



National
Qualifications
2026

2026 Statistics

Advanced Higher Paper 1

Question Paper Finalised Marking Instructions

These marking instructions have been prepared by examination teams for use by Qualifications Scotland appointed markers when marking external course assessments.

The information in this document may be reproduced in support of Qualifications Scotland qualifications only on a non-commercial basis. If it is reproduced, Qualifications Scotland must be clearly acknowledged as the source. If it is to be reproduced for any other purpose, written permission must be obtained from Qualifications Scotland.

General marking principles for Advanced Higher Statistics

Always apply these general principles. Use them in conjunction with the detailed marking instructions, which identify the key features required in candidates' responses.

The marking instructions for each question are generally in two sections:

- *generic scheme – this indicates why each mark is awarded*
- *illustrative scheme – this covers methods which are commonly seen throughout the marking*

In general, you should use the illustrative scheme. Only use the generic scheme where a candidate has used a method not covered in the illustrative scheme.

- Always use positive marking. This means candidates accumulate marks for the demonstration of relevant skills, knowledge and understanding; marks are not deducted for errors or omissions.
- If you are uncertain how to assess a specific candidate response because it is not covered by the general marking principles or the detailed marking instructions, you must seek guidance from your team leader.
- One mark is available for each •. There are no half marks.
- If a candidate's response contains an error, all working subsequent to this error must still be marked. Only award marks if the level of difficulty in their working is similar to the level of difficulty in the illustrative scheme.
- Only award full marks where the solution contains appropriate working. A correct answer with no working receives no mark, unless specifically mentioned in the marking instructions.
- Candidates may use any mathematically correct method to answer questions, except in cases where a particular method is specified or excluded.
- If an error is trivial, casual or insignificant, for example $6 \times 6 = 12$, candidates lose the opportunity to gain a mark, except for instances such as the second example in point (h) below.
- If a candidate makes a transcription error (question paper to script or within script), they lose the opportunity to gain the next process mark, for example

This is a transcription error and so the mark is not awarded.

This is no longer a solution of a quadratic equation, so the mark is not awarded.

$$x^2 + 5x + 7 = 9x + 4$$

$$x - 4x + 3 = 0$$

$$x = 1$$

The following example is an exception to the above

This error is not treated as a transcription error, as the candidate deals with the intended quadratic equation. The candidate has been given the benefit of the doubt and all marks awarded.

$$x^2 + 5x + 7 = 9x + 4$$

$$x - 4x + 3 = 0$$

$$(x - 3)(x - 1) = 0$$

$$x = 1 \text{ or } 3$$

(i) **Horizontal/vertical marking**

If a question results in two pairs of solutions, apply the following technique, but only if indicated in the detailed marking instructions for the question.

Example:

$$\begin{array}{cc} \bullet^5 & \bullet^6 \\ \bullet^5 & x = 2 \quad x = -4 \\ \bullet^6 & y = 5 \quad y = -7 \end{array}$$

Horizontal: $\bullet^5 x = 2$ and $x = -4$ Vertical: $\bullet^5 x = 2$ and $y = 5$
 $\bullet^6 y = 5$ and $y = -7$ $\bullet^6 x = -4$ and $y = -7$

You must choose whichever method benefits the candidate, **not** a combination of both.

(j) In final answers, candidates should simplify numerical values as far as possible unless specifically mentioned in the detailed marking instruction. For example

$\frac{15}{12}$ must be simplified to $\frac{5}{4}$ or $1\frac{1}{4}$ $\frac{43}{1}$ must be simplified to 43

$\frac{15}{0.3}$ must be simplified to 50 $\frac{4\cancel{5}}{3}$ must be simplified to $\frac{4}{15}$

$\sqrt{64}$ must be simplified to 8*

*The square root of perfect squares up to and including 144 must be known.

