



General Certificate of Secondary Education
2025

Centre Number

--	--	--	--	--

Candidate Number

--	--	--	--	--

Statistics

Unit 1 (With Calculator)

Higher Tier



[GST12]

GST12

FRIDAY 13 JUNE, AFTERNOON

TIME

2 hours.

INSTRUCTIONS TO CANDIDATES

Write your Centre Number and Candidate Number in the spaces provided at the top of this page.

You must answer the questions in the spaces provided.

Do not write outside the boxed area on each page or on blank pages.

Complete in black ink only. **Do not write with a gel pen.**

Questions which require drawing can be completed using an HB pencil.

Any working **must** be clearly shown in the spaces provided. Marks may be awarded for partially correct solutions.

Answer **all ten** questions.

INFORMATION FOR CANDIDATES

The total mark for this paper is 100.

Figures in brackets printed down the right-hand side of pages indicate the marks awarded to each question or part question.

You should have a calculator, ruler, compasses and protractor.

The formula sheet is on page 2.



HIGHER TIER FORMULA SHEET

$$\text{Standard deviation} = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$$

Spearman's Rank Correlation Coefficient

$$r_s = 1 - \left(\frac{6 \sum d^2}{n(n^2 - 1)}\right)$$





BLANK PAGE

DO NOT WRITE ON THIS PAGE

(Questions start overleaf)

[Turn over

15894



28GST1203

Answer **all** questions

- 1 Len records the time, in minutes, that customers at his petrol station spend refuelling their vehicles.

The results for a sample of 16 customers are as follows:

3.0	1.8	2.8	3.4
2.4	2.9	2.1	3.1
3.9	0.1	1.7	2.3
1.6	4.2	2.7	2.0

- (a) Which of these results appears to be an outlier?

Explain your choice.

[2]

The outlier is now removed.

- (b) Using the remaining 15 values, draw a stem and leaf diagram in the space below.

[3]



(c) From your stem and leaf diagram opposite, find:

(i) the median time taken by these customers to refuel;

Answer _____ minutes [1]

(ii) the range of the times taken by these customers to refuel.

Answer _____ minutes [1]

Len obtained his data from automatic timers which are attached to each fuel pump.

(d) Write down the name of this method of data collection.

_____ [1]

[Turn over



2 Kenzie wants to gather opinions about introducing cycle lanes outside her school. She decides to survey cyclists who pass her school.

(a) Why might Kenzie's results be biased?

[1]

One of the questions Kenzie asks is:

What should be done to reduce the heavy traffic on this road?

(b) (i) Give a reason why this is not a suitable question.

[1]

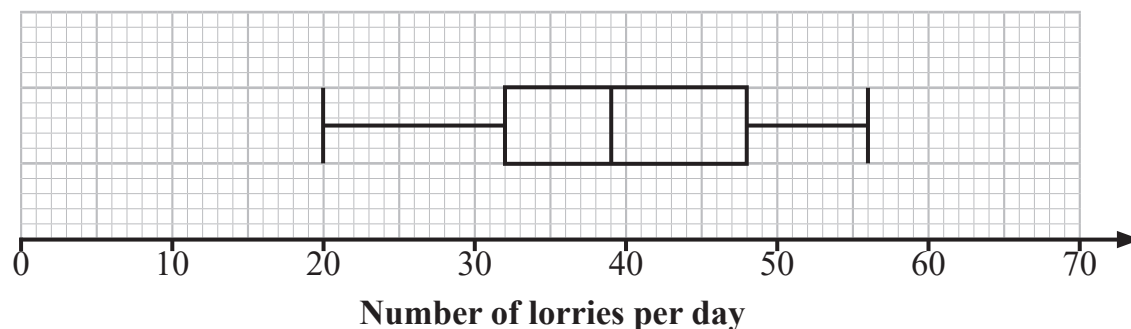
(ii) Give one example of qualitative data which Kenzie can record about the cyclists.

[1]



During one month, Noah recorded the number of lorries which passed by the same school each day.

His results are shown in the box plot below.



(c) Find the median number of lorries per day.

Answer _____ [1]



To improve safety around his school, the principal proposes three different plans.

He allows the teachers and pupils to vote on which plan they prefer.

The results are shown in the table below.

