GCSE
STATISTICS
8382/1H

Higher Tier Paper 1

Mark scheme

June 2024

Version: 1.0 Final

2 4 6 G 8 3 8 2 / 1 H / M S

Mark schemes are prepared by the Lead Assessment Writer and considered, together with the relevant questions, by a panel of subject teachers. This mark scheme includes any amendments made at the standardisation events which all associates participate in and is the scheme which was used by them in this examination. The standardisation process ensures that the mark scheme covers the students' responses to questions and that every associate understands and applies it in the same correct way. As preparation for standardisation each associate analyses a number of students' scripts. Alternative answers not already covered by the mark scheme are discussed and legislated for. If, after the standardisation process, associates encounter unusual answers which have not been raised they are required to refer these to the Lead Examiner.

It must be stressed that a mark scheme is a working document, in many cases further developed and expanded on the basis of students' reactions to a particular paper. Assumptions about future mark schemes on the basis of one year's document should be avoided; whilst the guiding principles of assessment remain constant, details will change, depending on the content of a particular examination paper.

No student should be disadvantaged on the basis of their gender identity and/or how they refer to the gender identity of others in their exam responses.

A consistent use of 'they/them' as a singular and pronouns beyond 'she/her' or 'he/him' will be credited in exam responses in line with existing mark scheme criteria.

Further copies of this mark scheme are available from aqa.org.uk

Glossary for Mark Schemes

GCSE examinations are marked in such a way as to award positive achievement wherever possible. Thus, for GCSE Statistics papers, marks are awarded under various categories.

If a student uses a method which is not explicitly covered by the mark scheme the same principles of marking should be applied. Credit should be given to any valid methods. Examiners should seek advice from their senior examiner if in any doubt.

M	Method marks are awarded for a correct method which could lead to a correct answer.
A	Accuracy marks are awarded when following on from a correct method. It is not necessary to always see the method. This can be implied.
B	Marks awarded independent of method.
ft	Follow through marks. Marks awarded for correct working following a mistake in an earlier step.
SC	Special case. Marks awarded for a common misinterpretation which has some mathematical worth.
M dep	A method mark dependent on a previous method mark being awarded.
B dep	A mark that can only be awarded if a previous independent mark has been awarded.
oe	Or equivalent. Accept answers that are equivalent. eg accept 0.5 as well as $\frac{1}{2}$
[a, b]	Accept values between a and b inclusive.
[a, b)	Accept values $a \leq$ value $< b$
3.14…	Accept answers which begin 3.14 eg 3.14, 3.142, 3.1416
Use of brackets	It is not necessary to see the bracketed work to award the marks.

Examiners should consistently apply the following principles

Diagrams

Diagrams that have working on them should be treated like normal responses. If a diagram has been written on but the correct response is within the answer space, the work within the answer space should be marked. Working on diagrams that contradicts work within the answer space is not to be considered as choice but as working, and is not, therefore, penalised.

Responses which appear to come from incorrect methods

Whenever there is doubt as to whether a student has used an incorrect method to obtain an answer, as a general principle, the benefit of doubt must be given to the student. In cases where there is no doubt that the answer has come from incorrect working then the student should be penalised.

Questions which ask students to show working

Instructions on marking will be given but usually marks are not awarded to students who show no working.

Questions which do not ask students to show working

As a general principle, a correct response is awarded full marks.

Misread or miscopy

Students often copy values from a question incorrectly. If the examiner thinks that the student has made a genuine misread, then only the accuracy marks (A or B marks), up to a maximum of 2 marks are penalised. The method marks can still be awarded.

Further work

Once the correct answer has been seen, further working may be ignored unless it goes on to contradict the correct answer.

Choice

When a choice of answers and/or methods is given, mark each attempt. If both methods are valid then M marks can be awarded but any incorrect answer or method would result in marks being lost.

Work not replaced

Erased or crossed out work that is still legible should be marked.

Work replaced

Erased or crossed out work that has been replaced is not awarded marks.

Premature approximation

Rounding off too early can lead to inaccuracy in the final answer. This should be penalised by 1 mark unless instructed otherwise.

Continental notation

Accept a comma used instead of a decimal point (for example, in measurements or currency), provided that it is clear to the examiner that the student intended it to be a decimal point.

