



Mark Scheme

Specimen Papers Set 1

Pearson Edexcel GCSE (9 – 1)
In Statistics (1ST0)
Higher (Calculator) Paper 2H

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 or 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.

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

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.

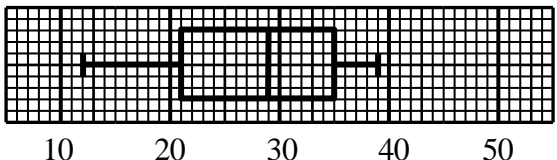
9 Range of answers

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

| | |
|--------------|---|
| M | method mark awarded for a correct method or partial method |
| A | 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 | unconditional accuracy mark (no method needed) |
| oe | or equivalent |
| cao | correct answer only |
| ft | follow through (when appropriate as per mark scheme) |
| sc | special case |
| dep | dependent (on a previous mark) |
| indep | independent |
| awrt | answer which rounds to |
| isw | ignore subsequent working |

| Question | Answer | Additional guidance | Mark |
|--------------|--|--|------------|
| 1 (a) | B1 eg those age 12 do not know which box to tick / could tick more than one option | B1 for answer recognising the overlap (poor 'reliability' as different boxes may be ticked by people of same age) Accept 'does not include under 10 or over 16' | (1) |
| (b) | B1 eg it is a leading question | B1 for equivalent wording recognising that it is a biased question | (1) |
| (c) | B1 eg it is an open question / answers will be difficult to handle | B1 for recognising it is open, or for explaining problems with using open questions Accept 'no time frame' | (1) |
| (d) | B1 eg members may not wish to answer (as questionnaire is not anonymous) | B1 for answer recognising that it is a sensitive question or that 'truthful' answers may not be obtained. (poor 'validity') | (1) |

| Question | Answer | Additional guidance | Mark |
|----------|---|--|------|
| 2 (a) |  <p style="text-align: center;">Memory test score</p> | <p>M1 for a box with two whiskers drawn with at least 3 values correct A1 fully correct</p> | (2) |
| (b) | <p>B1 for eg median for teenagers greater than median for parents</p> <p>B1 for eg IQR/range for parents greater than IQR for teenagers</p> <p>B1 for eg both negative skew</p> <p>depB1 for eg “The teenagers did better on the memory test than their parents” OR “The teenagers were more consistent on the memory test”</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</p> <p>depB1 for a correct contextual interpretation comparing medians or IQR/ranges (dependent upon relevant previous B1 being scored)</p> | (4) |
| (c) | <p>M1 56×0.75</p> <p>A1 42</p> | <p>M1 for identifying 75% or $\frac{3}{4}$ or 0.75</p> <p>A1 for 42</p> | (2) |
| 3 (a) | B1 eg Students who score highly in GCSE Mathematics also score highly in A-level Mathematics | B1 for a suitable hypothesis regarding relative performance in the two exams | (1) |
| (b) | <p>B1 GCSE is sat first</p> <p>It is plotted on the x-axis</p> | <p>B1 for an acceptable reason.</p> <p>Allow equivalent wording. Condone ‘horizontal’ axis.</p> | (1) |
| (c) | <p>B2 ft The scatter graph shows <u>positive correlation</u></p> <p>... which supports the hypothesis</p> | <p>B2 for a correct conclusion for their hypothesis (ft) and mention of positive correlation.(Otherwise B1 for identifying positive correlation)</p> | (2) |

| Question | Answer | Additional guidance | Mark |
|----------|--|---|------|
| (d) | B1 straight line with correct gradient B1 straight line through (578, 78) | 1 st B1 accept $0.6 < \text{gradient} < 0.8$ (not inclusive) If line does not extend at least from $x = 540$ to $x = 600$ then score max B1B0 | (2) |
| (e) | B1 ft eg for every extra mark at GCSE an extra 0.7 is scored for A-level | B1 ft for correct equivalent interpretation. Allow ft from their line. (eg 7 extra A-level marks for every extra 10 GCSE marks) | (1) |
| (f) | B1 (Will not be reliable because) 540 is outside the range of data / it is extrapolation | B1 for assessing the appropriateness of the method | (1) |
| 4 (a) | B1 (house) is not numeric data | B1 for any comment relating to a line graph being inappropriate for non-numeric data (eg reading values in between points is meaningless) | (1) |
| (b) | B1 diagram B / pie chart is best with attempt to give a reason B1 shows as a fraction / gives comparison with the total | B1 for correct choice with any attempt to justify choice B1 for equivalent description regarding comparison with the <u>total</u> . Accept converse statement about bar chart. | (2) |