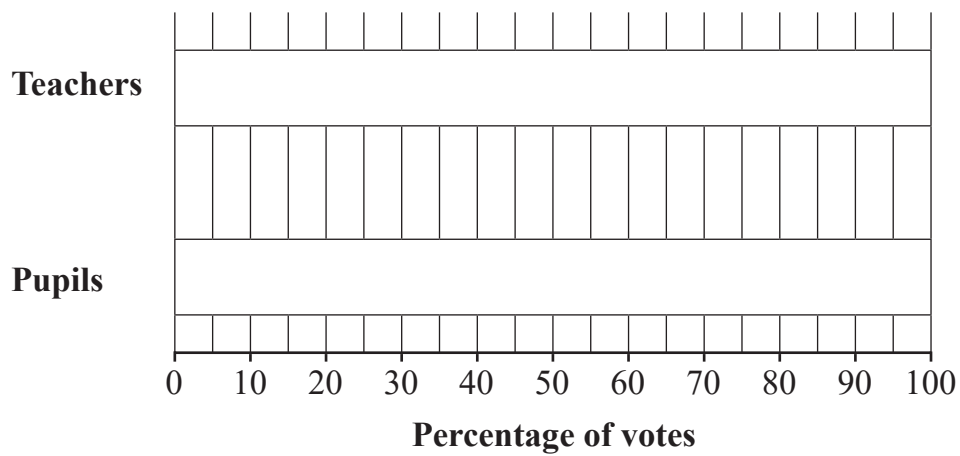
	Plan A	Plan B	Plan C
Teachers	20	16	44
Pupils	294	210	336

- (d) Explain why a compound percentage bar chart would be better than a dual bar chart to display this data.

[2]



(e) Use the data in the table opposite to complete the compound percentage bar chart below.



[3]

[Turn over



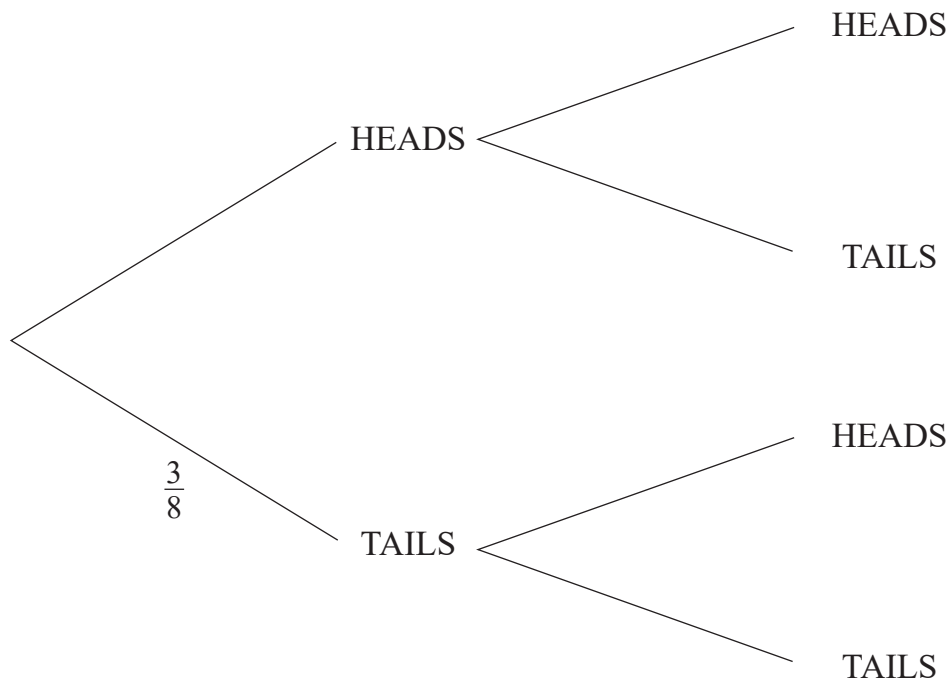
3 The probability that a certain coin shows TAILS when tossed is $\frac{3}{8}$

(a) How can you tell that this coin is biased?

[1]

Jo tosses the coin twice.

(b) Complete the missing probabilities on the tree diagram below.



[2]



(c) In the space below, list all four possible outcomes when Jo tosses this coin twice.

[2]

(d) Find the probability that the coin shows HEADS on both tosses.

Answer _____ [2]

(e) Calculate the probability that Jo obtains HEADS on one toss and TAILS on the other.

Answer _____ [2]

Jo tosses the coin 200 times.