(k) Do not penalise candidates for any of the following, unless specifically mentioned in the detailed marking instructions:

- working subsequent to a correct answer
- correct working in the wrong part of a question
- legitimate variations in numerical answers/algebraic expressions, for example angles in degrees rounded to nearest degree
- omission of units
- bad form (bad form only becomes bad form if subsequent working is correct), for example

$(x^3 + 2x^2 + 3x + 2)(2x + 1)$ written as

$(x^3 + 2x^2 + 3x + 2) \times 2x + 1$

$= 2x^4 + 5x^3 + 8x^2 + 7x + 2$

gains full credit

- repeated error within a question, but not between questions or papers

(l) In any 'Show that...' question, where candidates have to arrive at a required result, the last mark is not awarded as a follow-through from a previous error, unless specified in the detailed marking instructions.

(m) You must check all working carefully, even where a fundamental misunderstanding is apparent early in a candidate's response. You may still be able to award marks later in the question so you must refer continually to the marking instructions. The appearance of the correct answer does not necessarily indicate that you can award all the available marks to a candidate.

(n) You should mark legible scored-out working that has not been replaced. However, if the scored-out working has been replaced, you must only mark the replacement working.

- (o) If candidates make multiple attempts using the same strategy and do not identify their final answer, mark all attempts and award the lowest mark. If candidates try different valid strategies, apply the above rule to attempts within each strategy and then award the highest mark.

For example:

Strategy 1 attempt 1 is worth 3 marks.	Strategy 2 attempt 1 is worth 1 mark.
Strategy 1 attempt 2 is worth 4 marks.	Strategy 2 attempt 2 is worth 5 marks.
From the attempts using strategy 1, the resultant mark would be 3.	From the attempts using strategy 2, the resultant mark would be 1.

In this case, award 3 marks.

Marking Instructions for each question

Question			Generic scheme	Illustrative scheme	Max mark
1.	(a)		• ¹ state hypotheses	•¹ H₀: There is no association between a player's country and the number of times the player has played for that country H₁: There is an association between a player's country and the number of times the player has played for that country	1
Notes: 1. For • ¹ , also accept phrasing of 'independent/not independent'.					
Commonly Observed Responses:					
	(b)		• ² calculate expected frequency	$\left\{ \begin{array}{l} \frac{(\text{row 1 total})(\text{column 3 total})}{(\text{overall total})} \\ = \frac{(7+12+8+4+1)(8+6+3+6)}{132} \\ = 5.58 \end{array} \right.$	1
Notes: 1. For • ² , also accept: row total calculation of $32 - 26.43 = 5.57$ column total calculation of $23 - 17.43 = 5.57$ grid total calculation of $132 - 126.44 = 5.56$					
Commonly Observed Responses:					
	(c)		• ³ state reason	• ³ because only 60% (12 out of 20) of the expected values are ≥ 5 , and we need at least 80% of the expected values need to be at least 5	1
Notes: 1. For • ³ , there must be comparison of a statistic from the table of expected values to a threshold value. 2. For • ³ , also accept reference to '.... less than 20% of expected values can be less than 5'. 3. For • ³ , do not accept responses that only refer to threshold values.					
Commonly Observed Responses:					

Question			Generic scheme	Illustrative scheme	Max mark
1.	(d)		•⁴ appropriate critical value •⁵ deal with H_0 •⁶ appropriate conclusion	•⁴ $\chi^2_{9,0.95} = 16.919$ •⁵ as $13.942 < 16.919$ do not reject H_0 at the 5% significance level •⁶ there is insufficient evidence of an association between a player's country and the number of times the player has played for that country	3
Notes: 1. For • ⁴ and • ⁵ , also accept p -value of 0.1244, and '0.1244 > 0.05 so do not reject H_0 '. 2. For • ⁵ , must include reference to the chosen significance level. 3. For • ⁵ , do not accept reference to '95% level of significance', or similar. 4. For • ⁶ , refer back to mark • ¹ to check for consistent wording. 5. For • ⁶ , conclusion must be phrased in terms of H_1 . 6. For • ⁶ , do not accept conclusions that are too definite. Phrasing must include 'evidence to conclude...', or 'evidence to suggest...', or similar.					
Commonly Observed Responses:					
	(e)	(i)	• ⁷ appropriate comment	• ⁷ positive linear relationship between mass and height	1
Notes: 1. For • ⁷ , also accept: phrasing of '...between height and mass' (the same variable names, but interchanged). 'positive correlation'. 'positive linear association'. 2. For • ⁷ , do not accept: 'positive association'.					
Commonly Observed Responses:					
		(ii)	• ⁸ appropriate comment	• ⁸ (the slope for 'back' players is greater than the slope for 'forward' players which means) for every cm increase in height for 'back' players, it gives a greater increase in mass than for 'forward' players	1
Notes: 1. For • ⁸ , also accept: 'the slope for 'forward' players is less than the slope for 'back' players which means for every cm increase in height (for 'forward' players), it gives a lesser increase in mass (than for 'back' players)'. 'as height increases, mass increases at a faster rate for 'back' players than for 'forward' players'. 2. For • ⁸ , the relationship must be stated in context, citing mass and height and playing position.					
Commonly Observed Responses: Candidates who refer to 'strength of correlation', award no marks.					

