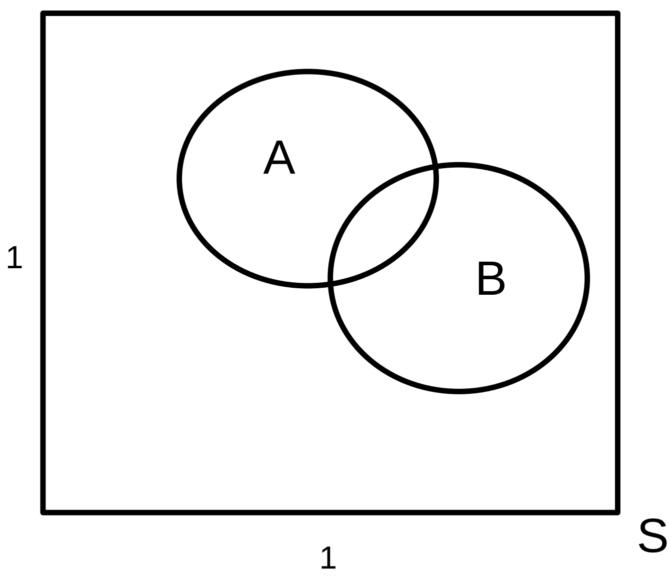
Areas in Venn diagrams obey the same rules as probability.



The chance that A doesn’t happen is:
$P(\bar{A}) = 1 - P(A)$.

Properties of probability

Areas in Venn diagrams obey the same rules as probability.



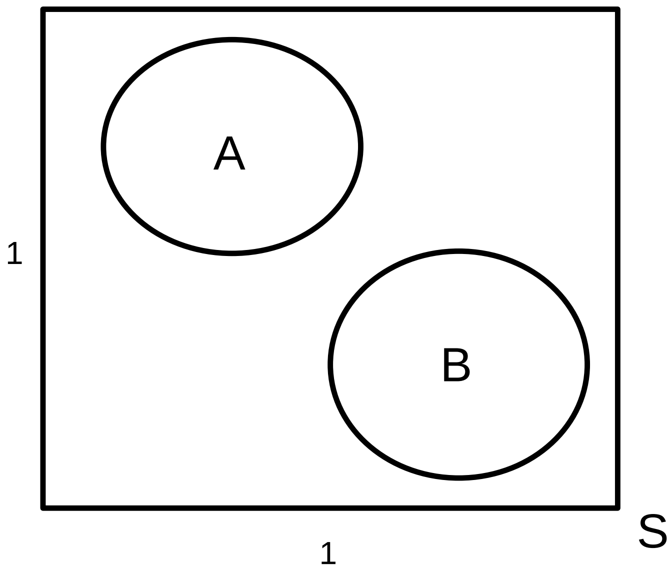
The chance that at least one of A and B happen is:

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

What do the symbols ∪ and ∩ mean in normal language?

Properties of probability

Areas in Venn diagrams obey the same rules as probability.



If two events are incompatible, then the chance that at least one of them occurs is:

$P(A \cup B) = P(A) + P(B)$.

What does the word "incompatible" mean in everyday language?

Probability and frequency

Probability and frequency both obey the same rules.

Previously, we asked “What proportion of the people in the sample voted for Daniel Peña?”

Now we could say: “If we choose one of the people in the sample at random, what is the chance they voted for Daniel Peña?

Would the answer be different?

	Luciano Parejo	Daniel Peña	Null votes	Abstentions
Getafe	108	111	8	12
Leganés	26	82	3	4
Colmenarejo	6	13	1	1

Conditional probability

The conditional probability of A given B is: $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$.

What does the | mean? Are P(A|B) and P(B|A) the same?

Assuming a voter is selected at random:

a) What is the probability they come from Getafe and voted for Daniel Peña?
b) What is the probability they come from Getafe?
c) Assuming they come from Getafe, what is the probability they voted for Daniel Peña?

	Luciano Parejo	Daniel Peña	Null votes	Abstentions
Getafe	108	111	8	12
Leganés	26	82	3	4
Colmenarejo	6	13	1	1

The multiplication law

This is just a rewriting of the conditional probability formula as:

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

Why is this useful?