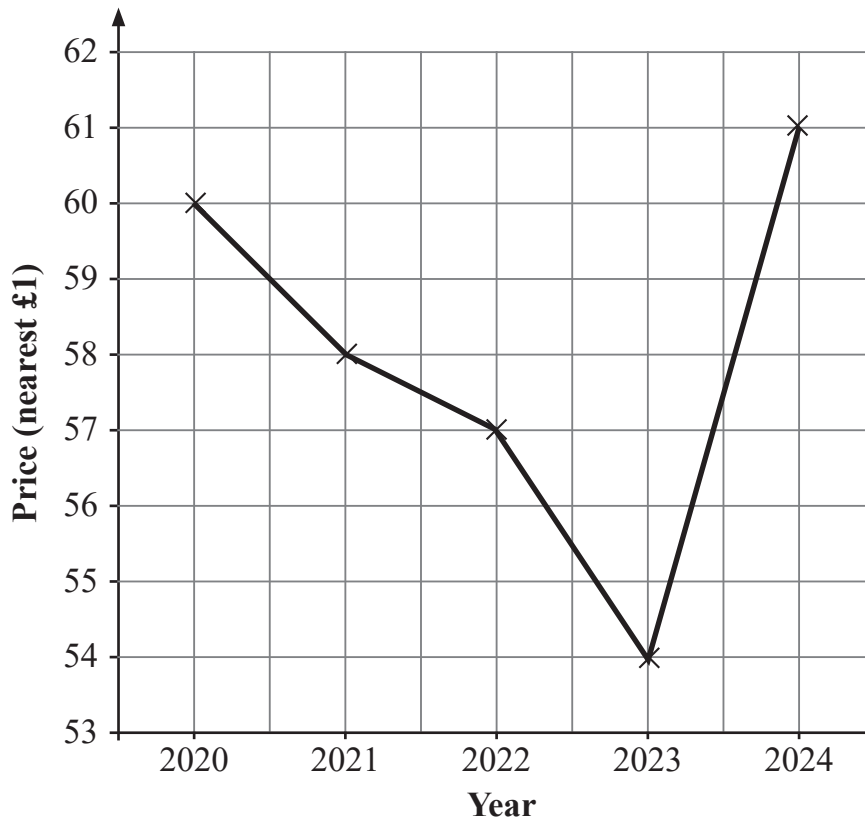
(f) How many times should Jo expect to obtain HEADS in these 200 tosses?

Answer _____ [2]

[Turn over



- 4 The time series graph below shows the price of a pair of shoes, to the nearest £1, between 2020 and 2024



- (a) Describe the trend in the price of this pair of shoes between 2020 and 2024

_____ [1]

- (b) (i) Between which two consecutive years was there the smallest change in price?

Answer _____ and _____ [1]

- (ii) How is this shown on the graph above?

_____ [1]



Using 2020 as the base year, Laura calculates index numbers for the years 2021 to 2024

(c) Write down the index number for 2020

Answer _____ [1]

Laura works out the index number for 2024 to be 101.7

(d) Explain the meaning of this value in relation to the price of this pair of shoes.

[3]

(e) Using 2020 as the base year, calculate the index number for 2023

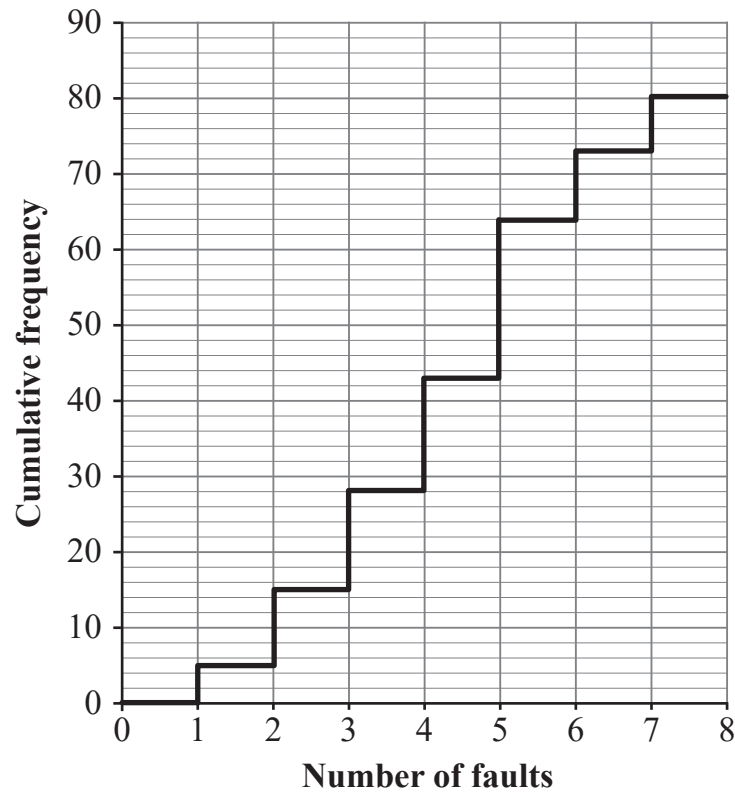
Answer _____ [3]

[Turn over



- 5 A quality control manager counts the number of faults in each car produced by a factory in one month.

