



Mark Scheme (Pre-Stand)

Summer 2025

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

Summer 2025

Question Paper Log Number P76387A

Publications Code 1ST0_2H_2506_MS

All the material in this publication is copyright

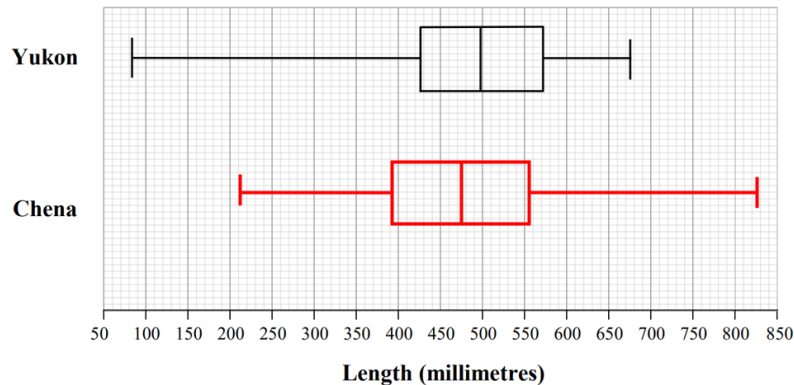
© Pearson Education Ltd 2025

Question number	Answer	Additional guidance	Mark
1 (a)	B1 Theoretical (probability)	B1 Condone misspelling	(1)
(b)	B1 B1 for any two of the following • Elijah has assumed the coin is fair • As the number of flips increases the probability of a heads will tend to the theoretical probability • Peter has (calculated) the experimental probability • The observed outcomes seem to match expected frequencies (assuming the coin is fair)		(2)

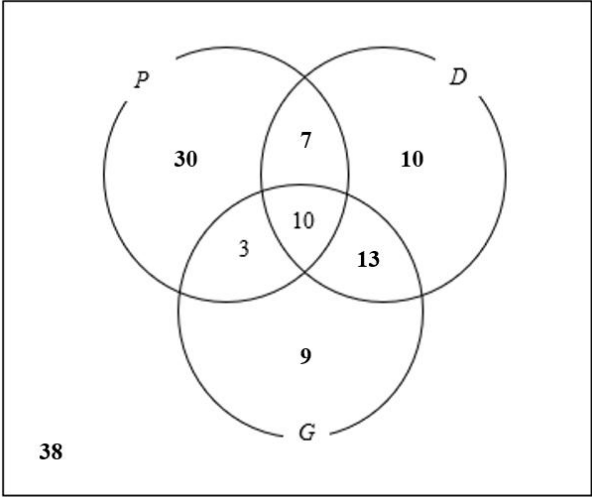
Question number	Answer	Additional guidance	Mark	Type
2	B1 e.g. - the hypothesis suggested is a question not a statement (so is inappropriate). B1 B1 B1 B1 for comments on data collection Sampling method - no indication of how (Elise) will sample the skulls from the museums - the sample size per mammal/species is small. - males and females for each species often differ in size so it is sensible to include both - finding/collecting all the data required will be difficult/time consuming. - Elise may find it difficult to access all the data required Issues with the collected data - there may be errors in recording of measurements/the data may not be reliable. - the data may need ‘cleaning’ before it can be processed. - there may be missing data Use of secondary data - do not know how the internet data was collected - the sources of the secondary data should be included / we do not know which museums/databases the team are using. - skulls / seeds and fruit may not be measured by the same people or in the same way or the measurements need to be consistent and measure what was intended	B1 for a comment on the hypothesis B1 for each of up to four comments on data collection	(5)	1

Question number	Answer	Additional guidance	Mark
3 (a)	B1 Secondary (data)	B1	(1)
(b)	B1 e.g. (LA) Clippers win more games than (Los Angeles) Lakers each year	B1 for a comparison made between the two teams for each year.	(1)
(c)	B1 e.g. (The number of games won by the (Los Angeles) Lakers year on year is) increasing/improving	B1 for a comment made on the overall trend seen.	(1)
(d)	B1 for 57% or 53% B2 for 57% and 53%	B1 for awrt to 57% or 53% B2 for awrt to 57% and 53% Note: The reference to round to nearest whole number in the question was merely to assist candidates but is not a requirement for the B1 marks.	(2)
(e)	B1 e.g. - LA Clippers (because) they won 54% of their games whereas the Los Angeles Lakers only won 52-53%. - LA Clippers, 54% > 53%	Note: for the Lakers, accept a value stated between 52 and 53% inclusive.	(1)

