Chapter 4: Probability

1. Random experiments, sample space, elementary and composite events.
2. Definition of probability.
3. Properties of probability.
4. Conditional probability, the multiplication law and independence.
5. The law of total probability and Bayes theorem.

Recommended reading:

- Chapters 13 and 14 of Peña and Romo (1997)

4.1: 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 Ω: The set of all the basic results of the experiment. For example {null, PSOE, PP, IU, …}

ELEMENTARY EVENT: Any of the basic results of the experiment. IU

COMPOSITE EVENT: They voted for a left wing party {PSOE, IU, …}

4.2: 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, then the probability of an event A is

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

¿What is the probability of seeing exactly two heads if we toss a fair coin twice?

¿In the voting example, is it reasonable to assume that all elementary events are equally probable?

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

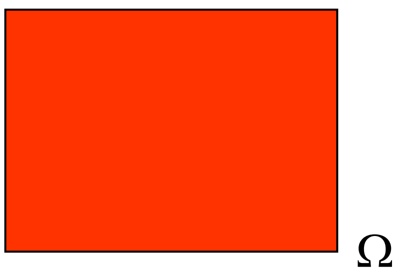
Each individual has their own probabilities which depend on their own personal knowledge, experience and uncertainty.

Philosophical interpretations

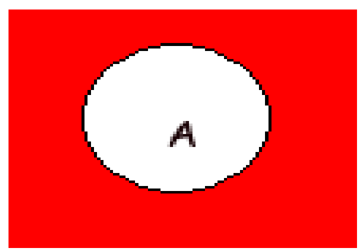
Philosophy is bunk!

4.3: Properties of probability

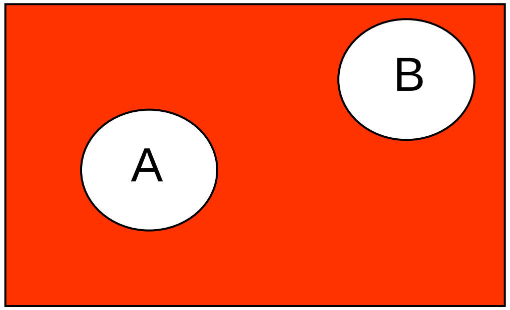
Venn diagrams obey the same rules as probability.

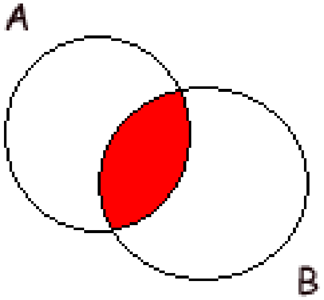


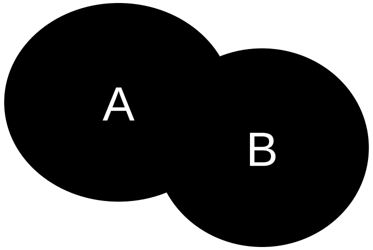
P(Ω) = 1



0 ≤ P(A) ≤ 1







How do we calculate P(A or B)?

- If A is an event in Ω then $0 \le P(A) \le 1$
- If $A=\{e_1, e_2, \ldots, e_k\}$, then $P(A) = P(e_1) + P(e_2) + \ldots + P(e_k)$
- $P(\Omega)=1$ y $P(\varnothing)=0$
- The law of complements: $P(\bar{A}) = 1 - P(A)$
- Addition law: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- If A and B are incompatible, then $P(A \cap B) = 0$ and

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

Example: Given the table (job versus family earnings)

	Bajo	Medio	Alto
Ama de casa	8	26	6
Obreros	16	40	14
Ejecutivos	6	62	12
Profesionales	0	2	8

we choose a person at random. Calculate the probability of:

a) Ama de casa b) Obrero c) Ejecutivo d) Profesional

e) Ingreso bajo f) Ingreso medio g) Ingreso alto

h) Ejecutivo con ingreso alto i) Ama de casa con ingreso bajo

4.4: Conditional probability, the multiplication rule and independence

The conditional probability of A given B is:

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

Another way of writing this is the multiplication rule

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

Deal two cards from a Spanish pack. What is the chance that they are both oros?

