Statistics GCSE

Paper 2

2025

Edexcel Higher

Variant 3 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 At a community sports club, 70% of members are male and 30% are female.
Sophie and Marcus plan to conduct a feedback survey.

Sophie decides to use simple random sampling to select 100 participants.
She uses the club's registration list as a sampling frame, assigning a number to each participant.
She then generates 100 random numbers and selects her sample accordingly.

Marcus decides to use quota sampling to collect a sample of 100 participants.
He plans to approach people at the club until 70 male and 30 female have been interviewed.

(a) Give two reasons why Sophie's method may **not** produce a sample of 100 participants.

(2 marks)

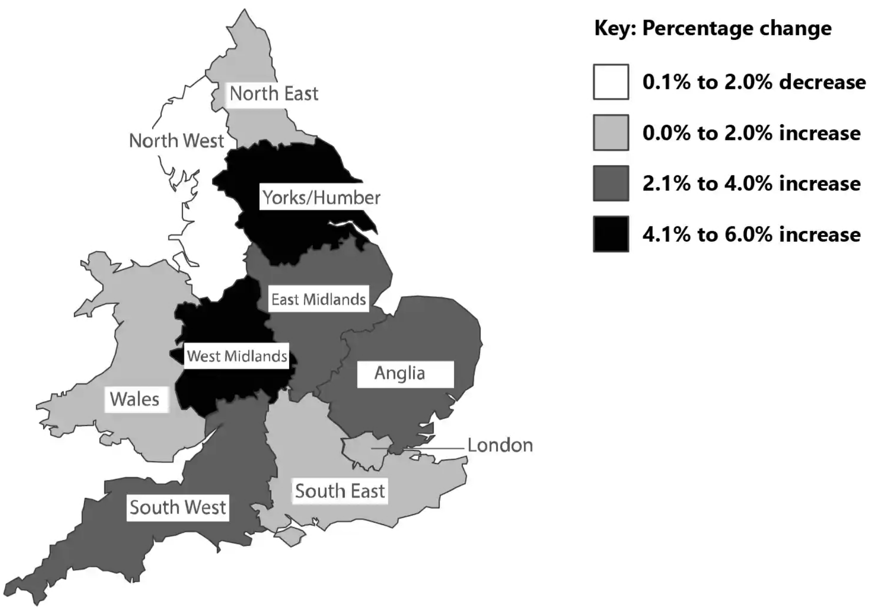
(b) Give **two** advantages of quota sampling.

(2 marks)

(c) Explain why the quota sample used by Marcus is not a random sample.

(1 mark)

2 The map below shows the percentage change in domestic tourist visits across different regions of England and Wales between 2012 and 2013.



(a) Write down the percentage change in domestic tourist visits between 2012 and 2013 in West Midlands.

(1 mark)

____

____

____

____

(b) There are 10 regions shown.
Find the number of regions that the domestic tourist visits **decreased**.

(1 mark)

____

(c) Erin states that domestic tourism in England and Wales increased overall between 2012 and 2013.
Explain why this may not be the case.

(1 mark)

____

____

____

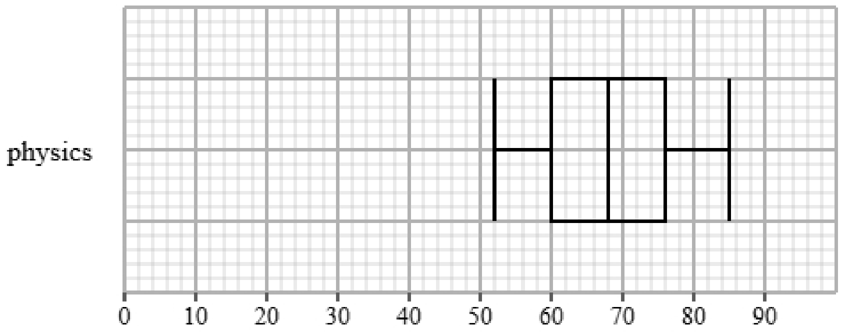
____

(d) State the statistical name for the type of map shown.

(1 mark)

____

3 Olivia collected the marks for physics and business students from a statistics exam.
Both groups took the same exam.
The box plot presents data on the marks for the physics students.