The results are shown in the stepped cumulative frequency diagram below.



- (a) What type of data is the number of faults?

Circle the two correct words below.

Secondary

Bivariate

Primary

Qualitative

Continuous

Discrete

[2]



(b) Use the stepped cumulative frequency diagram opposite to find:

(i) the total number of cars checked;

Answer _____ [1]

(ii) the modal number of faults per car;

Answer _____ [1]

(iii) the median number of faults per car;

Answer _____ [2]

(iv) the number of cars with more than 5 faults.

Answer _____ [2]

(c) (i) Show that the range between the 1st and 9th deciles is 4

[2]

(ii) Explain what is meant by your answer to part (c)(i).

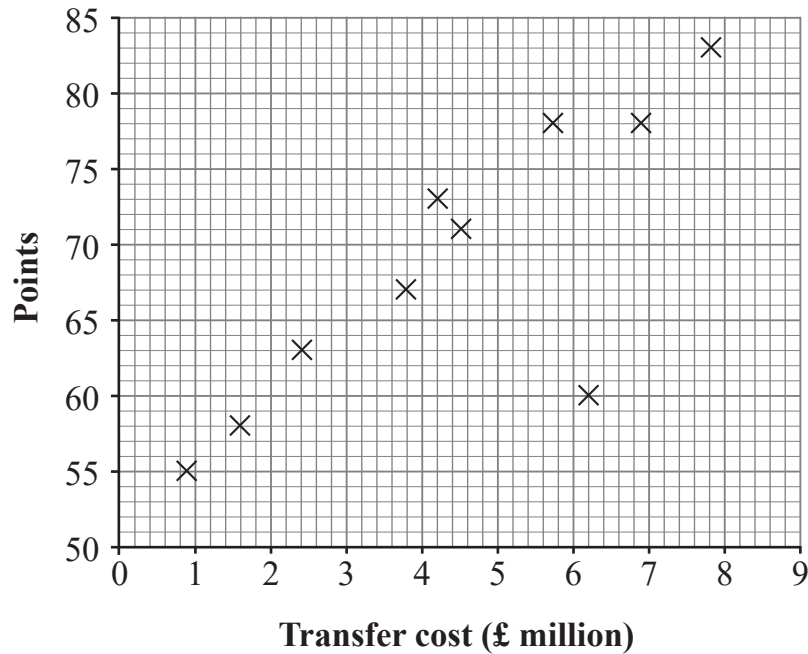
[1]

[Turn over



- 6 Jeffrey is investigating the relationship between the total transfer cost of a football team and the number of points achieved by that team in a season.

The results for 10 football teams are shown in the scatter diagram below.



- (a) Identify the explanatory variable in the scatter diagram above.

_____ [1]

- (b) Describe and interpret the correlation shown in the scatter diagram above.

_____ [2]



One of the managers was sacked for spending too much money on a poor quality team.

(c) Draw a circle around the point which is most likely to represent this team.

Give a reason for your choice.

[2]

With the point identified in part (c) removed, Jeffrey works out the equation of a line of best fit to be $y = 52.7 + 4.0x$

(d) Interpret, in context, the number 4.0 in this equation.

[2]

A team in the league had a transfer cost of £3.1 million.

(e) Use Jeffrey's equation to estimate the number of points achieved by this team in the season.

Answer _____ [2]

[Turn over



(f) Explain why Jeffrey should not use his equation to estimate the number of points achieved by a team with a transfer cost of £12 million.

[1]





BLANK PAGE

DO NOT WRITE ON THIS PAGE

(Questions continue overleaf)

[Turn over

15894



28GST1219

7 A machine is programmed to fill bags of cereal.

Every hour, five bags of cereal are sampled and the median mass of cereal for the sample is worked out.

(a) Explain why it would not be reasonable to check the mass of every bag of cereal.