| Question | Answer | Additional guidance | Mark |
|----------|---|--|------|
| 5 (a) | <p>M1 $\frac{2}{10}$ or 20% are tagged oe OR $\frac{25}{x} = \frac{2}{10}$ oe</p> <p>A1 $25/0.2$ or $x = \frac{25 \times 10}{2}$ oe</p> | <p>M1 for identifying proportion which are tagged</p> <p>A1 for a complete correct calculation that leads to 125</p> | (2) |
| (b) | <p>B1B1B1 for three correct points from</p> <ul style="list-style-type: none"> • Unreliable/poor (estimate) due to ... • Small sample • Sample not random / rabbits may not have mixed between samples • Long gap between samples / may have been a change in the population (eg may be births/deaths) • Tags may have come off between samples | <p>B1 for each of three correct comments from the list assessing the appropriateness of Richard's method to achieve reliable results. Accept equivalent statements. Allow each bullet point once only. Ignore excess statements if not contradictory.</p> | (3) |
| 6 (a) | <p>B1 Scatter (diagram) is most suitable graph</p> <p>B1 ... as data is bivariate/paired</p> <p>B1 line of best fit drawn (through the (double) mean point)</p> <p>B1 gradient (of line of best fit) is 0.8 / will tell her the relative rates of change in gross pay</p> <p>B1 points will be close to the line of best fit, OR there is strong correlation (so evidence is strong)</p> | <p>B1 for correct choice of graph</p> <p>B1 for reason explaining appropriateness for graph choice. Allow equivalent reference to bivariate data eg she was expecting to see correlation</p> <p>B1 for recognising that the double mean point should be used in drawing a best fit line</p> <p>B1 for correctly recognising that statistical reasoning for the comparison is based on gradient. Accept equivalent, eg steepness</p> <p>B1 for recognising how the high value of pmcc supports the validity of the conclusion OR that points will be close to a straight line</p> | (5) |

| Question | Answer | Additional guidance | Mark |
|--------------|---|--|------------|
| 8 (a) | B1 13 | B1 for correct evaluation of remaining counties | (1) |
| (b) | M1 $200 + \dots$ A1 $\dots \frac{(28-15)}{14} \times 200$ A1 = 386 awrt | M1 for identifying 2nd class and attempting to find median value within the class. A1 for use of correct fraction and class width (condone use of 28.5 for 28) A1 for value rounding to 386 (or awrt 393 if 28.5 used) | (3) |
| (c)(i) | B1 mean greater (than median) | B1 for equivalent statement recognising mean will be greater than median | (3) |
| (c)(ii) | B1 median more appropriate with attempt to give a reason B1 ...due to (positive) <u>skew</u> | B1 for correct choice AND attempt to justify B1 for stating data is skewed. This may be seen in either (i) or (ii) | |
| (d) | B1 Appropriate as data is grouped / continuous | B1 for conclusion recognising appropriateness with equivalent reason | (1) |
| (e) | M1 0.0125×400 A1 $r = 5$ $s = 8$ | M1 for attempting area of bar A1 cao | (2) |
| 9 (i) | B1 0.25 | Accept equivalent fraction or percentage | (1) |
| (ii) | M1 $\frac{100-93}{100-75}$ A1 = $\frac{7}{25}$ or 0.28 | M1 for attempt at ratio of absolute risks (condone 93/75 for M1) Accept as fraction decimal or percentage | (2) |