--	--	--	--

Question number	Answer	Additional guidance	Mark
4 (a)	M1 80 – ‘24’ (= 56) or for evidence of reading off graph at 440 and stating a value between 23- 24 A1 56 or 57	M1 for correct subtraction seen or for line drawn on graph indicating a valid attempt at reading off graph at 440 A1 56 or 57	(2)
(b)	B1 e.g. • Easy to do/convenient • Takes less time/more efficient	B1 for a correct comment about the advantages of using opportunity sampling	(1)
(c)	M1 for a box with two whiskers AND at least two values plotted correctly.  Yukon Chena Length (millimetres) A1 for all correct	For M1 and A1	(2)

(d)	B1 e.g. median length of Pike in Yukon is greater. B1 e.g. IQR is smaller for Yukon (≈ 145) than Chena (≈ 165) (or) e.g Range is smaller for Yukon (≈ 595) than Chena (≈ 615). B1 e.g. Yukon negatively skewed (or) Yukon symmetrical / Chena symmetrical (or) Chena positively skewed depB1 for any of the above interpreted in context e.g. - Yukon <u>Pikes are longer</u> (on average) or Chena <u>Pike are shorter</u> (on average) - The <u>lengths</u> of Pike in Chena are more widely dispersed, or the <u>lengths</u> of Pikes in Yukon are more consistent. - The spread of <u>lengths</u> above the average (median) for Yukon is smaller than the spread of <u>lengths</u> below the average, whereas the spread of lengths above the average for Chena is larger than the spread of lengths below the average. OR The spread of <u>lengths</u> above the average (median) is the same as the spread of <u>lengths</u> below the average for both Yukon and Chena.	B1 For a correct statistical statement comparing medians. Condone longer for comparison of medians. Must refer to <u>median</u>, not average for this B mark. Condone misspelling but medium is B0. B1 For a correct comparison of the IQRs/ranges. Condone wider for comparison of IQR/range. Calculated values not needed. B1 For correct comparison of the skews. For Chena also accept no skew but not neutral/normal/symmetrical skew. depB1 for a correct contextual interpretation comparing the medians/IQRs or skew	(4)
-----	---	--	-----

Question number	Answer	Additional guidance	Mark
5(a)	B1 for 7 and 13 in correct place B1 for 30, 10 and 9 in correct place B1 for 38 in correct place 		(3)
(b)	B1 $\frac{35}{120}$	B1 allow equivalent fraction, awrt 0.29 or awrt 29%	(1)
(c)	M1 $\frac{k}{40}$ or $\frac{\frac{23}{120}}{\frac{40}{120}}$ A1 $\frac{23}{40}$	M1 for realising 40 is the denominator or equivalent A1 allow equivalent fraction, 0.575 or 57.5%	(2)

(d)	B2 ft No, (Toga is incorrect) since $P(G D) \neq P(G)$ (B1 ft $P(G D) \neq P(G)$)	B2 ft for a correct conclusion with correct mathematical deduction that $P(G D) \neq P(G)$ (B1ft for correct mathematical comparison with no conclusion)	(2)

Question number	Answer	Additional guidance	Mark
6 (a)	B1 e.g. (Marcus might be wrong because) the pie charts do not show the number of livestock (they only show the proportion/percentage was greater for turkeys in Finland)	Allow an implication that the totals for Finland and Norway may be different.	(1)
(b)	M1 $\frac{2341780}{100} \times 35$	M1 for any full correct equivalent complete method.	(2)

	A1 819 623	A1 819623 Correct answer only with no working scores M1A1	
(c)	M1 $5 \times \frac{\sqrt{14\,888\,487}}{\sqrt{2\,341\,780}} \quad (= 12.6\dots)$ A1 12.6 (cm)	M1 for $5 \times \frac{\sqrt{14\,888\,487}}{\sqrt{2\,341\,780}} \text{ " or}$ $5 \div \frac{\sqrt{2\,341\,780}}{\sqrt{14\,888\,487}} \text{ "$ or $5 \times 2.521460\dots$ A1 awrt 12.6 Note: The reference to round to 1 decimal place in the question was merely to assist candidates but is not a requirement for the A1 marks	(2)

