



# **Mark Scheme**

## **Specimen Papers Set 2**

Pearson Edexcel GCSE (9 – 1)  
In Statistics (1ST0)  
Higher (Calculator) Paper 1H

## **Edexcel and BTEC Qualifications**

Edexcel and BTEC qualifications are awarded by Pearson, the UK's largest awarding body. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information visit our qualifications websites at [www.edexcel.com](http://www.edexcel.com) or [www.btec.co.uk](http://www.btec.co.uk). Alternatively, you can get in touch with us using the details on our contact us page at [www.edexcel.com/contactus](http://www.edexcel.com/contactus).

## **Pearson: helping people progress, everywhere**

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: [www.pearson.com/uk](http://www.pearson.com/uk)

## General marking guidance

These notes offer general guidance, but the specific notes for examiners appertaining to individual questions take precedence.

- 1** All candidates must receive the same treatment. Examiners must mark the last candidate in exactly the same way as they mark the first.

Where some judgement is required, mark schemes will provide the principles by which marks will be awarded; exemplification/indicative content will not be exhaustive. When examiners are in doubt regarding the application of the mark scheme to a candidate's response, the response should be sent to review.

- 2** All the marks on the mark scheme are designed to be awarded; mark schemes should be applied positively. Examiners should also be prepared to award zero marks if the candidate's response is not worthy of credit according to the mark scheme. If there is a wrong answer (or no answer) indicated on the answer line always check the working in the body of the script (and on any diagrams), and award any marks appropriate from the mark scheme.

**Questions where working is not required:** In general, the correct answer should be given full marks.

**Questions that specifically require working:** In general, candidates who do not show working on this type of question will get no marks – full details will be given in the mark scheme for each individual question.

- 3** **Crossed out work**

This should be marked **unless** the candidate has replaced it with an alternative response.

- 4** **Choice of method**

If there is a choice of methods shown, mark the method that leads to the answer given on the answer line.

If no answer appears on the answer line then mark both methods **as far as they are identical** and award these marks.

- 5** **Incorrect method**

If it is clear from the working that the "correct" answer has been obtained from incorrect working, award 0 marks.

**6 Follow through marks**

Follow through marks which involve a single stage calculation can be awarded without working as you can check the answer, but if ambiguous do not award.

Follow through marks which involve more than one stage of calculation can only be awarded on sight of the relevant working, even if it appears obvious that there is only one way you could get the answer given.

**7 Ignoring subsequent work**

It is appropriate to ignore subsequent work when the additional work does not change the answer in a way that is inappropriate for the question or its context. (eg an incorrectly cancelled fraction when the unsimplified fraction would gain full marks).

It is not appropriate to ignore subsequent work when the additional work essentially makes the answer incorrect (eg incorrect algebraic simplification).

**8 Probability**

Probability answers must be given as a fraction, percentage or decimal. If a candidate gives a decimal equivalent to a probability, this should be written to at least 2 decimal places (unless tenths).

Incorrect notation should lose the accuracy marks, but be awarded any implied method marks.

If a probability fraction is given then cancelled incorrectly, ignore the incorrectly cancelled answer.

**9 Range of answers**

Unless otherwise stated, when an answer is given as a range (eg 3.5 – 4.2) then this is inclusive of the end points (eg 3.5, 4.2) and all numbers within the range.

### Guidance on the use of abbreviations within this mark scheme

<b>M</b>	method mark awarded for a correct method or partial method
<b>A</b>	accuracy mark (awarded after a correct method; if no method or process is seen then full marks for the question are implied but see individual mark schemes for more details)
<b>B</b>	unconditional accuracy mark (no method needed)
<b>oe</b>	or equivalent
<b>cao</b>	correct answer only
<b>ft</b>	follow through (when appropriate as per mark scheme)
<b>sc</b>	special case
<b>dep</b>	dependent (on a previous mark)
<b>indep</b>	independent
<b>awrt</b>	answer which rounds to
<b>isw</b>	ignore subsequent working

Question	Answer	Additional guidance	Mark
<b>1</b> (a)	B2 Fully correct frequency polygon	B2 for all five points correctly plotted and joined with straight lines. Condone lines joining to 'x'-axis but not joining start to end. Otherwise: B1 for five points plotted correctly (not joined), OR for at least three points correct <b>and</b> joined with straight lines	<b>(2)</b>
(b)	<p>B1 Correct comparison of average eg Wilbur Farm has larger litters on average, OR Wilbur mode (10) &gt; Napoleon mode (8)</p> <p>B1 Correct comparison of dispersion eg Napoleon Farm has a greater spread of litter size, OR Napoleon range (6) &gt; Wilbur range (4)</p>	<p>1<sup>st</sup> B1 for a correct comparison of average. Must see a word implying average. (eg 'Wilbur Farm has larger litters' alone is B0)</p> <p>2<sup>nd</sup> B1 for a correct comparison of dispersion. May refer to spread/range/variation. For each mark, listing alone (eg mode/range) without comparison scores B0</p>	<b>(2)</b>