Assume that two voters are selected at random. What is the chance they both abstained?

	Luciano Parejo	Daniel Peña	Null votes	Abstentions
Getafe	108	111	8	12
Leganés	26	82	3	4
Colmenarejo	6	13	1	1

Independence

Two events, A and B, are said to be statistically independent if: $P(A \cap B) = P(A)P(B)$.

How do we interpret this?

Try dividing both sides by $P(B)$. Independence implies:

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

If we learn that B occurs, then the probability of A does not change.

Equally we have $P(B|A) = P(B)$ and observing A does not change the probability of B.

Independence

In reality, few events are independent. Exceptions are things like coin tosses where the result of the last toss doen't affect the next.

Are the events "come from Getafe" and "voted for Daniel Peña" independent? Why?

	Luciano Parejo	Daniel Peña	Null votes	Abstentions
Getafe	108	111	8	12
Leganés	26	82	3	4
Colmenarejo	6	13	1	1

Exercise

The following table gives the 17 cabinet ministers of the British coalition government, their Political party, age and sex.

Name	Ministry	Political Party	Age	Sex
David Cameron	Prime Minister	Conservative	43	Male
Nick Clegg	Deputy Prime Minister	Liberal Democrat	43	Male
William Hague	Foreign Affairs	Conservative	49	Male
George Osborne	Exchequer	Conservative	38	Male
Liam Fox	Defence	Conservative	48	Male
Kenneth Clarke	Justice	Conservative	69	Male
Patrick McCoughlin	Chief Whip	Conservative	52	Male
Theresa May	Home Secretary	Conservative	53	Female
Andrew Lansley	Health	Conservative	53	Male
David Laws	Treasury	Liberal Democrat	44	Male
Vince Cable	Business	Liberal Democrat	67	Male
Michael Gove	Education	Conservative	42	Male
Eric Pickles	Local Government	Conservative	58	Male
Chris Huhne	Energy and Climate Change	Liberal Democrat	55	Male
Danny Alexander	Scotland	Liberal Democrat	38	Male
Iain Duncan Smith	Work and Pensions	Conservative	56	Male
Dominic Grieve	Attorney General	Conservative	53	Male

Suppose that a cabinet member is selected at random.

a) What is the probability that they belong to the Liberal Democratic Party?
b) What is the probability that they are under 60 or female?
c) Are the events “aged under 60” and “female” independent? Why?
d) Supposing that the selected cabinet minister is Conservative, calculate the probability that they are aged under 50.

Exercise

The following table comes from a survey of 1140 American voters (508 Republicans, 462 Democrats and the remainder Independents) carried out by Quinnipiac University in December 2015. (Note that each row in the table sums to 100%).

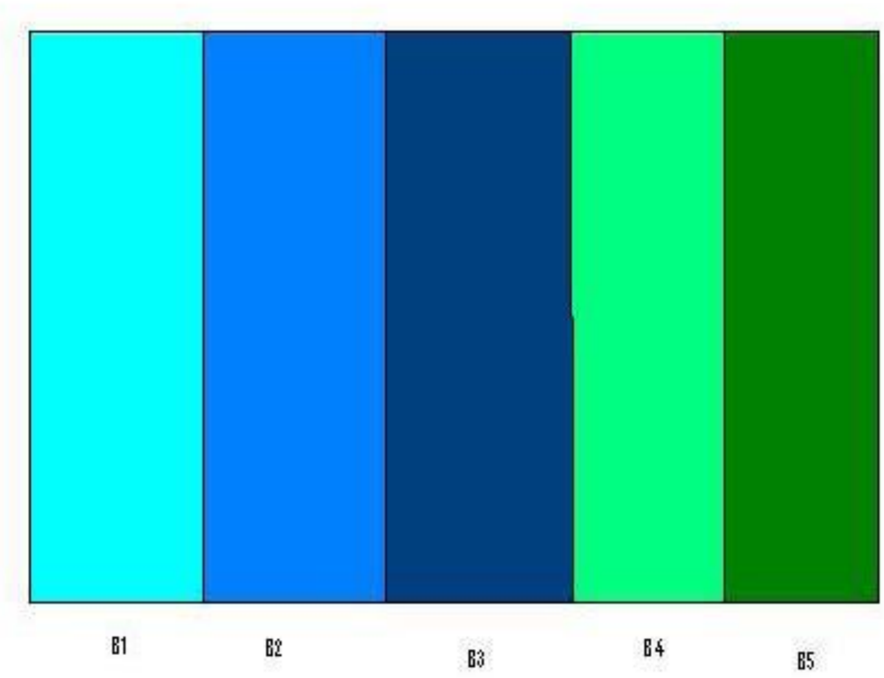
"Do you support or oppose accepting Syrian refugees into the U.S.?"