Q	Answer	Mark	Comments
1	$\frac{5}{50}$	B1	

Q	Answer	Mark	Comments
2	Ordinal	B1	

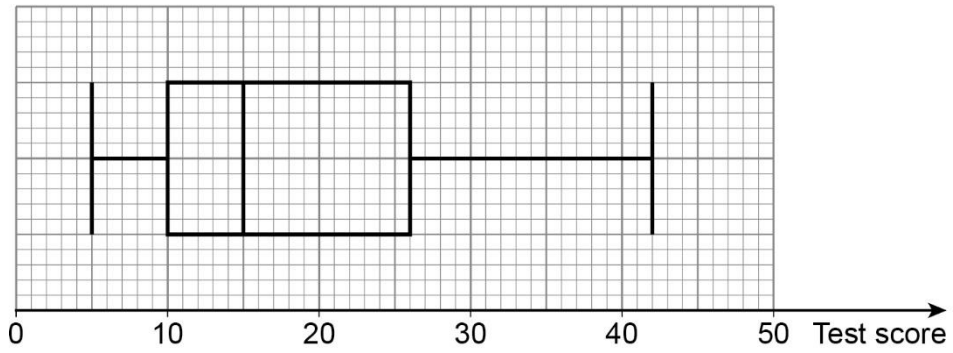
Q	Answer	Mark	Comments
3	Mode	B1	

Q	Answer	Mark	Comments
4	The range	B1	

Q	Answer	Mark	Comments
5(a)	Any suitable hypothesis, must be a prediction eg (students in school) A did better (in the French test than school B)	B1	oe
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Do not accept a question		
	Accept any tense		
	School A will do better (in the French test than school B) School B did better (in the French test than school A)	B1 B1	
	School A will have a higher mean/median/average	B1	
	I predict the schools will do similar	B1	
	Will school A do better (than school B)?	B0	

Q	Answer	Mark	Comments
5(b)	(All) Year 7 (students) in school A and school B	B1	oe
	Additional Guidance		
	All Year 7 students who sat the test in both schools All Year 7 students who sat the test		B1 B0
	All Year 7 students in the two schools All Year 7 students in two (different) schools		B1 B0
	Year 7 students that take French in both schools All Year 7 students (who study French)		B1 B0
	All students in both schools		B0

Q	Answer	Mark	Comments
5(c)	**Method A**		
	Correctly named as convenience (sampling)	B1	accept opportunity (sampling) or judgement (sampling)
	An advantage given about method A eg easy, quick, convenient	B1	oe
	Method B		
	Correctly named as random (sampling)	B1	comments about at random do not imply this mark
	An advantage given about method B eg every student has an equal chance of being picked or a wider range of students can be chosen or Sanjit cannot influence who is picked	B1	oe
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Condone random selection for Method B's name		
	Method B advantages: It's more representative Everyone has an equal chance of being picked More varied range of students More varied range of results/opinions It's not biased / It's fair (without further explanation)		 B1 B1 B1 B0 B0

Q	Answer	Mark	Comments
5(d)	(lower quartile =) 10	B1	may be seen in table
	(median =) 15	B1	may be seen in table
	(upper quartile =) 26	B1	may be seen in table
	Fully correct box plot with: • a suitable numbered linear scale with a label, eg Test score • values of their median, lower quartile and upper quartile correctly plotted and box drawn • minimum and maximum values correctly plotted and whiskers drawn to meet box	B3ft	ft their median and quartiles where 5 < their LQ < their median and their median < their UQ < 42 B2 2 bullet points satisfied B1 1 bullet point satisfied $\pm\frac{1}{2}$ small square
	Additional Guidance		
			
	Condone missing label		
	A box plot with no scale scores a maximum of B1B1B1B0		
	For a scale to be suitable it must cover all of the data values (5 to 42)		

Q	Answer	Mark	Comments
5(e)(i)	20	B1	

Q	Answer	Mark	Comments
5(e)(ii)	On average (school) B did better (than school A)	B1ft	oe ft their median from 5(d) and/or their median from 5(e)(i)
	Additional Guidance		
	If calculations are used, the outcomes must be correct		
	In general / Overall school B did better	B1	
	School A generally did worse	B1	
	School B did better	B0	
	School B's median was higher so they did better	B1	
	School B had a higher average/median	B0	