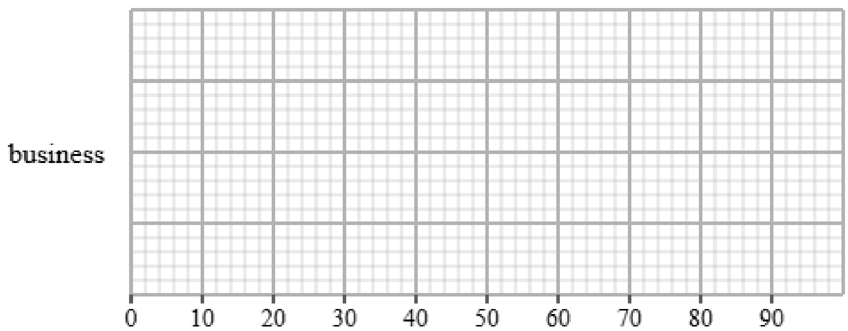


The table gives information about the marks for the business students.

Least tall	Lower quartile	Median	Upper quartile	Most tall
52	58	60	72	85

(a) Draw a box plot for the marks for the business students.

(2 marks)

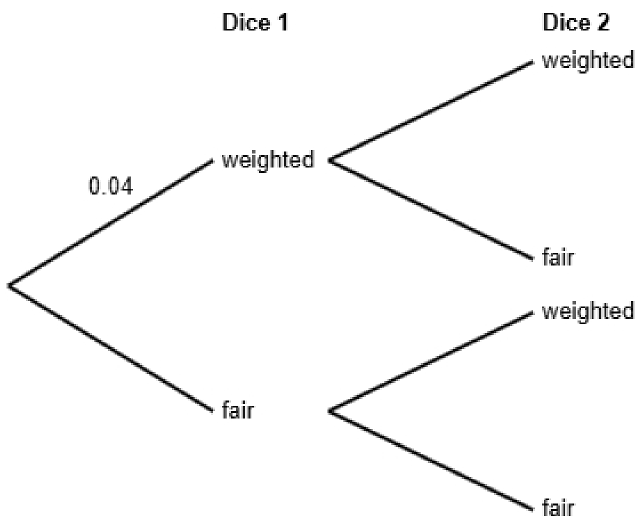


(b) Compare the two distributions of marks.

Give three comparisons and interpret one of these comparisons.

(4 marks)

4 A company found that 4% of its dice are slightly weighted and do not roll fairly.
The rest of the dice are perfectly balanced.
Jack picks two dice from a random batch.
He does not know if each die is weighted or fair.



(a) Complete the probability tree diagram.

(2 marks)

(b) Find the probability that both of Jack's dice are fair.

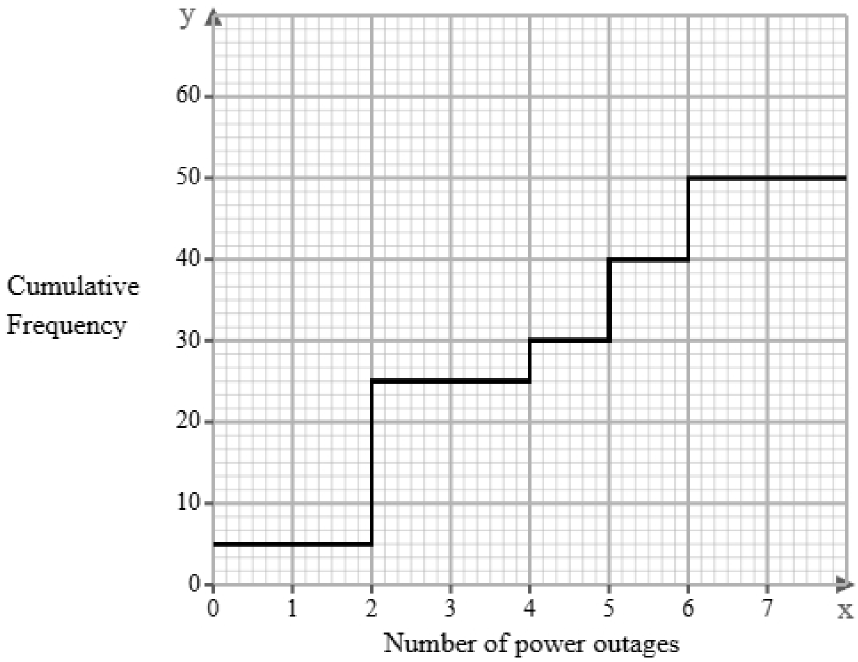
(2 marks)

____

(c) Jack states that the probability that exactly one dice is weighted is less than 8%
Find out whether or not Jack is correct.

