

SOLUTIONS

GCSE Maths: Statistics

Mark schemes for the Grade 4 starter questions

Questions	20
Total marks	61
Sub-topics covered	5
Papers drawn from	19
Years	2017 to 2025

The official Edexcel mark scheme for every question in the Statistics question booklet, in the same order.

Question numbers match: Question 7 here answers Question 7 there.

M marks are for method, A marks for accuracy, B marks stand alone.
'oe' means or equivalent; 'cao' means correct answer only.

S1 Sampling and Inference

Question 1

June 2019 · Paper 2 · Q3

3 marks

3	(i)	65	M1	for working with proportion eg $10 \div 30 \times 195 (=65)$	Condone use of 200 for 195
			A1	cao	
	(ii)		statement	C1	
Acceptable examples sample is representative (otherwise answer wrong) random sample (otherwise answer will be different) the 30 students are from the 195 (otherwise not accurate) 10 out of every 30 want to go to the Theme Park (otherwise answer will be different/wrong) there is no bias Not acceptable examples There would be more than 10 people who want to go to the Theme Park I rounded my answer					

S2 Tables, Charts and Diagrams

Question 2

June 2019 · Paper 3 · Q4

2 marks

4	statement	B2	Two different statements Acceptable eg should be joined with straight lines (not curve)/should use a ruler 1 st (quarter) not shown/plotted/labelled/not all quarters labelled does not show all 4 seasons 9.5 missing from vertical axes/not linear vertical (number) axis does not start at 0/the y axis starts at 6 the graph does not begin at 0, it starts at 6 it is not clear what 2, 3, 4 on the x-axis mean the scale of years doesn't make sense there is lack of clarity about what the numbers on the x axis represent graph is curved line Not acceptable eg no value plotted for 2 in 2016 it does not start at 0 (no reference to vertical axis)/missing 0 they should not have connected the dots like that the numbers on the x axis are repeated the numbers along the x axis 2, 3, 4 the years on the x axis have not been written properly does not follow a sequence it needs a discontinuity wiggle on the axis no title (B1) One statement eg from those above.)	Ignore additional statements provided no contradiction
	statement	(B1)		

S3 Diagrams for Grouped Data

Question 3

November 2019 · Paper 2 · Q1

2 marks

1	Two statements	C2	Two different statements Acceptable There is no 'frequency' label / y-axis is not labelled / no title for the y-axis The polygon should not be closed / have a line at the bottom / have first and last points connected (15, 6) has been plotted incorrectly / at (15, 8) / (The first point is at) 8 rather than 6 / First point is on an incorrect frequency Not acceptable There is no title / Points should be joined with a curve / x-axis doesn't start at 0 There is no label / The axes have not been labelled (x and y) The points haven't (all) been plotted correctly $10 < w \leq 20$ and $30 < w \leq 40$ have been plotted wrong The first point is plotted incorrectly, its at (15, 7) not (15, 6) The points have been joined up wrong / Points should not be joined in the shape of a triangle / They've connected all the points Done the midpoints rather than the numbers on the right side / The points are in the middle (C1 for one statement eg from those above)	Ignore additional statements provided no contradiction
---	----------------	----	--	--

Question 4

June 2022 · Paper 3 · Q5

2 marks

5	40 missing from frequency scale Incorrect point (50, 5)	C2	Two different statements Acceptable eg (50, 5) / the last point is incorrect the last point should be at (45,5) the last point plotted was placed incorrectly for his last point he has plotted by the end of the data and for the rest he has plotted by the middle he did not use the midpoint, he used 50 instead of 45 40 missing (from vertical axis) vertical scale is not linear the frequency doesn't increase in the same intervals the vertical axis is not right Not acceptable eg the last point should be at (40, 5) bottom of the polygon should be connected he didn't start the graph at the origin he did not draw a polygon he has plotted the first 4 points at midpoint (C1 One acceptable statement)	Ignore additional statements provided no contradiction
---	---	----	---	--

Question 5

June 2023 · Paper 1 · Q4

2 marks

4	Frequency polygon drawn	B2	for fully correct frequency polygon with points plotted at the midpoints	Joining must be with line segments Accept points plotted within half a small square Ignore any histogram drawn and any part of a frequency polygon outside range of first and last points plotted
	(2.5, 8), (7.5, 24) (12.5, 13) (17.5, 11) (22.5, 4)	(B1	for all points plotted correctly but not joined with line segments or points plotted at correct heights not at midpoints but consistently within each interval and joined with line segments or correct frequency polygon with one point incorrect or correct frequency polygon with first and last points joined directly)	for example, at 0, 5, 10,...or at 5, 10, 15,...