Q	Answer	Mark	Comments
5(f)(i)	14	B1	

Q	Answer	Mark	Comments
5f(ii)	(The scores for school) A were less consistent (than school B) or (the scores for) both schools had similar consistency	B1ft	oe ft their interquartile range from 5(d) and their interquartile range from 5(f)(i)
	Additional Guidance		
	If calculations are used, the outcomes must be correct		
	School B was more consistent	B1	
	B had a lower variation than A	B1	
	School B scores varied less	B1	
	School A's scores were more spread out than School B's	B1	
	The spread of school A and school B scores was similar	B1	
	The scores were similar	B0	
	The interquartile range of school A was bigger than the interquartile range of school B	B0	
	Any reference to comparing the statistical range	B0	

	Answer	Mark	Comments
6	Comment that the mode is not suitable/representative with a reason eg 70% of people did more exercise than the mode / only represents a small proportion or it was the lowest number in the list	B1	
	Comment that the median is suitable/representative with a reason eg represents the data well/best as it ignores extreme values	B1	
	Comment that the mean is not suitable/representative with a reason eg distorted or altered by 14 and/or 30 or correctly references the extremes/outliers or only 3 people did more exercise than the mean	B1	

Additional Guidance	
Ignore any non-contradictory or irrelevant statements	
Ignore any reference to the degree of suitability/unsuitability of the averages	
Any values stated must be correct	
Their decision may be implied eg unfair No decision or an incorrect decision made	B0
The mode ignores all the higher values so isn't representative	B1
The mode is unsuitable as the majority exercised more than once	B1
The mode ignores the higher values (14 and/or 30) so isn't representative	B0
The mode is unsuitable as it doesn't take all/other values into account	B0
The mode is unsuitable because there are values other than one	B0
Median is suitable as it (is the middle value and) does not include the extremes	B1
Median is suitable as it is the middle value	B0
The median is suitable because it doesn't take all values into account	B0

Q	Answer	Mark	Comments
7(a)	(Group) 5	B1	

Q	Answer	Mark	Comments
7(b)(i)	(More results/trials) will lead to a more reliable/fairer result	B1	oe
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	The first set of results may not be representative (so repeat it)		B1
	Uses more data does not mention reliability		B0
	Reference to accuracy		B0
	More representative		B0
	Comparing number of 1s in each group		B0

Q	Answer	Mark	Comments
7(b)(ii)	Ticks 'No' and gives a correct reason or ticks 'Cannot tell' and gives a correct reason	B1	eg if it was fair we would expect 20 1s but we got 11 / less than this (so it is not fair) or the probability of 1 should be 0.2 but the results indicate a probability of 0.11 (which is too low so it is not fair) eg 100 trials is too few to give a reliable estimate
	Additional Guidance		
	'No' or 'Cannot tell' could be contained in their response if none of the boxes indicated		
	Any values stated must be correct		
	Comparing the number of 1s in each group		B0

Q	Answer	Mark	Comments
8(a)	350 – 70 or 280 or 166 + 154 or 320 or (70 + 166 + 154) – 350 or (–)40	M1	oe may be seen on diagram
	166 – their 40 or 154 – their 40 or 126 or 114	M1dep	may be seen on diagram
	Fully correct solution	A1	
	Additional Guidance		
	Award the best mark from Venn diagram or in working for up to M2		

Q	Answer	Mark	Comments
8(b)(i)	$\frac{70}{350}$	B1	oe fraction, decimal or percentage
	Additional Guidance		
	Ignore an incorrect simplification/conversion if correct answer seen		

Q	Answer	Mark	Comments
8(b)(ii)	$\frac{126}{166}$	B2ft	oe fraction, decimal or percentage ft their 126 that must be less than 166 B1ft their 126 as numerator or their 166 as denominator if their $\frac{126}{166} < 1$
	Additional Guidance		
	Ignore an incorrect simplification/conversion if correct answer seen		

Q	Answer	Mark	Comments
9(a)	It is unlikely there is the same proportion/number of each group (in the population) or there may not be 30 records for a given age group or should take a stratified sample	B1	oe eg there are/may be fewer babies than seniors eg there may not be 30 babies
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Should sample them all / sample size too small		B0
	It may not be representative without a reason		B0

Q	Answer	Mark	Comments
9(b)	42 ÷ 20 (= 2.1) or any correct scaling on the vertical axis using frequency density eg one or more points correctly marked on the vertical axis	M1	may be seen on graph
	80 × their 0.2 (= 16) or 40 × their 0.65 (= 26)	M1dep	ft their 0.2 and their 0.65 from their scale 16 or 26 implies M1M1dep
	frequency = 16 for $0 \leqslant x < 80$ in table and frequency = 26 for $120 \leqslant x < 160$ in table	A1	
	26 ÷ 20 (= 1.3) or 10 ÷ 20 (= 0.5)	M1	implied by at least one bar of correct height $\pm\frac{1}{2}$ small square
	Complete correct histogram including a correct vertical scale	A1	$\pm\frac{1}{2}$ small square

