statsGCSE

Statistics GCSE

Paper 1

Edexcel Higher - 2026

Higher Tier

Variant 2 1ST0/1H

Instructions

- Write all answers in the spaces provided.
- Answer all questions.
- You must show all your working.
- There may not be enough space to show all your working out.

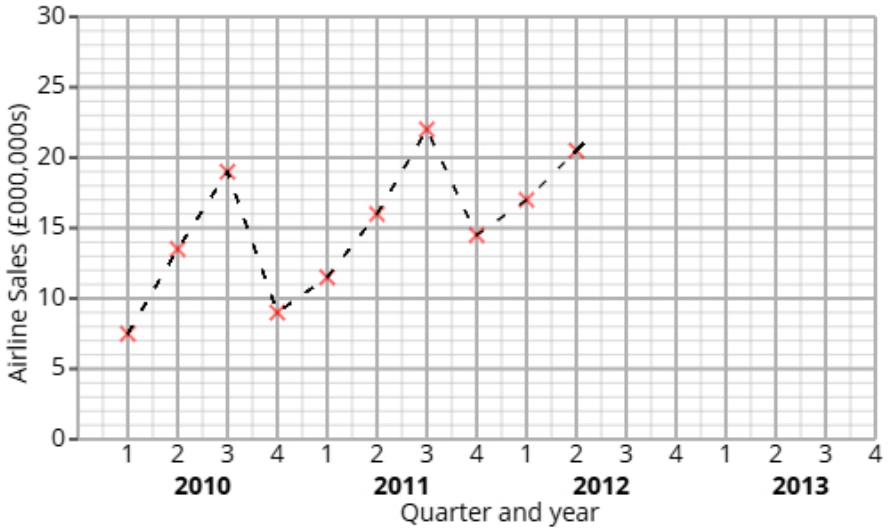
Information

- This is a practise paper to aid your revision for your exams.
- This site, and all that work on it, have no affiliation or relationship with any exam board.
- This site is not endorsed by any company or charity, unless we state otherwise.

Advice

- You can get support for all these questions at our website: www.statsgcse.com
- This paper and more are available on our site with questions that change subtly after each attempt.
- Good luck!

1 The time series graph shows information about the the sales for an airline from 2010 to 2012.



Liam calculates the 4-point moving averages from the time series graph, which are shown below.

12.5 13.5 14 14.5 16 17.5 18.5

(a) Identify and interpret in context one example of seasonality displayed in the time series graph.

(2 marks)

(b) Liam uses the time series graph to estimate that there was £1950000 of airline sales in Q1 2013

i) Plot the moving averages onto the time series graph and draw a trend line from 2010 to 2012.

ii) Describe the trend.

iii) By using the average seasonal effect for Q1, show that Liam's estimate is reasonable.

(7 marks)

(c) Explain why a 4-point moving average is appropriate.

(1 mark)

2 A fair 3-sided spinner is numbered 1, 2, 3.

A fair 5-sided spinner is numbered 1, 2, 3, 4, 5.

The spinners are used to play a game. Both spinners are spun and the total score is recorded.

		5-sided spinner				
		1	**2**	**3**	**4**	**5**
3-sided spinner	**1**	2	3			
	2	3				
	3					

The game is won when the total is at least 6.

Maya plays the game once.

(a) Complete the sample space diagram.

(2 marks)

(b) Find the probability that Maya wins the game.

(2 marks)

3 Liam organises two fitness programs, Program A and Program B, to help people improve their stamina.

He wants to compare the two programs to see which improves stamina better.

The table shows number of participants who passed and failed the fitness test for each program.

	Passed	Failed	Total
Program A	10	15	25
Program B	12	36	48

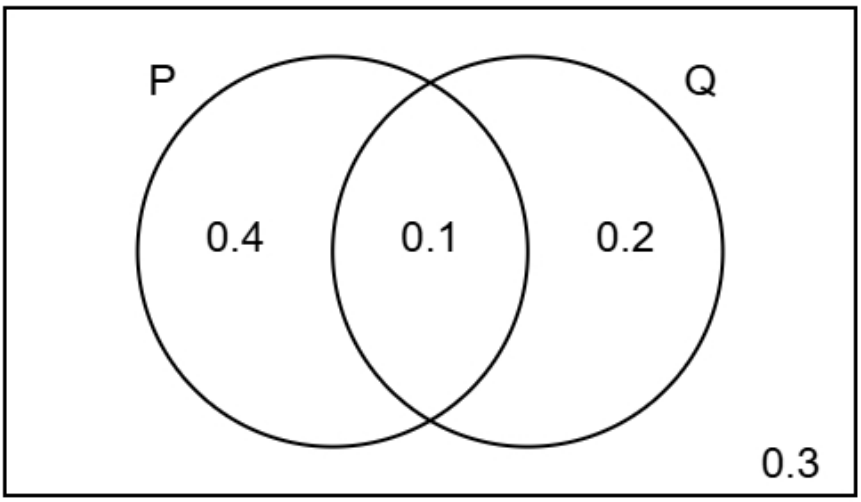
(i) Find the relative risk of failing the fitness test having taken Program A compared to Program B.