	Support %	Oppose %	Unsure/ No answer %
12/16-20/15	43	51	5
Republicans	13	82	5
Democrats	74	22	4
Independents	42	51	6

Suppose that one of the participants in the survey is chosen at random.

a) What is the probability that they are Republican?
b) Assuming they are Republican, calculate the probability that they oppose accepting Syrian refugees into the U.S.
c) Are the two events "Republican" and opposing accepting Syrian refugees into the U.S. independent? Why?
d) If the person selected opposes accepting Syrian refugees, what is the chance that they are Republican?

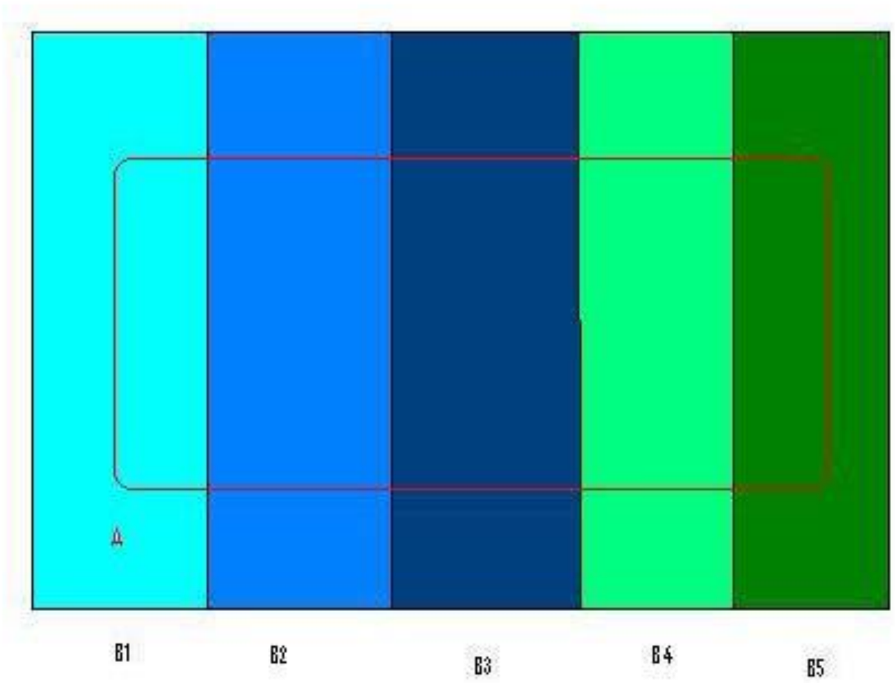
Supplementary material: The law of total probability and Bayes theorem



The events B_1, …, B_k form a partition if:

$S = B_1$ or B_2 or … or B_k
B_i and $B_j = \phi$

The law of total probability



For an event A, $P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots + P(A \cap B_k)$

$$= P(A|\, B_1)P(B_1) + P(A|\, B_2)P(B_2) + \ldots + P(A|\, B_k)P(B_k)$$

Bayes theorem

Given that A has ocurred, the probability of B_i is:

$$P(B_i | A) = P(A \cap B_i)/P(A) = P(A|B_i)P(B_i)/P(A)$$

Example

We have 3 urns: A with 3 red balls and 5 black, B with 2 red balls and 1 black and C with 2 red balls and 3 black. We pick an urn at random and take a ball out. If the ball is red, what is the chance we picked it from urn A?