Additional Guidance continues on the next page

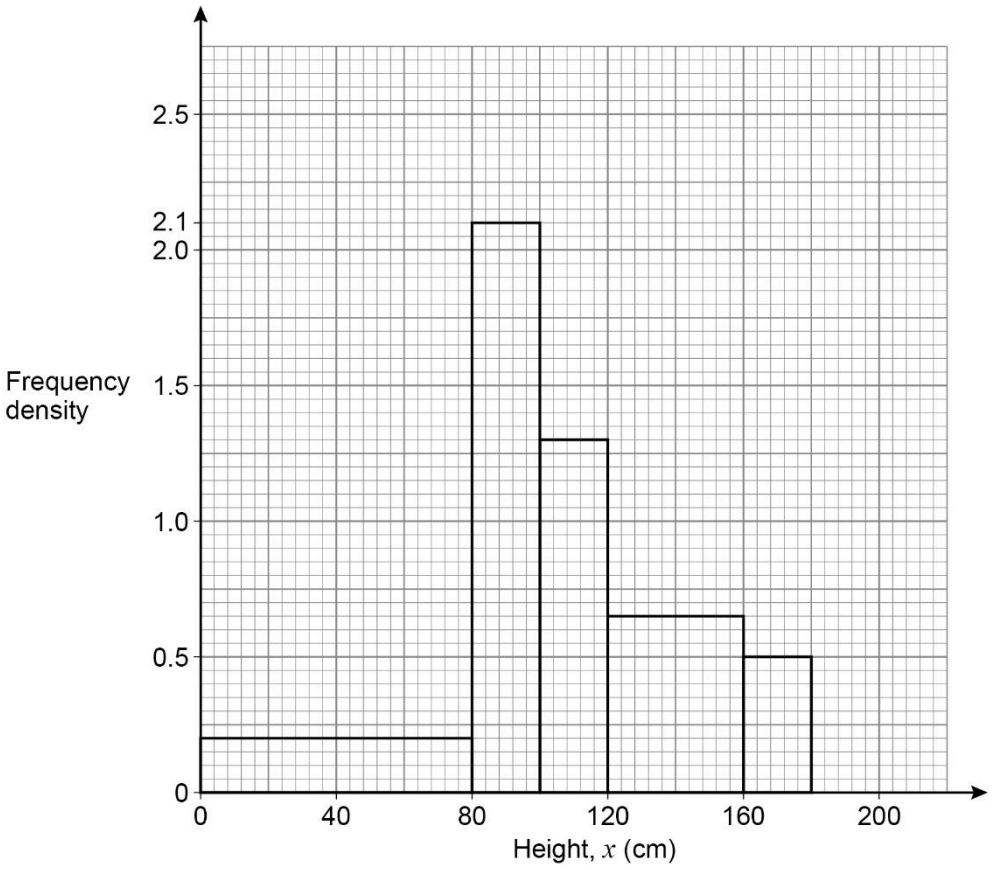
9(b) cont

Additional Guidance

Correctly completed table:

Height, x (cm)	**Frequency**
$0 < x \leqslant 80$	16
$80 < x \leqslant 100$	42
$100 < x \leqslant 120$	26
$120 < x \leqslant 160$	26
$160 < x \leqslant 180$	10

Correctly completed histogram



Use of frequencies instead of frequency density for histogram cannot score final two marks

Q	Answer	Mark	Comments
10(a)	Finds cumulative percentages up to at least the 201 – 350 group or finds cumulative percentages down to at least the 351 – 500 group	M1	oe calculations eg 41 + 15 + 19 + 14 or 89 or 41 + 15 + 19 + 14 + 6 or 95 or 5 + 6 = 11
	Ticks 'No' with justification and correct working seen	A1	
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Justification could be stating it is in group 351 – 500 or that 11 > 10 etc		
	States 90th percentile is group 351 – 500 without working		M0
	One misread value in calculation with calculation structure correct		M1A0
	'No' and 41 + 15 + 19 + 14 = 89, 89 < 90 'No' and 41 + 15 + 19 + 14 = 89		M1A1 M1A0