(ii) Give an interpretation of your answer to part (i).

(4 marks)

4 The Venn diagram shows information about the probabilities of two events occurring.

The events are labelled as P and Q.



(a) Find the probability of event Q happening.

(1 mark)

(b) Find P(P and Q)

(1 mark)

(c) Find $P(Q \mid P)$

(2 marks)

(d) Two different events events X and Y are independent.

$P(X) = 0.5$

$P(Y) = 0.9$

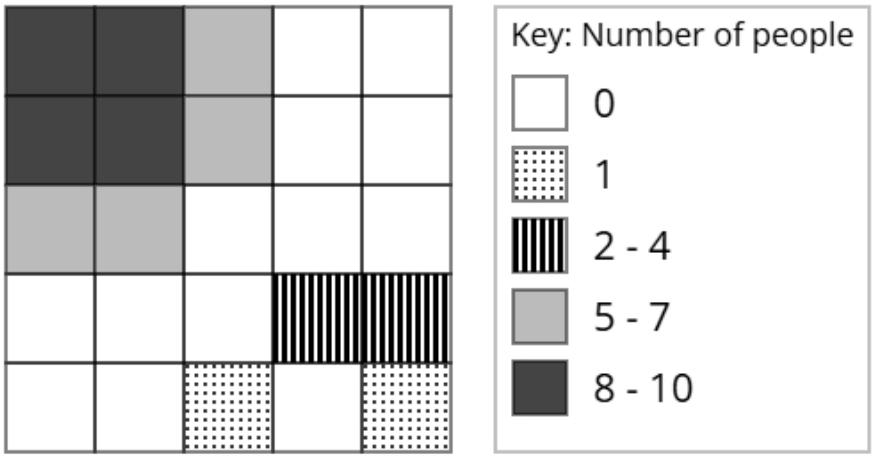
Find $P(X \text{ and } Y)$

(2 marks)

5 The choropleth map below represents a shopping centre that has been divided into 25 squares of equal area.

Dave has collected data about the popularity of different parts of the shopping centre.

The number of people recorded in each square on one Tuesday morning is shown.



(a) Calculate an estimate of the total number of people that were recorded on Tuesday.

(3 marks)

(b) Dave would like to open a smoothie cart in the shopping centre.

After analysing the data, he decides that he should open the smoothie cart in the corner of the shopping centre shown at the top left of the choropleth map.

Using the information in the choropleth map, assess the validity of Dave's conclusion.

(2 marks)

(c) Aisha argues that the method used by Dave to collect the data is not appropriate for reaching a reliable conclusion.

Assess whether Aisha's argument is correct and give a reason.

(1 mark)

6 A scientist is conducting an experiment to investigate how caffeine affects concentration.

She plans to use a matched pairs design.

Layla is one of the participants in the study.

As part of the experiment, she takes four concentration tests.

Each test has a different weighting.

The table below shows the weightings and Layla's scores for each test.

Test	Weighting	Score
A	1	15
B	2	20
C	3	18
D	4	25

(a) Explain the concept of matched pairs in an experimental design.

(2 marks)

(b) Calculate the weighted mean score for Layla's four tests.

(3 marks)

7 Aisha is trying to estimate the rabbits population in a park.

She first captures and tags 30 rabbits, then releases them.

One week later, she catches a second sample of 15 rabbits.

Using the Petersen capture-recapture method, she estimates the total number of rabbits to be 150.

(a) How many of the 150 rabbits in Aisha's second sample were tagged?

(2 marks)

(b) Discuss how reliable Aisha's estimate is by considering the assumptions required for using the Petersen capture-recapture method.

(3 marks)

8 A school headmaster wants to find out teachers have left the school premises during school hours without permission in the last 2 months.

Amelia suggests using the random response technique to ask the teachers whether they teachers have left the school grounds without permission.

(a) Explain why Amelia has suggested using the random response technique for this situation.

(1 mark)

(b) The school headmaster uses the random response question below:

> Think of a month of the year.
>
> If the month is after June, do not read the question and tick box A.
>
> If the month is June or before, answer the question truthfully.
>
> Have you left the school premises during school hours without permission?
>
> If yes, tick box A. If no, tick box B.
>
> A ☐ B ☐

State if this question is appropriate and give reasons for your answer.

(2 marks)

(c) The final questionnaire will be distributed to a sample of teachers.