(3 marks)

5 The cumulative frequency step polygon shows information about number of power outages recorded in a small town over 50 days.



(a) Give a reason why a cumulative frequency step polygon has been used to display this data.

(1 mark)

____

____

____

____

(b) Find the mode of the number of power outages.

(1 mark)

____

(c) Find the number of days where there were:
i) exactly 3 power outages.
ii) more than 3 power outages.

(3 marks)

i) Exactly 3 power outages: __________

ii) More than 3 power outages: __________

(d) In 40 days fewer than x power outages were recorded.
Find the value of x

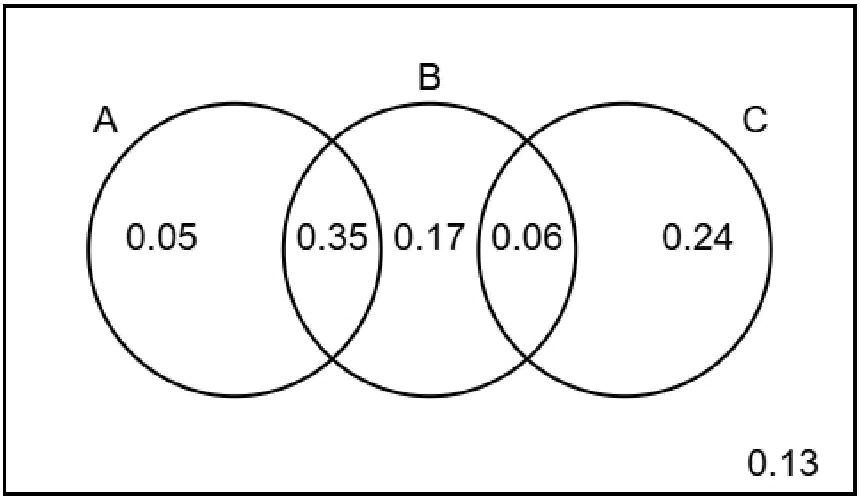
(1 mark)

____

(e) Nigel believes the interquartile range of number of power outages is 8.
Explain why the interquartile range for this data cannot be 8.

(1 mark)

6 The Venn diagram illustrates the probabilities associated with events A, B, and C.



(a) Identify the **two** events that are mutually exclusive, giving a reason for your answer.

(2 marks)

____

____

____

____

____

____

____

____

(b) Find P(B)

(1 mark)

____

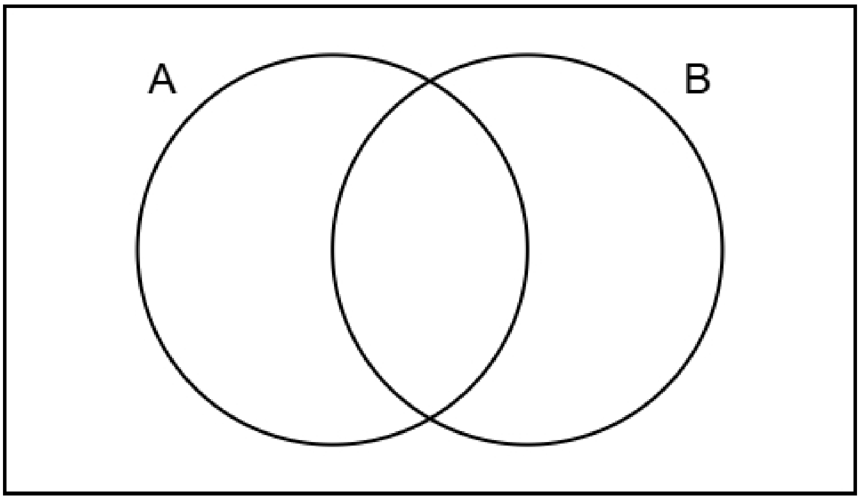
(c) Find P(A or C)

(2 marks)

____

(d) Complete the Venn diagram to show **only** the probabilities for events A and B.

(2 marks)



7 Aiden is researching customer ratings and critic ratings for 10 restaurants.

(a) Suggest a diagram that Aiden could draw to determine if there is a relationship between customer ratings and critic ratings for the 10 restaurants.

(1 mark)

(b) i) Calculate Spearman's rank correlation coefficient from the table.
ii) Interpret your answer in the context of Aiden's research. You will need to refer to the effects of any outliers.