Q	Answer	Mark	Comments
10(b)	Comments on uncertainty of number of farms eg do not know the total amount of farms in either country or the graph shows only percentages of farms not numbers of farms	B1	
	Comments on uncertainty over number of cattle eg as the data is grouped we don't know the actual number of cattle in each farm or as the final group is open ended the largest English farms may be much larger (than the largest Scottish farms)	B1	
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	England being bigger scores B0 unless linked to number of farms		
	Secondary data could be unreliable/outdated		B0
	England might have lots of cattle that are not in farms		B0

Q	Answer	Mark	Comments
10(c)	Positive (Skew)	B1	accept comments about "Skewed to the right" or "skewed right" oe
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Ignore any reference to the 'degree of skewness'		
	Answers involving correlation		B0

Q	Answer	Mark	Comments
11(a)(i)	Correct identification of at least 3 relevant values for at least one country eg [9.25,9.75] (+) [9.75,10.25] (+) [9.75,10.25] (+) [12.75,13.25] or [15.75,16.25] (+) [15.75,16.25] (+) [15.25,15.75] (+) [15.75,16.25]	M1	may be marked on graph or seen in a calculation
	42.5 or 63.5	A1	
	Spain 42.5 and Turkey 63.5	A1	

Q	Answer	Mark	Comments
11(a)(ii)	(The percentage of the population that is below 40 is) lower in Spain (than in Turkey)	B1 ft	oe ft their 11(a)(i)
	Additional Guidance		
	Any values stated must be correct		
	Ignore any non-contradictory or irrelevant statements		
	Spain is only 42.5% and Turkey is 63.5% Spain is 42.5% and Turkey is 63.5%		B1 B0
	Turkey has a higher percentage so Turkey has a younger population Turkey has a younger population (than Spain) Turkey has more people under 40 (than Spain)		B1 B0 B0

Q	Answer	Mark	Comments
11(b)	There is a larger percentage/proportion of people in the older age groups for Spain (than in Turkey)	B1	oe
	Additional Guidance		
	Any values stated must be correct		
	Must refer to percentage/proportion		
	Comments referring to number of people		B0
	The bars are wider for Spain at the top of the diagram		B1
	In Spain a larger proportion are over 40		B1
	There is no 90+ bar for Turkey but there is for Spain There is no 90+ bar for Turkey		B1 B0

Q	Answer	Mark	Comments
12	$\frac{11.6 \times 58744\,600}{1000}$ (= 681 437.36) or 681 437	M1	oe
	their 681 437.36 × (1 – 0.096) (= 616 019.3734) or 616 019(.0…)	M1dep	oe
	$\frac{\text{their } 616019.3734 \times 1000}{59719700}$	M1dep	
	10.3 or better	A1	answer may be seen in table allow 10 with correct working seen SC2 [10.48, 10.5]

Q	Answer	Mark	Comments
13	$\frac{41.8-42.7}{4.4}$ or $\frac{68.8-70.7}{11.7}$	M1	
	Running = – 0.2(0…) or Cycling = – 0.16(2…)	A1	oe
	Running = – 0.2(0…) and Cycling = – 0.16(2…) and ticks Yes	A1	oe with both values in a comparable form
	Additional Guidance		
	Any supporting statements must be correct		

Q	Answer	Mark	Comments
14(a)	72 or 48 or 40(%) or 0.4(0)	M1	may be seen as part of a fraction or on diagram oe
	60(%)	A1	
	Additional Guidance		
	40(%) on answer line		M1A0

Q	Answer	Mark	Comments
14(b)(i)	Lower Quartile = 7 and Upper Quartile = 11	B1	
	Interquartile range = 4	B1 ft	ft their quartiles

Q	Answer	Mark	Comments
14(b)(ii)	their LQ – 1.5 × their IQR or their UQ + 1.5 × their IQR or 7 – 6 (=1) or 11 + 6 (=17)	M1	ft their quartiles and interquartile range if their IQR < 13
	One value correctly evaluated or 1 or 17 seen	A1ft	ft their quartiles and interquartile range if their IQR < 13
	Both values correctly evaluated or 1 and 17 seen and correct box ticked for their values	A1 ft	
	Additional Guidance		
	Ticking a box with no working shown		M0A0