Question number	Answer	Additional guidance	Mark
7(a)	M1 $0.9 + 0.6 + 0.4 + 0.2 + 0.1$ A1 $= 2.2$	M1 A1	(2)
(b)	M1 $30 + (50 - 32.03) \times \frac{45 - 30}{19.52}$ 43.808 ... A1 44 years (or M1 $30 + \frac{17.12}{19.52} \times 15$ 43.155... A1 43 years)	M1 for correct method of interpolation using the tabulated data A1 for a correct median value awrt 44 (or M1 for correct method of interpolation using the tabulated data A1 for a correct median value awrt 43)	(2)
(c)	B1eg ‘44’ (years)...is greater than 27 (years) so Kiran’s claim is correct. 51.55% > 50% , the median age falls in the $45 \leq y < 59$ interval, so the new median age > 22.	B1 for a correct comparison with a conclusion. Allow a comparison with the tabulated data and comparing to an interval the median falls within.	(1)

(d)	B1eg • Loss of accuracy/introducing further error in calculations • for the class intervals has assumed the data is evenly spread	B1 Comment explaining the implication of merging or assuming evenly spread data	(1)
-----	--	---	-----

Question number	Answer	Additional guidance	Mark
8 (a)	B1 e.g. - data is bivariate - data is in pairs		(1)
(b)	B1 e.g. height is the independent variable (so should be plotted on the x -axis)		(1)
(c)	B1 for no (linear) correlation. .	B1 Know that the PMCC measures the strength (of a linear) correlation.	(1)
(d)	B2 e.g. (Since) $r = -0.5$ implies a (weak) negative (linear) correlation ... as the long jump distance increases the 100 m race time decreases (B1 e.g. $r = -0.5$ implies a (weak) negative (linear) correlation	B2 for interpretation of the PMCC and a correct contextual interpretation Accept converse when describing the correlation e.g. as the race time decreases the long jump (distance) increases Ignore reference to strength of correlation. (B1 for correct interpretation of PMCC)	(2)
(e)	B2 for interpolating/inside of range for race times so is reliable. (B1 for reference to interpolation/inside of range for race times with no or incorrect conclusion)		(2)

	Or B2 for sample size is small (12) so the result would be unreliable. (B1 for reference to the sample size being too small with no or incorrect conclusion)		
--	---	--	--

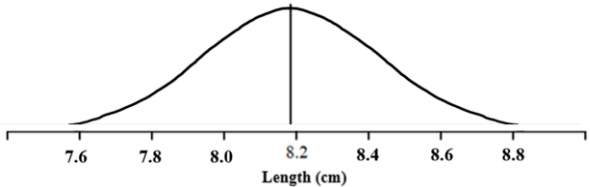
Question number	Answer	Additional guidance	Mark
9(a)	B1 e.g. (The control measures put in place) avoid any bias from their sampling method.	B1 A comment on trying to eliminate bias with regards to the water samples, maintaining reliability and validity.	(1)
(b)	M1 $7 + (3 \times \sigma) = 9.4$ A1 ($\sigma =$) 0.8 oe	M1 For correct calculation seen for the upper action limit to find the standard deviation A1 for correct or equivalent answer	(2)
(c)	B1 e.g. $10 > 9.4$ (which is 3 standard deviations above the mean) so it is very unlikely (for this level of lead) to occur.	B1 for knowledge that for a normal distribution, values more than three standard deviations from the mean are very unusual	(1)
(d)	M1 e.g. - All samples means are within warning limits M1 e.g., - all the sample ranges within warning limits A1 No action is required	M1 correct comment with reference to the control chart for means M1 correct comment with reference to the control chart for ranges	(3)

		A1 dependent on scoring both of the M marks	
--	--	---	--

Question number	Answer	Additional guidance	Mark
10(a)	M1 $222000 / 224000 \times 100$ A1 99.1 M1 ft for $\sqrt[5]{111.1 \times 106.7 \times 105.5 \times 105.4 \times "99.1"}$ A1 ft 105.5 B1 ft (The average increase is) "5.5%" <u>per year</u>	M1 for correct calculation of chain base index number A1 for awrt 99.1 M1 ft for correct calculation of the geometric mean of 5 chain base index numbers A1 ft for awrt 105.5 B1 ft for a correct interpretation of the value of the geometric mean.	(5)
(b)	M1 $3 \times \frac{235884 - 213250}{119492}$ A1 $0.5682... = 0.6$ So, positive skew	M1 correct use of Pearsons Skew formula. A1 awrt 0.6 and stating the skew is positive	(2)