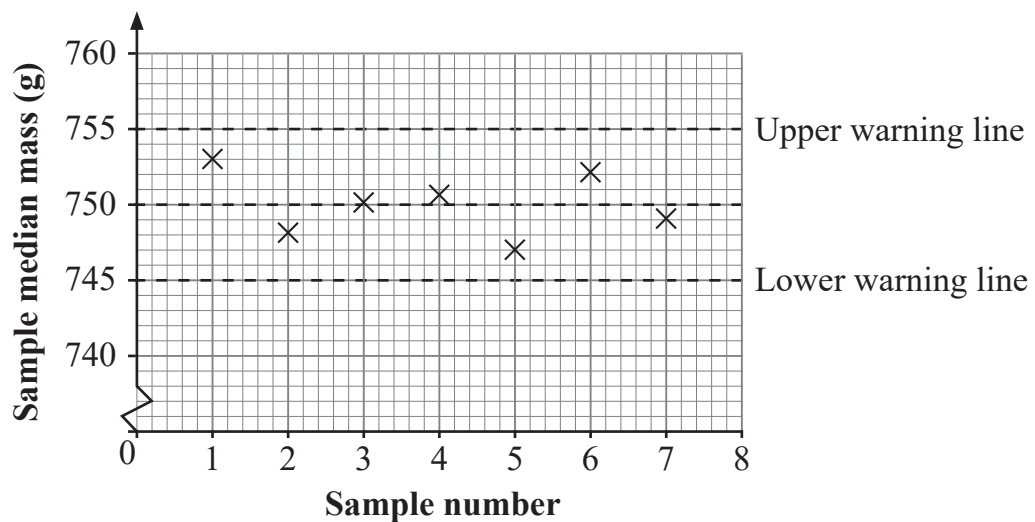
[1]

(b) Identify one consequence of underfilling the bags of cereal.

[1]

Simon monitors the median mass of each hourly sample.

The first seven sample medians are plotted on the control chart below.



(c) Write down the value of the target mass.

Answer _____ g [1]



- (d) On the control chart opposite, draw and label action lines at 8 g above the target mass and 8 g below the target mass. [2]

The masses of the bags of cereal in the eighth sample are as follows:

758 g 748 g 750 g 757 g 759 g

- (e) Work out the median of the eighth sample and plot it on the control chart. [2]

- (f) Explain how the value of the eighth sample median should be interpreted and what, if anything, needs to be done.

[2]

The masses of the bags of cereal in the ninth sample are as follows:

754 g 758 g 754 g 721 g 755 g

- (g) Use this data to explain why the median may not be suitable for working out the average mass of the samples of cereal.

[2]

[Turn over



8 Janet recorded the cost of car insurance for a random sample of 60 customers.

Her results are shown in the table below.

Cost, £	Frequency, f		
$250 < x \leq 300$	2		
$300 < x \leq 350$	13		
$350 < x \leq 400$	17		
$400 < x \leq 450$	25		
$450 < x \leq 500$	2		
$500 < x \leq 550$	1		

(a) (i) Calculate an estimate of the mean cost of car insurance.

You may use the blank columns in the table to help with your working.

Answer £ _____ [4]

(ii) Suggest one way Janet could make the estimate of the mean of her sample more reliable.

_____ [1]



Janet had a total of 176 customers.

(b) Estimate the total cost of car insurance for all 176 customers.

Answer £ _____ [2]

Janet considered decreasing the cost of car insurance for all her customers by 10%.

(c) For her random sample of 60 customers, what effect would this decrease have on:

(i) the estimated mean;

_____ [1]

(ii) the estimated standard deviation?

_____ [1]

[Turn over



- 9 The table below shows the number of guests staying in a hotel each quarter between 2022 and 2024

		Year		
		2022	2023	2024
Number of guests	Q1	340	344	352
	Q2	388	400	404
	Q3	440	452	464
	Q4	360	364	380

Rounded to the nearest whole number, where appropriate, the first seven 4-point moving averages for this data are as follows:

382 383 386 389 390 392 393

- (a) Calculate the next two 4-point moving averages.