Independence

A and B are said to be independent if:

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

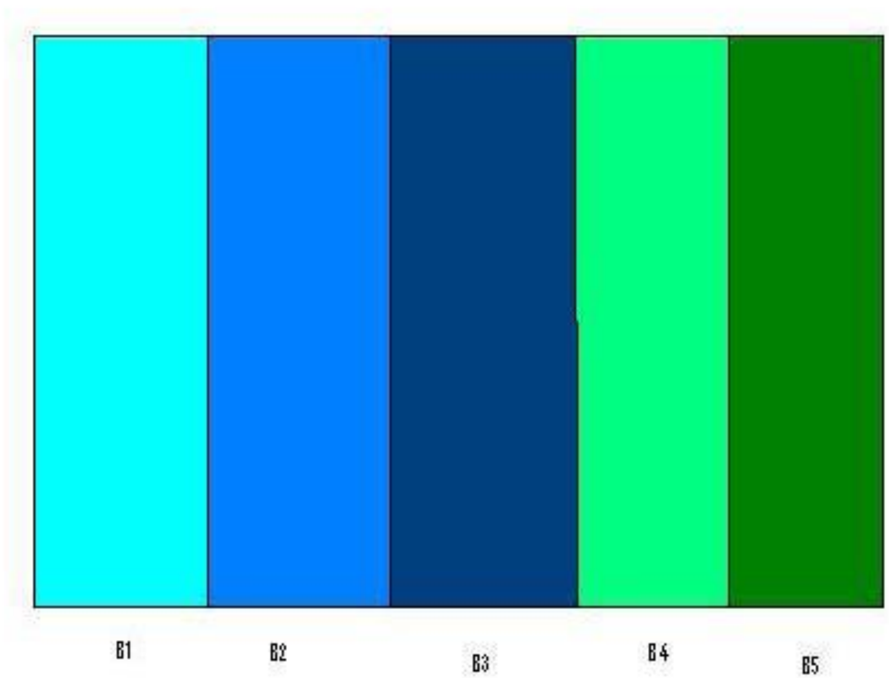
This is equivalent to saying that P(A|B) = P(A) and means that knowing that B has ocurred does not change the uncertainty about A.   Equally we have P(B|A) = P(A) and observing A does not change the probability of B.

Example

A tv station surveyed 2500 personas to find out how many of them watched a debate and/ or a film shown at two different times: 2 100 saw the film, 1 500 watched the debate and 350 didn’t see either of the programs. If we pick a person at random from the survey group:

a) What is the chance that they saw both the film and the debate?
b) What is the chance that they saw the film, assuming that they watched the debate?
c) If they saw the film, what is the chance they saw the debate?

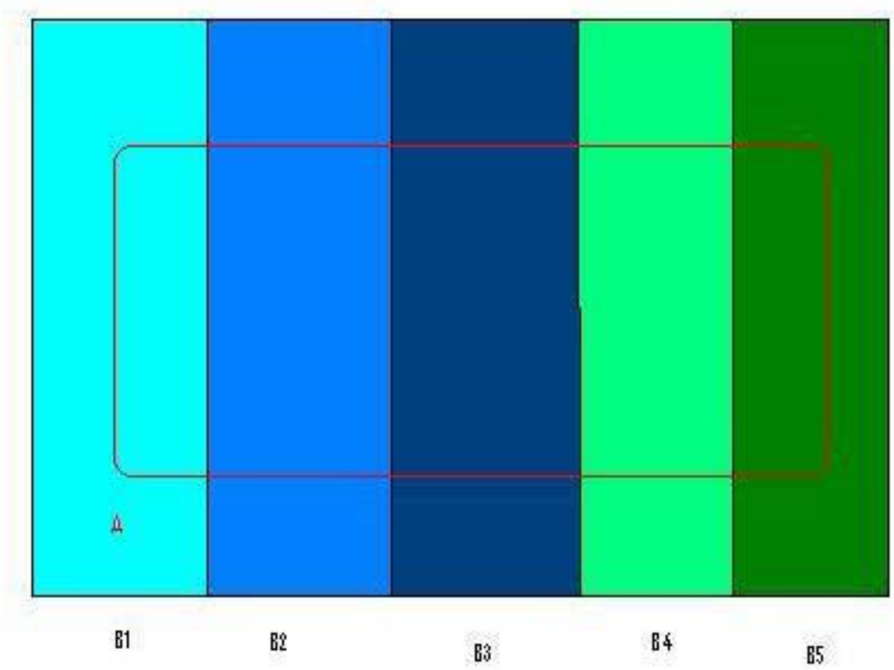
4.5: The law of total probability and Bayes theorem