Q	Answer	Mark	Comments
15(a)	The area/radius of the pie chart for (Country) A is larger (than the pie chart for Country B) (so there are more students finishing university in 2021 in Country A than in Country B)	B1	oe
	Additional Guidance		
	Any comment comparing angles of the sectors		B0
	(Country) A's pie chart is bigger (than the one for Country B) The pie chart is bigger		B1 B0

Q	Answer	Mark	Comments
15(b)	23(°)	B1	accept angles in range [21, 25] may be seen as part of a calculation or on diagram
	$\frac{4^2(\times\pi)}{3.6^2(\times\pi)}$ (= 1.2345679…) or $\frac{3.6^2(\times\pi)}{4^2(\times\pi)}$ (= 0.81)	M1	oe eg $\left(\frac{4}{3.6}\right)^2$
	618 000 × their 1.2(…) or 618 000 ÷ their 0.81 or $\frac{618000}{3.6^2} \times 4^2$ or 762 962.(9…) or 762963	M1dep	oe scale factor may be given as $\left(\frac{4}{3.6}\right)^2$ or $\left(\frac{3.6}{4}\right)^2$
	$\frac{\text{their } 23}{360} \times$ their 762 963	M1dep	
	Answer in the interval [44505, 53000]	A1	with no errors in working
	Additional Guidance		
	Correct answer with no angle and no working seen		B0M0M0M0A0

Q	Answer	Mark	Comments
16(a)	First sample is too small	B1	oe
	Coloured chalk dust will easily wash off or coloured chalk dust may transfer from one animal to another or coloured chalk dust will be obvious to other animals so they will not be able to mix into the population	B1	oe
	Time between samples is too small (because they won't have time to mix) or 50 is too many / will take too long to catch or previously captured animals may be wary and harder to recapture	B1	oe
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Answers may be given in any order		
	May leave the area / numbers might change		B0
	May not recapture any of the marked animals because there were only 5 May not recapture any of the marked animals		B1 B0
	It will take a long time because 50 is a large number of animals It will take a long time		B1 B0

Q	Answer	Mark	Comments
16(b)	$\frac{40}{P} = \frac{8}{35}$ or (P =) $\frac{40 \times 35}{8}$	M1	oe P = population (accept any other letter/symbol)
	175	A1	

Q	Answer	Mark	Comments
17(a)	Bell-shaped curve with a peak at 156 which must approach axis at each end	B1	symmetrical bell-shaped curve $\pm\frac{1}{2}$ square
	The height of the curve at 136 and 176 should be more than 1 small square	B1	
	Curve extends as far as 126 and 186 with a height of no more than one small square above the axis at these points	B1	condone the graph touching the axis at 126 and 186
	Additional Guidance		
	No curve		B0B0B0
	Ignore the curve before 126 and after 186		

Q	Answer	Marks	Comments
17(b)	Sight of one of the following values: $[0.\dot{6}, 0.683]$ or [0.95, 0.955] or $[0.\dot{3}, 0.342]$ or [0.475, 0.478]	M1	oe
	their $\frac{[0.\dot{6}, 0.683]}{2}$ + their $\frac{[0.95, 0.955]}{2}$ or $[0.\dot{3}, 0.342]$ + [0.475, 0.478]	M1dep	oe
	[0.808, 0.82]	A1	oe
	Additional Guidance		
	Answer of [0.808, 0.82] with no working		M1M1depA1
	Accept percentages written without percentage signs up to M1M1dep		

<table>
<tr><th>Q</th><th>Answer</th><th>Mark</th><th>Comments</th></tr>
<tr><td rowspan="11">17(c)</td><td>The dashed curve is much steeper/narrower (than the solid curve)
or
both curves have same mean</td><td>B1</td><td>oe
eg the results are more consistent

the word ‘mean’ must be used to gain this mark</td></tr>
<tr><td colspan="3">Additional Guidance</td></tr>
<tr><td colspan="2">Ignore any reference to standard deviation or median as the dashed line is not known to be a Normal distribution</td><td></td></tr>
<tr><td colspan="2">Peaks are at the same point
It has a higher peak</td><td>B0
B0</td></tr>
<tr><td colspan="2">It spikes in the middle
It goes above the solid line</td><td>B0
B0</td></tr>
<tr><td colspan="2">The data is more concentrated in the middle
The data is concentrated in the middle
It is less spread out
It isn’t very spread out
It goes over a shorter range
It only goes over a short range</td><td>B1
B0
B1
B0
B1
B0</td></tr>
</table>