Question	Answer	Additional guidance	Mark
<b>2</b> (a)	B1 Each (student) has the same chance of selection, oe	B1 for a correct equivalent explanation of randomness	<b>(1)</b>
(b)	B1 All the students at her school	B1 for an answer indicating <b>all</b> the students, BUT an answer indicating population size (850) alone scores B0	<b>(1)</b>
(c)	<p>B1 One number (940) is out of range (so cannot be used)</p> <p>B1 One number (310) is repeated (so needs replacing)</p>	<p>1<sup>st</sup> B1 for a statement recognising that only numbers on the sampling frame can be used</p> <p>2<sup>nd</sup> B1 for recognising the repeated number has to be rejected.</p>	<b>(2)</b>
(d)	B2 Terri's method is not appropriate as the large population would make it impracticable	<p>B2 for a complete answer rejecting Terri's idea with an appropriate reason. eg it would take too long.</p> <p>OR</p> <p>B1 for an incomplete answer, eg correct reasoning without a conclusion, OR correct conclusion with an attempt at reasoning.</p>	<b>(2)</b>

Question	Answer	Additional guidance	Mark
<b>3</b> (a)	M1 $\frac{890}{781} \times 100$  A1 114	M1 for correct calculation for index number  A1 for an answer in the range 113.9 to 114.0	(2)
(b)	B1 ft eg both prices have increased OR male price has gone up by a greater percentage  B1 ft eg male price has gone up by 14%	B1 ft for a statement implying that both prices have increased B1 ft for giving correct percentage for at least one index number  Note: male prices have gone up by 1% more than female prices scores B2	(2)

Question	Answer	Additional guidance	Mark
<b>4</b> (a)	B1 for 0.7, 0.4 and 0.8 in correct positions		(1)
(b)	M1 $0.3 \times 0.6$ or “0.7” $\times$ “0.8” M1 $0.3 \times 0.6 +$ “0.7” $\times$ “0.8” A1 ft 0.74	1 <sup>st</sup> M1 for a correct product of (their) probabilities 2 <sup>nd</sup> M1 for complete method using their probabilities Allow ft provided probabilities are between 0 and 1	(3)
(c)	M1 $\frac{0.3 \times 0.6}{0.3 \times 0.6 + \text{“0.7”} \times \text{“0.8”}}$  A1 0.243...	M1 for correct method for conditional probability Allow ft provided probabilities are between 0 and 1  A1 for awrt 0.24	(2)

Question	Answer	Additional guidance	Mark
5 (a)(i)	M1 $\frac{23430 + 22880 + 27430 + 25710}{4}$ A1 24862.5	M1 for summing the correct four quarters and dividing by 4  A1 for 24900, 24860, 24863 or 24862.5	(2)
(ii)	B1 ft for correctly plotting their point at height '24862.5' between Q2 and Q3 2016		(1)
(b)	B1 for reference to the truncated vertical axis OR reference to the graph only showing houses sold over £40 000	B1 for a correct comment assessing why the graph is misleading	(1)
(c)	B1 for quarter 3		(1)
(d)	B1 Upwards oe  B1 The number of houses sold is increasing as time goes by	B1 for a correct description of the trend. Accept positive but positive correlation is B0 B1 for a contextualised interpretation	(2)
(e)	B1 For a complete interpretation of the gradient of the trend line in context eg <ul style="list-style-type: none"> <li>The number of houses sold is <b>increasing</b> at an average rate of 325 <b>per quarter</b></li> </ul>	B1 for a complete contextualised interpretation of the gradient of the trend line	(1)
(f)	B1 for correctly identifying a problem with extrapolation eg <ul style="list-style-type: none"> <li>the trend may not continue</li> </ul>	B1 for a correct comment on the problems of extrapolation	(1)