The events $B_1, \ldots, B_k$ form a partition if:

$\Omega = B_1$ 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?

Example

In the recent rectoral elections, the votes emitted by postgraduate students were as follows:

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

A person is selected at random amongst the postgraduate voters:

a) What is the probability that they voted for Daniel Peña?
b) Assuming they come from Leganés, what is the probability they voted for Daniel Peña?
c) Are the two events “voted for Daniel Peña” and “come from Leganés” independent? Why?

Example

Crisis-hit Spanish town votes to grow cannabis commercially to pay off debts

A Spanish town has voted in favour of growing cannabis to help pay off its spiralling debt.

The inhabitants of Rasquera were consulted in a referendum on whether to rent out seven hectares of municipal land for use as a cannabis farm.

Mayor Bernat Pellisa believes the cash-strapped town council can pay off its £1.25million debt within two years by following the scheme.

Some 56.3 per cent of the population voted in favour of allowing cannabis to be grown, with 43.7 per cent against, in Tuesday's referendum.

The small town plans to lease farmland to a cannabis consumers' association which will grow, cultivate and smoke the drug.
…
The turnout was 68 per cent of the town's registered population of 804 over 18s.

a) How many people from Rasquera voted?
b) If two people from Rasquera are chosen at random, what is the probability that they both voted in favour of growing marijuana?

Daily Mail, 11th April 2012:

Example

The following table is taken from a US student survey on presidential preferences and ideologies. From the information in the table, are the events “preference for Bush” and “Conservative ideology” independent? Why?

TABLE 3. PRESIDENTIAL PREFERENCE BY IDEOLOGY

IDEOLOGY

PRES PREF	Liberal	Slightly Liberal	Moderate	Slightly Cons.	Conser-vative	*Total*
Kerry	13	10	3	1	0	27
Bush	1	2	3	3	5	14
Total	14	12	6	4	5	41

$a \approx 0.7$

Source: POLI 300, Student Survey, Fall 2007

Example

The following table gives the 17 cabinet ministers of the new 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? (**0,5 points**)
b) What is the probability that they are under 60 or female? (**0,5 points**)
c) Are the events “aged under 60” and “female” independent? Why? (**0,5 points**)
d) Supposing that the selected cabinet minister is Conservative, calculate the probability that they are aged under 50. (**0,5 points**)
e) Calculate the mode and median of the distribution of ages. (**0,5 points**)

Example

The following table is taken from the CIS Barometer of May 2013.

PREGUNTA 1
Para empezar, refiriéndonos a la situación económica general de España, ¿cómo la calificaría Ud.: muy buena, buena, regular, mala o muy mala?

	Hombre		Mujer		TOTAL	
	%	(N)	%	(N)	%	(N)
Buena	1.3	(16)	0.6	(8)	1.0	(24)
Regular	8.1	(98)	8.5	(107)	8.3	(205)
Mala	37.8	(458)	31.2	(392)	34.5	(850)
Muy mala	52.6	(637)	59.6	(748)	56.1	(1385)
N.C.	0.2	(2)	0.1	(1)	0.1	(3)
TOTAL	100.0	(1211)	100.0	(1256)	100.0	(2467)

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

a) What is the probability that they are female?
b) Supposing that they say that the economic situation is good (*buena*), what is the probability that they are female?
c) Are the two events being female and saying *buena* independent? Why?