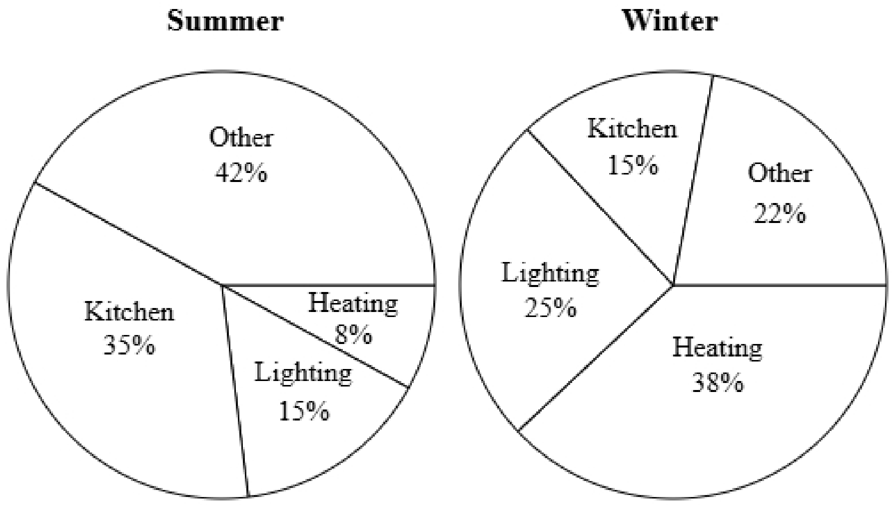
(5 marks)

Restaurant	Average Customer Rating (out of 5)	Customer Rank	Critic Rank	d	d^2
Bella's Bistro	4.8	1	3	-2	4
Caffe Vino	2.9	8	7	1	1
Downtown Grill	4.4	3	2	1	1
El Camino	3.2	7	8	-1	1
Fireside Diner	3.7	6	6		
Grand Oak Tavern	4.5	2	5		
Harbor View Restaurant	4.1	4	4		
Joy's Kitchen	1.8	10	9		
Olive Garden Grill	3.9	5	6		
The Golden Spoon	2.7	9	10		

(c) Discuss whether Aiden should have used Pearson's product moment correlation coefficient instead of Spearman's rank correlation coefficient to measure the correlation.

(3 marks)

8 The pie charts show the electricity usage in a house in Summer and Winter.



(a) Explain why the pie charts do not show a decrease in kitchen electricity usage in Winter compared to Summer.

(1 mark)

____

____

____

____

(b) The electricity usage in the kitchen in Summer was 224.735 kWh.
Find the electricity usage used for lighting in Summer.

(2 marks)

__________ kWh

(c) In Summer the total electricity usage was 640 kWh (nearest ten).
In Winter the total electricity usage was 830 kWh (nearest ten).

Liam wants to use the totals to draw pie charts.

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

(5 marks)

9 A company manufactures car tires.
The tires have a target pressure of 32 psi.
The company uses quality assurance to monitor the pressure of each tire.
Samples of the tires are taken from the production line at regular intervals and the mean pressure of the tires in each sample is found.
The sample means should be normally distributed with a mean of 32 psi and a standard deviation of 0.8 psi.

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

(2 marks)

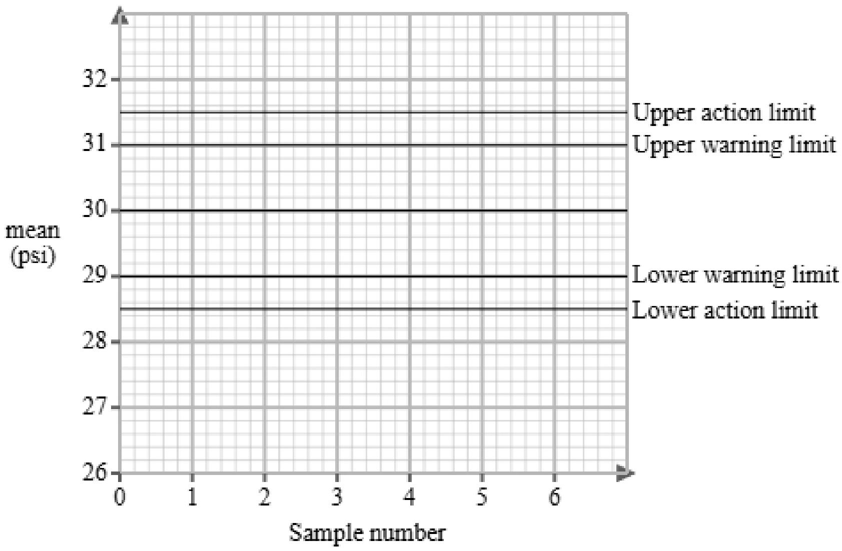
__________ psi

(b) The upper action limit will be set closer to the target pressure of 32 psi.
Describe the effect this will have on the frequency of production process stoppages.