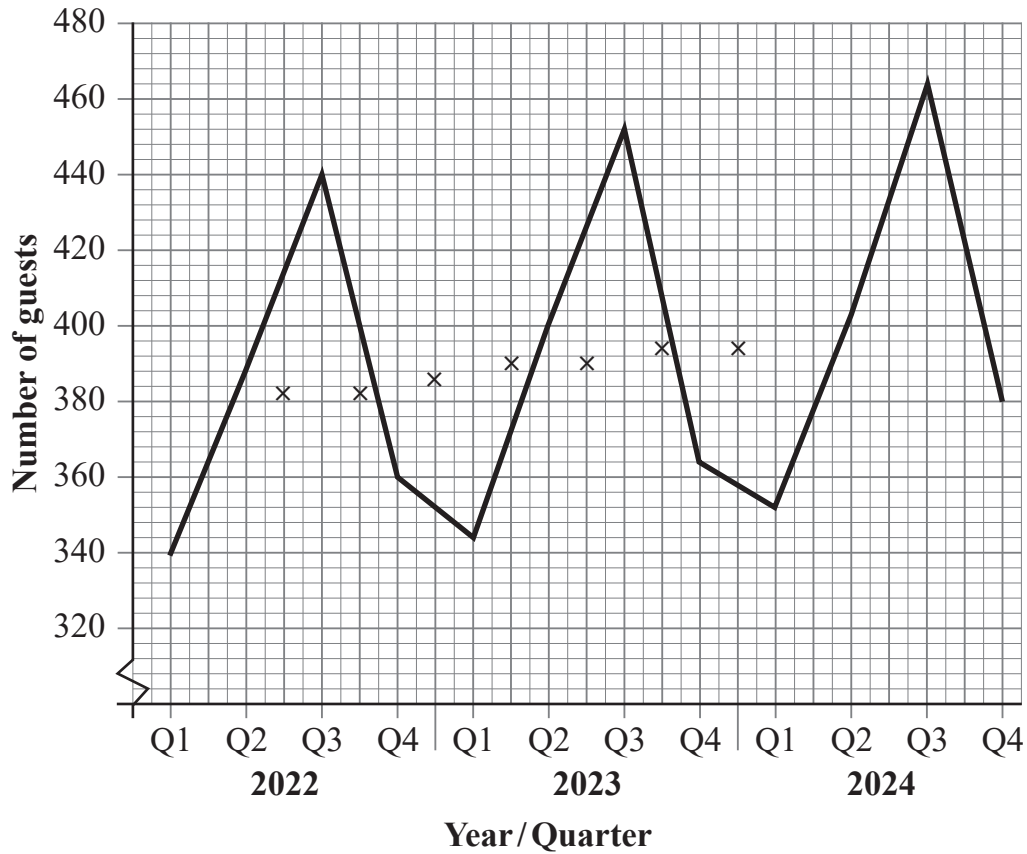
Answer _____ and _____ [2]

- (b) Explain why a 4-point moving average was used.

_____ [1]



The first seven moving averages have been plotted on the time series graph below.



(c) Plot the two moving averages you calculated in part (a) and draw a trend line. [3]

(d) Describe the trend in the number of guests in the hotel over the three-year period.

[1]

(e) Using a value taken from your trend line, estimate the number of guests expected in the hotel for Quarter 1 of 2025

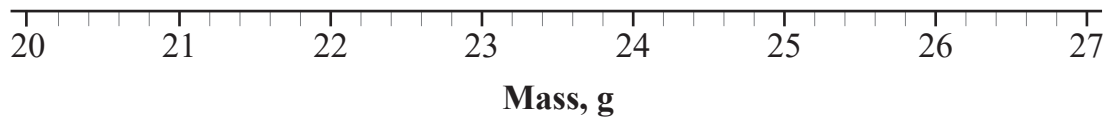
Answer _____ [3]

[Turn over



10 The masses of biscuits are normally distributed with a mean of 23.4 g and a standard deviation of 0.8 g.

(a) Sketch this distribution on the diagram below.



[2]

A biscuit was found to have a mass of 25.9 g.

(b) Use a suitable calculation to decide if a biscuit of mass 25.9 g is unusual for this distribution.

[4]



(c) Calculate an estimate of the percentage of biscuits which will have a mass between 21.8 g and 24.2 g.

Answer _____ % [5]

15894



28GST1227

THIS IS THE END OF THE QUESTION PAPER

DO NOT WRITE ON THIS PAGE

For Examiner's use only	
Question Number	Marks
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Total Marks	
--------------------	--

Examiner Number

Permission to reproduce all copyright material has been applied for.
In some cases, efforts to contact copyright holders may have been unsuccessful and CCEA will be happy to rectify any omissions of acknowledgement in future if notified.

15894



28GST1228