Question	Answer	Additional guidance	Mark
6 (a)	B1 for eg it is not practical to collect primary data in this case	B1 for correct justification for collecting secondary data	(1)
(b)	B1 for eg box plots do not allow you to see correlation	B1 for a correct comment on the appropriateness of box plots Accept the data is bivariate	(1)
(c)	<p><b>Diagrams</b> B2 for scatter graphs AND appropriate reason eg allows comparison of correlation OR if B2 not scored... B1 for scatter graphs with an attempt at a reason or for scatter graph and to identify correlation (no comparison)</p> <p><b>Calculations</b> B2 for reference to a comparing measures of correlation AND reason eg</p> <ul style="list-style-type: none"> <li>calculate Spearman rank correlation coefficients AND to compare the strength of the correlations</li> <li>use Pearson product moment correlation coefficients AND to compare the strength of the correlations</li> </ul> <p>OR if B2 not scored... B1 for reference to a measure of correlation for either male or female lions with an attempt at a reason OR for reference to correlation coefficients with an attempt at a reason</p> <p>B2 for reference to regression equations AND reason eg</p> <ul style="list-style-type: none"> <li>calculate the regression equations AND compare the increase in height at shoulder per 1kg increase in weight</li> </ul> <p>OR if B2 not scored... B1 for reference to regression equations with an attempt at a reason OR for reference to a regression equation for either male or female lions with an attempt at a reason.</p>	<p>B2 for assessing the appropriateness of statistical methods for presenting data</p> <p>B2B2 for assessing the appropriateness of statistical methods for processing/analysing data</p>	(6)

Question	Answer	Additional guidance	Mark
7 (a)	M1 for box with two whiskers AND at least two values plotted correctly A1 for all correct		(2)
(b)	<p>B1 for eg</p> <ul style="list-style-type: none"> <li>median reaction time for 30-39 year olds is greater than for 20-29 year olds</li> <li>median reaction times increase as age increases</li> </ul> <p>B1 for eg</p> <ul style="list-style-type: none"> <li>IQR for 40-49 year olds greater than IQR for 20-29 year olds</li> <li>30-39 year olds have the greatest IQR</li> <li>range for 30-39 year olds greater than range for 20-29 year olds</li> <li>range of reaction times increases as age increases</li> </ul> <p>B1 for eg</p> <ul style="list-style-type: none"> <li>20-29 year olds positively skewed, 30-39 year olds negatively skewed</li> </ul> <p>B1 for eg reaction times increase as you get older (on average) OR reaction times for the 20-29 year olds most consistent</p>	<p>B1 for a correct statistical statement comparing the medians</p> <p>B1 for a correct comparison of the IQRs or ranges</p> <p>B1 for a correct comparison of the skews Allow positive or negative skew for 40-49 year olds box plot</p> <p>B1 for a correct contextual interpretation comparing medians or IQR/ranges/skews</p>	(4)
(c)	<p>B2 for appropriate AND reference to minimising the effect of extraneous variables oe</p> <p>OR if B2 not scored B1 for appropriate with an attempt at a reason OR B1 for reference to minimising the effect of extraneous variables oe</p>	<p>B2 for assessing the appropriateness of the use of matched pairs to compare the reaction time groups</p> <p>OR B1 for an attempt at assessing the appropriateness of the use of matched pairs to compare the reaction time groups</p>	(2)
(d)	B1 for quantitative AND multivariate		(1)

Question	Answer	Additional guidance	Mark
<b>8</b> (a)	B1 for reference to respondents are more likely to be honest if answering the question	B1 for a correct justification of the appropriateness of the method	<b>(1)</b>
(b)	<p>M1 for <math>0.5 \times (426 + 354)(= 390)</math></p> <p>M1 for <math>\frac{426 - "390"}{"390"}</math></p> <p>OR for <math>2 \times (426 - "390")</math> AND <math>0.059 \times (426 + 354)</math></p> <p>OR for <math>2 \times ("390" - 354)</math> AND <math>0.059 \times (426 + 354)</math></p> <p>A1 for 0.092(307...)</p> <p>OR for 72 AND 46(.02)</p> <p>A1 for comparison of '0.092(307...)' and 0.059 with appropriate conclusion about effectiveness eg <math>0.092 &gt; 0.059</math> so the random response technique appears to be effective as more people admitted to avoiding tax</p> <p>OR for comparison of '72' AND '46.02' with appropriate conclusion about effectiveness</p> <p>B1 for eg we don't know the population used in the research</p>	<p>M1 for method to estimate the number of people who answered yes because they got 1, 2 or 3</p> <p>M1 for method to estimate the proportion of people who have avoided tax OR for method to find 5.9% of the total number of respondents</p> <p>A1 for a correct proportion OR for comparable values based on the proportions</p> <p>A1 for comparison of values with appropriate conclusion about effectiveness</p> <p>B1 for identifying a limitation of the conclusion</p>	<b>(5)</b>