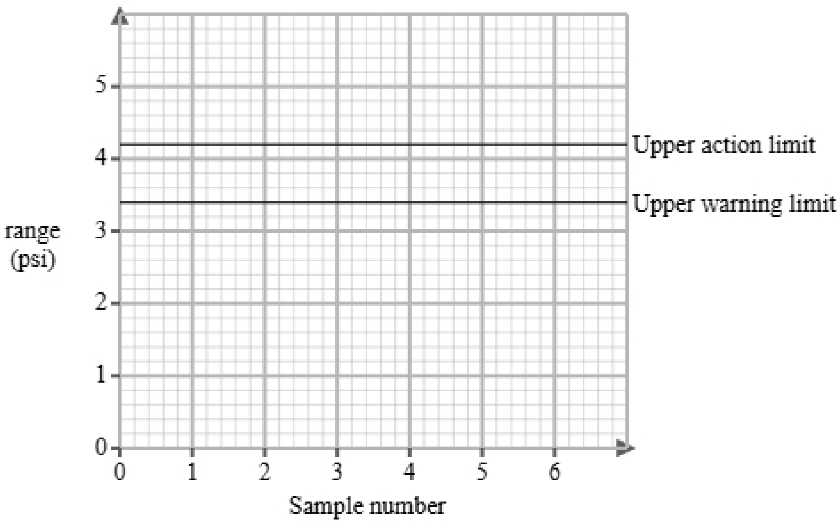
(1 mark)

(c) The company also manufactures bike tires and uses quality assurance to monitor the pressure of each tire.
Here are the control charts for the sample means and for the sample ranges of the pressures of the tires.

Control chart for means



Control chart for ranges



A sample is taken and is found to have a mean of 31.1 psi and a range of 3.2 psi.
Use the sample mean and range to determine what action, if any, needs to be taken.

(2 marks)

10 The heights of a group of ten-year-old girls have a mean of 137.5 cm and a standard deviation of 4.9 cm.

(a) Sophie is ten-year-old girl with a standardised score of 0.
Find Sophie's height.

(1 mark)

__________ cm

(b) Amelia and Ruby are both ten-year-old girls in the group.
Amelia's standardised score for height is 1.1 cm.
Ruby's standardised score for height is -0.3 cm.
Amelia is taller than Ruby.
How much taller is Amelia?

(3 marks)

__________ cm

(c) Fatima takes a sample of 3 ten-year-old girls from the group.
She wants to calculate the standardised score for the sample mean of their ages.

(i) Discuss the appropriateness of using 137.5 cm as the mean in the calculation of the standardised score,
(ii) Discuss the appropriateness of using 4.9 cm as the standard deviation in the calculation of the standardised score.

(4 marks)

11 A study took place in Brazil to find if there was a relationship between caffeine intake and reaction time of university students.
The researchers found the equations of the regression lines for the relationship between caffeine intake (x mg) and reaction time (y ms) for male students and female students before and after exams.

The table below gives the equations of the regression lines.

	before exams	after exams
male students	$y = -0.2x + 350$	$y = -0.25x + 340$
female students	$y = -0.1x + 330$	$y = -0.15x + 320$

(a) Interpret in context the figure -0.15 in the regression equation for female students after the exams.

(1 mark)

(b) Compare the relationships between caffeine intake and reaction time in male and female students. Include in your comparisons reference to whether it is before or after exams.

(3 marks)

(c) The researchers would like to use a normal distribution as a model for the reaction time of male students before exams.

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)

12 A factory produces lightbulbs, 5% of which are defective.
An inspector randomly selects 3 bulbs and the number of defective lightbulbs is recorded.

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

(2 marks)

(b) Calculate the probability, as a fraction, that all 3 of the lightbulbs are defective.

(2 marks)

(c) Calculate the probability, as a fraction, that at least 2 of the lightbulbs are defective.

(3 marks)