Stats1 Chapter 5: Probability

Calculating Probability

Experimental

i.e. Dealing with collected data.

Chp1: Data Collection

Methods of sampling, types of data, and populations vs samples.

Chp2: Measures of Location/Spread

Statistics used to summarise data, including mean, standard deviation, quartiles, percentiles. Use of linear interpolation for estimating medians/quartiles.

Chp3: Representation of Data

Producing and interpreting visual representations of data, including box plots and histograms.

Chp4: Correlation

Measuring how related two variables are, and using linear regression to predict values.

Theoretical

Deal with probabilities and modelling to make inferences about what we 'expect' to see or make predictions, often using this to reason about/contrast with experimentally collected data.

Chp5: Probability

Venn Diagrams, mutually exclusive + independent events, tree diagrams.

Chp6: Statistical Distributions

Common distributions used to easily find probabilities under certain modelling conditions, e.g. binomial distribution.

Chp7: Hypothesis Testing

Determining how likely observed data would have happened 'by chance', and making subsequent deductions.

This Chapter Overview

This chapter is a recap of the concepts you learnt at GCSE.
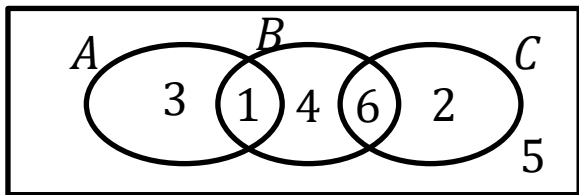
1 :: Basic Probability

"I throw two fair die. Calculate the probability the sum of the two dice is more than 6."

2 :: Venn Diagrams

"Out of 50 students, 12 play both piano and drums, 30 play piano and 25 play drums. Find the probability a randomly chosen student plays neither instrument."

3 :: Mutually Exclusive/Independent Events

Determine whether A and B are independent.

4 :: Tree Diagrams

"The probability I hit a target is 0.3. If I hit it, the probability I hit again on the next shot is 0.4. If I miss, the probability I hit on the next shot is 0.1. If I shoot 3 times, what's the probability I hit on the first and third shot?"

Probability concepts

✎ An **experiment** is a repeatable process that gives rise a number a number of **outcomes**.

✎ An **event** is a set of <u>one or more</u> of these outcomes.

(We often use capital letters to represent them)

S
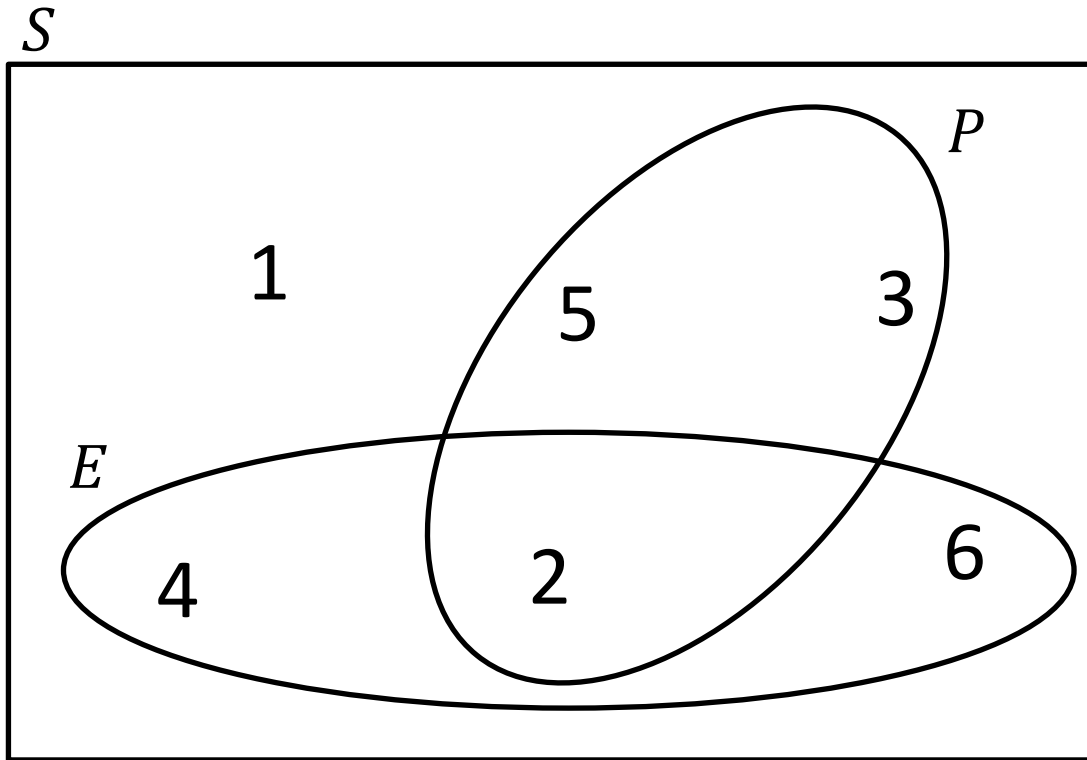
$E =$ "rolling an even number"
$P =$ "rolling a prime number"

✎ A **sample space** is the set of all possible outcomes.

Because we are dealing with sets, we can use a **Venn diagram**, where
- the numbers are the individual outcomes,
- the sample space is a rectangle and
- the events are sets, each a subset of the sample space.