Question	Answer	Additional guidance	Mark
9 (a)	M1 for $\frac{100}{n} = \frac{12}{60}$ oe A1 for 500		(2)
(b)	B2 for reliable/not reliable with a correct supporting reason eg reliable and samples are a good size/reliable and time interval between samples is not too long/not reliable and may catch greater proportion of injured or ill turtles so not random  OR if B2 not earned B1 for reliable/not reliable with an attempt at a reason OR for identifying a factor which would impact reliability without a conclusion	B2 for a correct comment assessing the reliability of the conclusion  OR if B2 not earned B1 for an incomplete attempt to assess the reliability of the conclusion	(2)
(c)	M1 for $\sqrt{\frac{256\ 510}{24} - (102.58)^2}$ A1 for 12.85(53...)	M1 for a correct calculation for standard deviation  A1 for 12.8-12.9	(2)
(d)	M1 for $88.3 - 3 \times 7.07$ or $88.3 + 3 \times 7.07$ M1 for $88.3 - 3 \times 7.07$ and $88.3 + 3 \times 7.07$ A1 for 67.09 and 109.51 A1 for eg $67.09 < 74$ and $106 < 109.51$ and 'yes' oe	M1 for a correct calculation for one outlier limit M1 for a correct calculation for both outlier limits A1 for both correct limits, accept 67.1 and 109.5 B1 for comparison of minimum value and lower limit for outliers AND comparison of maximum value and upper limit for outliers together with correct conclusion	(4)

Question	Answer	Additional guidance	Mark
10	B2 for Method C AND reference to dividing the total of the deadlifts by the total number of people OR B1 for Method C with attempt at reason	B2 for a complete assessment of the appropriate choice with reason OR B1 for an incomplete assessment of the appropriate choice	(2)

Question	Answer	Additional guidance	Mark
11 (a)	<p>M1 Test A: <math>\frac{16.3 - 14.4}{1.5}</math>, Test B: <math>\frac{21.6 - 19.8}{2.4}</math></p> <p>A1 Test A: 1.26(6...)</p> <p>A1 Test B: 0.75</p> <p>B1 Better/faster performance in Test B, relative to the other students, with a reason</p> <p>B1 Lower standardised score in Test B, oe</p>	<p>M1 for either correct calculation</p> <p>A1 for 1.26-1.27</p> <p>A1 for 0.75</p> <p>B1 for contextual interpretation of results. B0 if no reason.</p> <p>B1 for statistical reasoning, using standardised scores, to support conclusion.</p>	(5)
(b)	<p>M1 for reference to 68% or just over <math>\frac{2}{3}</math></p> <p>A1 for 34%</p>		(2)

Question	Answer	Additional guidance	Mark
12 (a)	<p>B2 for reference to (most) areas with high numbers of bees corresponding to areas with high concentrations of wild flowers AND reference to the area at the bottom right (A4-A6, B4-B6) having a high concentration of bees, but a low concentration of wild flowers</p> <p>OR if B2 not earned...</p> <p>B1 for reference to (most) areas with high numbers of bees corresponding to areas with high concentrations of wild flowers OR reference to the area at the bottom left (A4-A6, B4-B6) having a high concentration of bees, but a low concentration of wild flowers</p>	<p>B2 for a complete answer assessing the conclusion based on the diagram</p> <p>OR if B2 not earned...</p> <p>B1 for an incomplete answer assessing the conclusion based on the diagram</p>	(2)
(b)	<p>B1 eg when deciding you should consider the total population...</p> <p>B1 ...if the total population is different then you should use comparative pie charts OR ...if the total population is the same then you can use pie charts</p>	<p>B1B1 for a complete answer assessing the appropriateness of pie charts and comparative pie charts</p>	(2)
(c)	<p>M1 for <math>\frac{349}{236} = \frac{r^2}{5^2}</math></p> <p>A1 for 6.08(03...)</p>	<p>M1 for a correct equation connecting the two radii</p> <p>A1 for awrt 6.1</p>	(2)

Question	Answer	Additional guidance	Mark
<b>13</b> (a)	B1 for a scatter graph showing positive non-linear correlation		<b>(1)</b>
(b)	B1 for 'as the amount of money spent on ski equipment increases, time to complete the ski course decreases'	B0 if more than one box ticked	<b>(1)</b>
(c)	B1 for eg the PMCC value shows correlation, but not causation		<b>(1)</b>