The teachers are made up from teachers with no other role, middle leaders, and senior leaders.

They work either full-time or part-time.

The table shows how many teachers there are in each category

		Role		
		No Other Role	Middle Leader	Senior Leader
Employment status	Full-time	46	24	6
	Part-time	23	13	3

The school headmaster plans to take a stratified sample based on role and employment status and requires a minimum of 20 individuals from each stratum.

If the calculated sample size for a particular stratum is a decimal, he will round it to the nearest integer.

Determine the smallest total sample size that ensures at least 20 people are selected from each stratum.

(2 marks)

9 Jordan is investigating the profits made by two different shops, BrightWay Retail and CornerStone Retail.
Jordan has obtained the annual percentage profits made by BrightWay Retail for the years 2016 to 2020 and the annual percentage profits made by CornerStone Retail for the years 2017 to 2020.

The table below gives this information.

Year	Percentage profit (%)	
	BrightWay Retail	CornerStone Retail
2016	1.8	
2017	2.1	1.3
2018	2.4	1.9
2019	2.9	1.5
2020	3.4	6.5

Jordan concludes that the average annual percentage profit made by CornerStone Retail over the 4 years is greater than the average annual percentage profit made by BrightWay Retail over the 5 years.
By using appropriate geometric means, assess Jordan's conclusion.
You must show your working.

(5 marks)

10 Aisha has collected data about the heights, in cm, of basketball players in a school.

The table gives some of the percentiles of Aisha's data.

Percentile	Height (cm)
97.5th	199.5
80th	189.6
60th	184.6
40th	180.4
20th	175.4
5th	168.5
2.5th	165.5

(a) Find the 2.5th to 97.5th interpercentile range.

(1 mark)

_____ cm

(b) One of the basketball players from the sample is selected at random.

Find the probability that their height is between 168.5 cm and 180.4 cm.

(1 mark)

_____ %

(c) Give a reason why it is appropriate for Aisha to use the mean and the standard deviation to summarise this data

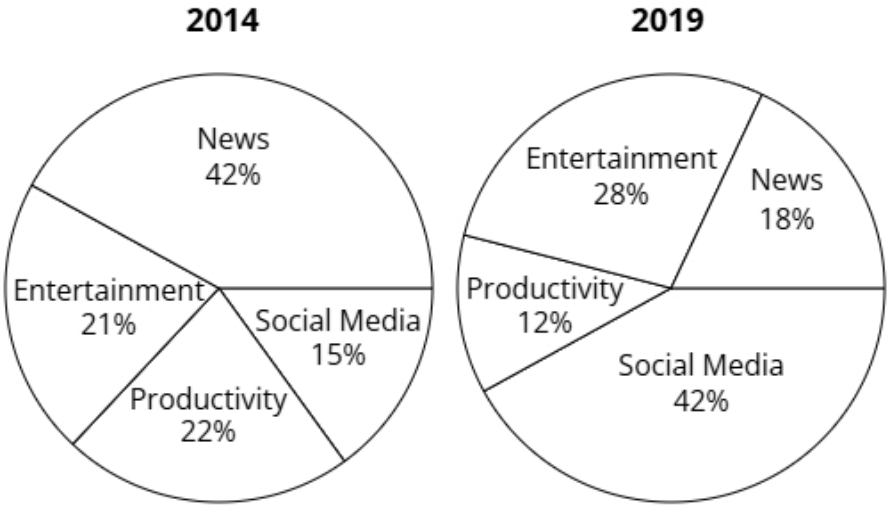
(1 mark)

(d) Aisha claims that the heights of the basketball players can be modelled using a normal distribution with mean 182.5 cm and standard deviation 8.5 cm.

Use the data to assess the validity of Aisha's claim.

(4 marks)

11 The pie charts show a country's mobile phone app downloads in 2014 and 2019.



In 2014 the total number of app downloads was 4577000 (nearest thousand).

In 2019 the total number of app downloads was 55345000 (nearest thousand).

Santi wants to use the totals to draw pie charts.

Explain, giving reasons, how Santi can use the totals to draw these pie charts.

(5 marks)

12 A company produces chocolate chip cookies.

The cookies have a target mass of 60 g.

The company uses quality assurance to monitor the mass of each cookie.

Samples of the cookies are taken from the production line at regular intervals and the mean mass of the cookies in each sample is found.

The sample means should be normally distributed with a mean of 60 g and a standard deviation of 1.2 g.

(a) Find the upper action limit for the sample means for the cookies.

(2 marks)

_____ g

(b) The upper action limit will be set closer to the target mass of 60 g.