Question 6

June 2025 · Paper 3 · Q1

2 marks

1	Points plotted at (5, 12) (15, 25) (25, 10) (35, 8) (45, 5) and joined with line segments	B2	for correct plotting of all 5 points at midpoints and joining with line segments	Ignore any histogram drawn Ignore any part of the frequency polygon outside of the range of the first and last point plotted.
		(B1)	for correct plotting of all 5 points at midpoints joined with a curve or missing / no line segments or for a frequency polygon with one point incorrect or for a frequency polygon with first and last point joined directly or for plotting of all 5 points at the correct heights consistent within intervals (including end points) and joining with line segments	for example, at 0, 10, 20, 30, 40 or at 10, 20, 30, 40, 50

Question 7

November 2024 · Paper 3 · Q1

2 marks

1	Points plotted at (130, 12) (150, 28) (170, 30) (190, 22) (210, 8) and joined with line segments	B2	for correct plotting of 5 points and joining with line segments	Ignore any histogram drawn Ignore any part of the frequency polygon outside of the range of the first and last point plotted.
		(B1)	for points plotted correctly at midpoints or for a frequency polygon with one point incorrect or for a frequency polygon with first and last point joined directly or for joining the points with line segments at the correct heights consistent within intervals (including end points))	for example, at 120, 140, 160, 180, 200 or at 140, 160, 180, 200, 220

S4 Comparing Distributions

Question 8

November 2019 · Paper 1 · Q6

3 marks

6	No (supported)	P1	for process to find total weight of the 4 red bricks, eg. $5 \times 4 (= 20)$ or for process to find total weight of the 5 blue bricks eg. $9 \times 5 (= 45)$	May be seen next to statements 20 must be clearly referenced to the red bricks. $5 + 9 + 6 = 20$ gets no marks
		P1	for process to find total weight of all 10 bricks, eg. “20” + “45” + 6 (= 71)	
		C1	No with correct supporting evidence Acceptable examples No, it is 7.1 She is wrong, it is 0.1 more No, (the total weight is) 71 not 70 Not acceptable examples Yes No, it is 71	Candidates working in grams will need to give 7100 and 7000 for example as comparable figures

Question 9

November 2019 · Paper 3 · Q4

3 marks

4	18.6	M1	for finding 4 products within intervals (including end points)	<table><tr><td>Min fx</td><td>Max fx</td></tr><tr><td>5</td><td>10</td></tr><tr><td>20</td><td>30</td></tr><tr><td>105</td><td>140</td></tr><tr><td>160</td><td>200</td></tr></table>	Min fx	Max fx	5	10	20	30	105	140	160	200
		Min fx	Max fx											
		5	10											
		20	30											
105	140													
160	200													
M1	for $\Sigma“fx” \div (1+ 2 + 7 + 8)$ or $(7.5\times 1 +12.5\times 2 + 17.5\times 7 +22.5\times 8) \div (1+ 2 + 7 + 8)$ or (“7.5” + “25” + “122.5” + “180”) $\div$ “18” or “335” $\div$ “18”	$\Sigma“fx”$ must come from 4 products fx within intervals (including end points)												
A1	for 18.6(111...)													

Question 10

November 2020 · Paper 1 · Q7

3 marks

7	20 or 24 or 168	B1	for identification of the range of the girls (20) or the range (24) or the median (168) of the boys	
	Comparison	C2	for a correct comparison of medians and a correct comparison of ranges supported by correct figures eg the median height for girls (165) is less than the median height for boys (168) and the range for girls (20) is less than the range for boys (24) At least one comparison must be in context referring to height or quoting cm.	Simply quoting values for median, range is insufficient; they must be compared.
		(C1)	for a correct comparison of medians or a correct comparison of ranges that could fit their incorrect figure(s))	Context not necessary for C1

Question 11

November 2017 · Paper 1 · Q5

4 marks

5	(a)		365	M1	$\bar{f}x$ with x consistent within intervals eg 200×1 , 300×11 , 400×5 , 500×0 , 600×3 , if 200, 3300, 2000, 0, 1800 are seen without working then condone 1 error (dep) $\Sigma \bar{f}x \div \Sigma f$ eg “7300” $\div 20$
	(b)	Comment		M1 A1 C1	cao for comment about outliers affecting mean

Question 12

November 2023 · Paper 1 · Q4

4 marks

4	(a)	3.5	P1	for a process to find the total length of the 5 sticks, eg $4.2 \times 5 (= 21)$ or for forming an equation, eg $\frac{7+4x}{5} = 4.2$	
			P1	for complete process to find the mean eg “(21” – 7) $\div$ 4	
			A1	oe	
	(b)	Explanation	C1	for explanation Acceptable examples it reduced the mean my answer will be less the answer will be 1 it will be 2.5 less Not acceptable examples the mean will be more my answer will change it would decrease the lengths of the other sticks	If figures are given as part of the answer they must be correct, but can allow ft

Question 13

November 2023 · Paper 3 · Q3

4 marks

3	(a)	24.6	M1	for finding 5 products within intervals (including end points) with not more than one error, may be seen near table. eg 2×12.5 (= 25), 8×17.5 (= 140), 13×22.5 (=292.5), 21×27.5 (= 577.5), 6×32.5 (= 195) or for 1230	<table><tr><td>Min $\bar{f}x$</td><td>Max $\bar{f}x$</td></tr><tr><td>20</td><td>30</td></tr><tr><td>120</td><td>160</td></tr><tr><td>260</td><td>325</td></tr><tr><td>525</td><td>630</td></tr><tr><td>180</td><td>210</td></tr></table>	Min $\bar{f}x$	Max $\bar{f}x$	20	30	120	160	260	325	525	630	180	210
			Min $\bar{f}x$	Max $\bar{f}x$													
			20	30													
			120	160													
260	325																
525	630																
180	210																
M1	for $\Sigma \bar{f}x \div \Sigma f$ eg “25” + “140” + “292.5” + “577.5” + “195”) $\div$ “50” or “1230” $\div$ “50”	$\Sigma \bar{f}x$ must come from 5 products, $\bar{f}x$ within intervals (including end points)															
A1	for 24.6 oe																