You do not need to use set notation like ∩ and ∪ in this module (but ordinarily you would!)

Example

Two fair spinners each have four sectors numbered 1 to 4. The two spinners are spun together and the sum of the numbers indicated on each spinner is recorded.
Find the probability of the spinners indicating a sum of
(a) exactly 5   (b) more than 5

Spinner 1

+	1	2	3	4
1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8

Spinner 2

$$P(5) \quad = \frac{4}{16} = \frac{1}{4}$$

$$P(> 5) = \frac{6}{16} = \frac{3}{8}$$

If the sample space is the amalgamation of two underlying experiments, a table is a helpful way to list the outcomes.

Another Example

The table shows the times taken, in minutes, for a group of students to complete a number puzzle.

Time, t (min)	$5 \leq t < 7$	$7 \leq t < 9$	$9 \leq t < 11$	$11 \leq t < 13$	$13 \leq t < 15$
Frequency	6	13	12	5	4

A student is chosen at random. Find the probability for a group of students to complete a number puzzle
(a) In under 9 minutes   (b) in over 10.5 minutes.

$$P(\leq 9) = \frac{6 + 13}{40} = \frac{19}{40}$$

$$P(\geq 10.5) = \frac{3 + 5 + 4}{40} = \frac{3}{10}$$

You're effectively doing **linear interpolation** here: 10.50 to 11 is a quarter of the $9 \leq t < 11$ interval, therefore we use a quarter of the 12 people. This is an estimate because we're assuming the people are equally distributed across the interval.

Pearson Applied Year 1/AS
Page 31

1 Two coins are tossed. Find the probability of both coins showing the same outcome.

2 Two six-sided dice are thrown and their product, X, is recorded.

  a Draw a sample space diagram showing all the possible outcomes of this experiment.

  b Find the probability of each event:

    i $X = 24$ ii $X < 5$ iii X is even.

3 The masses of 140 adult Bullmastiffs are recorded in a table.
One dog is chosen at random.

  a Find the probability that the dog has a mass of 54 kg or more.

  b Find the probability that the dog has a mass between 48 kg and 57 kg.

The probability that a Rottweiler chosen at random has a mass under 53 kg is 0.54.

  c Is it more or less likely that a Bullmastiff chosen at random has a mass under 53 kg? State one assumption that you have made in making your decision.

Mass, m (kg)	Frequency
$45 \leqslant m < 48$	17
$48 \leqslant m < 51$	25
$51 \leqslant m < 54$	42
$54 \leqslant m < 57$	33
$57 \leqslant m < 60$	21
$60 \leqslant m < 63$	2

Hint
Use interpolation.

4 The lengths, in cm, of 240 koalas are recorded in a table.

One koala is chosen at random.

a Find the probability that the koala is female.

b Find the probability that the koala is less than 80 cm long.

c Find the probability that the koala is a male between 75 cm and 85 cm long.

Koalas under 72 cm long are called juvenile.

d Estimate the probability that a koala chosen at random is juvenile. State one assumption you have made in making your estimate.

Length, l (cm)	Frequency (male)	Frequency (female)
$65 \leqslant l < 70$	4	14
$70 \leqslant l < 75$	20	15
$75 \leqslant l < 80$	24	32
$80 \leqslant l < 85$	47	27
$85 \leqslant l < 90$	31	26

5 The histogram shows the distribution of masses, in kg, of 70 adult cats.

a Find the probability that a cat chosen at random has a mass more than 5 kg. **(2 marks)**

b Estimate the probability that a cat chosen at random has a mass less than 6.5 kg. **(3 marks)**
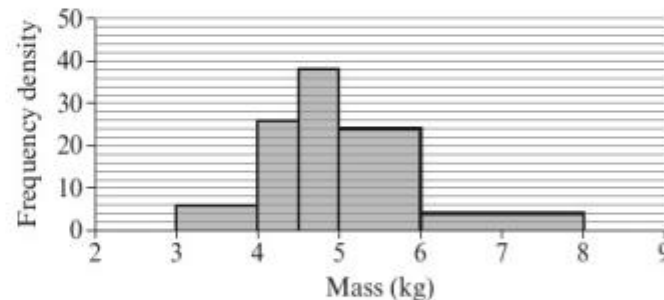
Challenge

Samira picks one card at random from group A and one card at random from group B.

She records the product, Y, of the two cards as the result of her experiment. Given that x is an integer and that $P(Y \text{ is even}) = P(Y \geqslant 20)$, find the possible values of x.

A: 2 7 5
B: 4 x

1 $\frac{1}{2}$

2 a

		Second roll					
		1	**2**	**3**	**4**	**5**	**6**
First roll	**1**	1	2	3	4	5	6
	2	2	4	6	8	10	12
	3	3	6	9	12	15	18
	4	4	8	12	16	20	24
	5	5	10	15	20	25	30
	6	6	12	18	24	30	36

b i $\frac{1}{18}$ **ii** $\frac{2}{9}$ **iii** $\frac{3}{4}$

3 a $\frac{2}{5}$

b $\frac{5}{7}$

c Less likely; frequency uniformly distributed throughout the class.

4 a $\frac{19}{40}$ **b** $\frac{109}{240}$ **c** $\frac{71}{240}$

d $\frac{2}{15}$; distribution of lengths of koalas between 70 and 75 cm is uniform.

5 a $\frac{16}{35}$

b $\frac{32}{35}$

Challenge:
5, 7 or 9