Question number	Answer	Additional guidance	Mark
11(a)	M1M1 $\frac{(32.5 \times 25) + (37.5 \times 18) + (42.5p) + (47.5 \times 14) + (52.5 \times 3)}{25 + 18 + p + 14 + 3}$ (= 40.5) $[2310 + 42.5p = 40.5(60 + p)]$ $[2310 + 42.5p = 2430 + 40.5p]$ $[42.5p - 40.5p = 2430 - 2310]$ $[2p = 120]$ $[p = 120 \div 2]$ A1 $p = 60$	M1 for consistent use of fx with x within interval (including end points) and attempt to sum Use of mid-interval gives $812.5 + 675 + 42.5p + 665 + 157.5$ May also be seen as $2310 + 42.5p$ M1 the sum $25 + 18 + p + 14 + 3$ seen May also be seen as $60 + p$ The following all imply M1M1 $2310 + 42.5p = 40.5(60 + p)$ $2310 + 42.5p = 2430 + 40.5p$ $42.5p - 40.5p = 2430 - 2310$ $2p = 120$ $p = 120 \div 2$ A1 for $p = 60$ If $p = 60$ given with no working award M1M1A1	(3)
(b)	B1 e.g. - we do know the total number of cars that have been surveyed, only those that exceeded the speed limit. - We cannot work out the percentage of cars that broke the speed limit		(1)

Question number	Answer	Additional guidance	Mark
12(a)	M1 for 22.85 – 21.33 (or sight of 1.52) Girls M1 for 23.45 – 21.86 (or sight of 1.59) Boys A1ft girls IPR is smaller than boys IPR. For the award of this mark the boys IPR and girls IPR must be clearly identified. Follow through their values to award A1. ALT M1 5th percentile B – G 21.86 – 21.33 = 0.53 M1 75th percentile B – G 23.45 – 22.85 = 0.6 A1ft Hence Boys <u>IPR</u> > Girls <u>IPR</u> Follow through their values to award A1	Only award the A mark if at least one M mark has been awarded. Only award the A mark if at least one M mark has been awarded.	(3)
(b)	B1ft the eye length of (6 year old) girls are more consistent [than the eye lengths of (6 year old) boys.] Or, (6 year old) boys eye lengths are less consistent [than (6 year old) girls eye lengths].		(1)

Question number	Answer	Additional guidance	Mark
13 (a)	B1 $(8.6 - 8.2) \div 2 (= 0.2)$	B1 for the correct calculation seen	(1)
(b)	B2 bell shaped curve centred on 8.2 and tails ending at 7.6 and 8.8  (B1 bell shaped curve centred on 8.2)	B2 There must be an attempt at a symmetrical bell-shaped curve centred around 8.2 and at least one value indicated on the axis to emphasise the 7.6 and 8.8 tails. Condone the tails being labelled as 7.6 and 8.8 at different points on the axis. (B1 for an attempt at a symmetrical bell-shaped curve centred around 8.2 but with no or incorrect scale applied to axis to indicate correct position of tails)	(2)
(c)	B1 e.g. • '<u>unlikely</u>' (since) 7.5 cm is an outlier • '<u>unlikely</u>' (since) 7.5 cm is smaller than 7.6 • '<u>unlikely</u>' (since) 7.5 cm is smaller than $\mu - 3\sigma$	B1 for indicating that the 7.5 cm is an unlikely (or equivalent phrase) and stating this is because 7.5 cm is an outlier or some other equivalent reason.	(1)

Question number	Answer	Additional guidance	Mark
13 (d)	<u>Tabulation</u> Comments to support Liam's tabulation being appropriate: B1 B1 for any two from (Appropriate since...) • they do not overlap • equal class intervals are suitable if data is closely grouped within a range of values and not more spread out in others • as normally distributed data would be more closely clustered in the middle of the range of values • They are exhaustive (since no butterfly lives longer than 15 days) OR Comments to support Liam's tabulation being inappropriate: B1 B1 for any one from (Inappropriate since...) • (for a normal distribution) Less data is collected for the extreme class intervals, so the groups would be merged for lower and higher group intervals. • (for normal distributions) the class intervals need not be the same width / frequency densities would be more appropriate for different sized class intervals. • more class intervals should have been included <u>Histogram</u>	B1 B1 for two comments on the class intervals. Candidate must decide if the tabulation method is appropriate or not and give two reasons according. One of each is not permitted. B1 for comment on the use of a histogram	(5)