Question 13 · continued

(b)	No, with reason	C1	for No and reason Acceptable examples No, the median is in the interval $25 < T \leq 30$ No, the median is in the group containing the 25(.5)th temperature No, she did not take into account frequency No, the frequencies are not the same for each group. Not acceptable examples No, the median is 27.5 No, the median is higher than 22.5 $25 < T \leq 30$ Yes, ...	Any incorrect statement as part of a correct response can be ignored unless it contradicts the statement.
-----	-----------------	----	---	---

Question 14

November 2024 · Paper 3 · Q6

4 marks

6	Comparison	B1	for correctly identifying the median of (School) A as 57	
		C1	(ft, dep on value for the median being stated) for making a correct comparison of <u>median</u> eg 'the median of (School) A is higher than the median of (School) B'	If median for school B is stated in the comparison it must be correct
		B1	for correctly identifying the range of (School) A as 49	
		C1	(ft, dep on value for the range being stated) for making a correct comparison of <u>range</u> eg 'the range of (School) A is higher than the range of (School) B'	If range for school B is stated in the comparison it must be correct

S6 Scatter Graphs and Correlation

Question 15

November 2018 · Paper 2 · Q2

2 marks

2	statements	C1	for lobf incorrect Acceptable examples lobf lobf does not suit all points/not a lobf lobf wrong since hits x axis/is inaccurate/should be amongst the crosses lobf goes through the origin/through one point Not acceptable examples no correlation/there is no title	
		C1	for height scale not linear Acceptable examples 150 missing Height not linear / Height numbers going up wrong Not acceptable examples 150 graph does not start at 140/graph does not start at 0 height should start at 170	

Question 16

June 2018 · Paper 3 · Q1

3 marks

1	(a)	negative	B1	cao	
	(b)	Explanation	C1	for a correct explanation, eg "not in line with the trend of the other points" "does not fit in with the correlation" "is far away from the other points or line of best fit"	
	(c)	Comment	C1	for an explanation eg "point would be outside of the range of the scatter diagram"	Ignore any description of a relationship and any reference to strength of correlation

Question 17

November 2024 · Paper 2 · Q2

3 marks

2	(a)	positive	C1	cao	Ignore any description of a relationship and any reference to strength of correlation
	(b)	lobf drawn	C1	for straight line passing between (140, 20) and (140, 22.5) and between (220, 30) and (220, 32.5)	
	(c)	26.5 – 29.5	C1	for answer in range 26.5 – 29.5 or ft single line with positive gradient	

Question 18

November 2021 · Paper 3 · Q1

4 marks

1	(a)	(100,18)	B1	cao	
	(b)	12.8 to 14.8	M1	for a method to read off eg line of best fit or line up from 370 or for a point on the grid at (370, y) where y lies between 12.8 and 14.8	
			A1	for an answer in the range 12.8 to 14.8	
	(c)	Decision and statement	C1	for decision and statement	
Acceptable examples No, as this point can be disregarded from the general trend No, ignore this point No, the correlation is positive No, because even with an outlier you can still have a negative or positive correlation. No, there is still a correlation. No, as you can use the rest of the data to determine a correlation. No, as outlier does not affect the majority No as a line of best fit can still be drawn No, it is an anomaly Not acceptable examples Yes, Outliers can be ignored [no decision] No, the outlier can be ignored so the correlation is negative No there are other things that can affect the test					

Question 19

November 2022 · Paper 2 · Q1

4 marks

1	(a)	(2, 1)	B1	cao	Accept negative correlation Ignore any comment about strength Any numbers used in the description must be within tolerance
	(b)	Description	C1	correct description, eg as the amount of rainfall decreases the number of hours of sunshine increases	
	(c)	3 to 4	M1	for a suitable line of best fit drawn, or for a point marked at (x, 7), or a horizontal line drawn from 7 across to (x, 7) where x is in the range 2.5 to 4	
			A1	answer in the range 3 to 4	

Question 20

June 2017 · Paper 1 · Q1

5 marks

1	(a)		10,19	B1	cao	
	(b)		positive	C1	positive (correlation)	
	(c)		12 to 13	M1	for an appropriate line of best fit drawn, or a point marked at (x, 16.4) or a horizontal line drawn from 16.4 across to (x, 16.4) where x is in the range 12 to 13	
				A1	hours given in the range 12 to 13	
	(d)		explanation	C1	(yes) e.g. as the majority of points for high temperature appear when there are more hours of sunshine (positive correlation)	