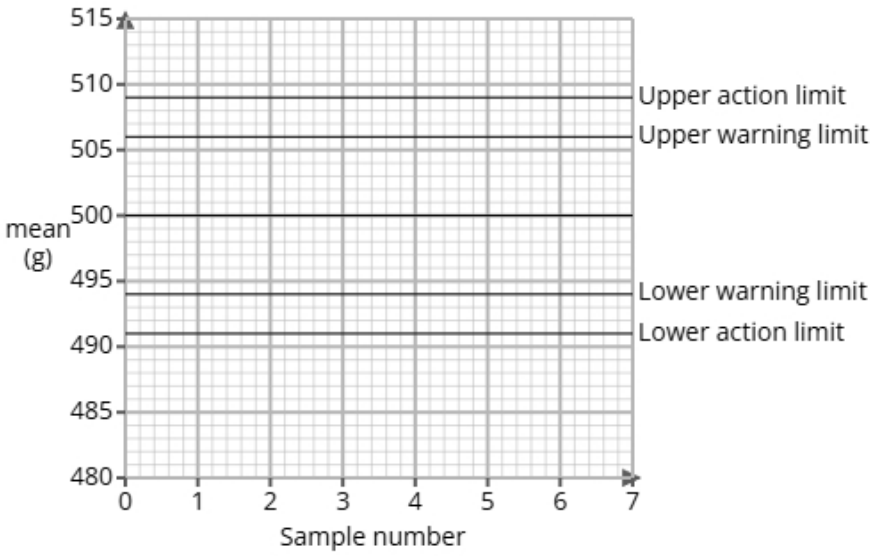
Describe the effect this will have on the frequency of production process stoppages.

(1 mark)

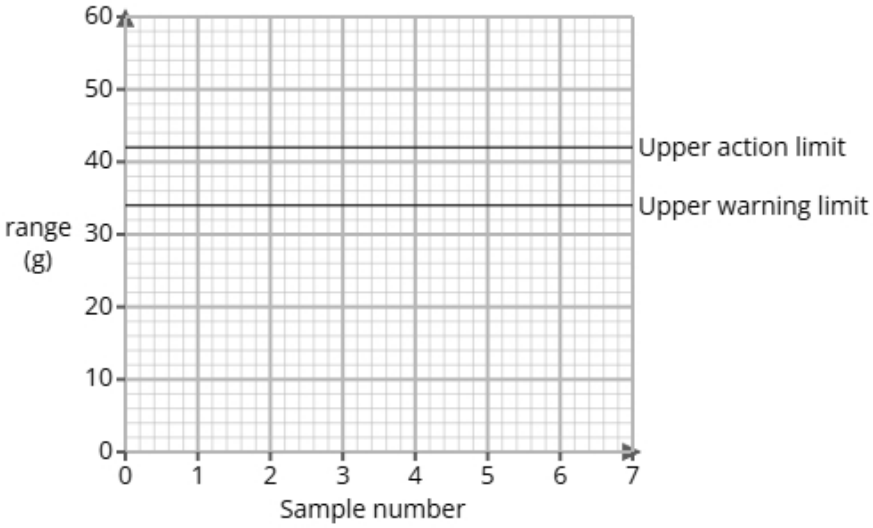
(c) The company also produces sponge cake and uses quality assurance to monitor the mass of each cake.

Here are the control charts for the sample means and for the sample ranges of the masses of the cakes.

Control chart for means



Control chart for ranges



A sample is taken and is found to have a mean of 496 g and a range of 31 g.

Use the sample mean and range to determine what action, if any, needs to be taken.

(2 marks)

13 A study took place in Japan to find if there was a relationship between screen time and sleep duration of teenagers.

The researchers found the equations of the regression lines for the relationship between screen time (x hours) and sleep duration (y hours) for male teenagers and female teenagers on school nights and weekend nights.

The table below gives the equations of the regression lines.

	school nights	weekend nights
male teenagers	$y = -0.4x + 7.5$	$y = -0.2x + 8.2$
female teenagers	$y = -0.3x + 7.8$	$y = -0.15x + 8.5$

(a) Compare the relationships between screen time and sleep duration in male and female teenagers. Include in your comparisons reference to whether it is a school night or weekend night.

(3 marks)

(b) The researchers would like to use a normal distribution as a model for the sleep duration of male teenagers on school nights.

i) Explain how they could check whether a normal distribution is a suitable model by drawing a histogram.

ii) Explain how they could check whether a normal distribution is a suitable model by calculating the averages and the standard deviation.

(3 marks)

14 A fair coin is tossed 4 times.

The number of heads obtained is recorded.

(a) Identify two conditions needed so that a binomial distribution is a suitable model for the number of heads recorded.

(2 marks)

(b) Calculate the probability, as a fraction, that all 4 of the coins land on heads.

(2 marks)

(c) Calculate the probability, as a fraction, that at least 2 of the coins land on heads.

(3 marks)