| Question | Answer | Additional guidance | Mark |
|----------|--|--|------|
| 10 (a) | M1 Adila: $\frac{60-75}{10}$ Brandon: $\frac{60-65}{4}$ A1A1 $= -1.5$ $= -1.25$ | M1 for attempt at standardised scores for both (condone $\pm(x - \text{mean})/\text{sd}$ for M1) A1 for each correct answer. (Accept equivalents) | (3) |
| (b) | B2 Adila had the best journey due to the lower standardised score | B2 for a complete conclusion in context making reference to standardised score values. Accept converse about B. (Otherwise B1 for correct comparison of standardised scores without conclusion) If B0 scored allow B1 for 'both quicker than normal' | (2) |
| (c) | B1 eg the journeys have different distributions / means / standard deviations | B1 for recognising that values are from different distributions making standardised scores appropriate | (1) |

| Question | Answer | Additional guidance | Mark |
|----------|---|---|------|
| 11 (a) | B1 eg correct, as highest <u>each</u> quarter 3 (or lowest <u>each</u> quarter 1) | B1 for an equivalent statement recognising a regular pattern | (1) |
| (b) | B1 eg 4-point as seasonal pattern repeats every four values / data given a quarterly values | B1 for equivalent statement recognising period of 4 values | (1) |
| (c) | Description of the key stages required: B1 ft Plot (4-point) moving averages (on the time series graph) B1 Draw a trend line (using the moving average points) B1 Calculate the mean seasonal effect for Quarter 1 B1 (Extend the trend line to) find the trend line value for Q1 2014 B1 Add the mean seasonal effect to the trend line value | Accept equivalent wording for each stage but words in bold are required. B1 ft for knowing that moving average points need plotting (allow follow through from (b)) B1 for knowing a trend line needs to be found B1 for interpreting how values for a particular quarter have an average variation from the trend (accept a description of the process) B1 for knowing to use extended trend line B1 for interpretation that trend line value + mean seasonal effect provides the required estimate | (5) |

| Question | Answer | Additional guidance | Mark |
|----------|---|--|------|
| 12 (a) | <p>M1 $23.7 = \frac{\text{number of births} \times 1000}{179300000}$</p> <p>A1 4249410</p> <p>M1 $\frac{3659289 \times 1000}{331900000}$</p> <p>A1 11.0</p> | <p>M1 for a correct equation for 1960 Allow use of 179.3 A1 4249410 or awrt 4250000 Accept 4.24941 million or awrt 4.25 million</p> <p>M1 for correct calculation for 2021 Allow use of 331.9 A1 Awrt 11.0 Accept 11</p> | (4) |
| (b) | <p>B1 for e.g.</p> <ul style="list-style-type: none"> we know the birth rate per 1000 not the number of births we don't know about rates of death / migration | <p>B1 for correct reason which indicates that the figures are relative to the population of the country OR for correct reason which indicates that there are other factors to consider in population change</p> | (1) |

| Question | Answer | Additional guidance | Mark |
|----------|--|--|------|
| 13 (a) | B1 binomial | Condone poor spelling if meaning is clear | (1) |
| (b) | B1 $9 \times \frac{2}{3}$ | B1 for knowledge of calculation for mean (accept $n \times p = 6$) May be in words. | (1) |
| (c) | <p>M1A1 $\left(\frac{2}{3}\right)^6 \times \left(\frac{1}{3}\right)^3 \times {}^9C_6$ ($= \frac{1792}{6561}$ or awrt 0.273)</p> <p>M1 “0.273” $\times 15$ A1 = 4</p> | <p>M1 for using $p^6 \times (1 - p)^3$ A1 for fully correct expression using 9C_6 or 9C_3 or 84 from Pascal’s triangle. (M1A1 implied by awrt 0.273 - may be from calculator)</p> <p>M1 for a probability $\times 15$ A1 for 4 (accept awrt 4.1)</p> | (4) |
| (d) | <p>B1 any two assumptions from:</p> <ul style="list-style-type: none"> • Independent flips (of the coins) • Probability of Heads is constant • Heads/Tails are the only possible outcomes • Fixed number of coins/flips (used each time) | <p>B1B1 allow equivalent wording but at least one must be in context Ignore excess statements if not contradictory.</p> | (2) |