Question			Generic scheme	Illustrative scheme	Max mark
1.	(f)		•⁹ deal with H_0 •¹⁰ appropriate conclusion	•⁹ as $p\text{-value} < 0.05$, we reject H_0 •¹⁰ we have evidence to suggest that the linear model is useful for predicting mass from height for 'forward' position players	2
Notes: 1. For • ¹⁰ , do not accept 'we have evidence to suggest that there is a linear relationship between the masses and heights of players playing in 'forward' positions.' (as this is the conclusion for a hypothesis test on ρ).					
Commonly Observed Responses:					
	(g)		• ¹¹ calculate fitted value	• ¹¹ 116.64	1
Notes:					
Commonly Observed Responses:					
	(h)		•¹² interpret the prediction interval •¹³ interpret the confidence interval	•¹² 95% of the time, the mass of an individual 'forward' position player (of height 202 cm) is expected to be between 101 kg and 132 kg •¹³ 95% of the time, the mean mass of 'forward' position players (with height 202 cm) is expected to be between 115 kg and 118 kg	2
Notes: 1. For • ¹² and • ¹³ , any number of decimal places is acceptable. 2. For • ¹² do not accept '95% confidence' in the response, as it is a prediction interval.					
Commonly Observed Responses:					
	(i)		• ¹⁴ appropriate comment	• ¹⁴ it is not suitable as you are extrapolating beyond the range of observed data	1
Notes:					
Commonly Observed Responses:					

Question			Generic scheme	Illustrative scheme	Max mark
2.	(a)		•¹ appropriate description •² appropriate suggestion	•¹ the garages that did not appear on the MOT cost search website •² using the original data set's details (for example, phone them up and ask them)	2
Notes: 1. For • ² , also accept 'visit them in person, using their postal address'.					
Commonly Observed Responses:					
	(b)		•³ appropriate explanation	•³ a confidence interval ought to capture the population mean value (and not the maximum value)	1
Notes:					
Commonly Observed Responses:					
	(c)		•⁴ appropriate description	•⁴ population standard deviations (of cost) were not known for either sample	1
Notes: 1. For • ⁴ , also accept ' population variances (of cost) were not known for either sample'.					
Commonly Observed Responses:					
	(d)		•⁵ correct value for s •⁶ state assumption	•⁵ $\sqrt{\frac{(18-1) \times 8.627261^2 + (34-1) \times 8.099399^2}{18+34-2}}$ $= 8.282647$ •⁶ standard deviations of the two populations of costs are equal	2
Notes: 1. For • ⁶ , also accept 'variances of the two populations of costs are equal'.					
Commonly Observed Responses:					

Question			Generic scheme	Illustrative scheme	Max mark
2.	(e)		•⁷ correct parameters •⁸ use continuity correction •⁹ standardise to calculate z	•⁷ $E(W) = 477 \quad V(W) = 2703$ •^{8,9} $\frac{404.5 - 477}{\sqrt{2703}} = -1.39$	3
Notes:					
Commonly Observed Responses: Candidate A does not use continuity correction. $z = \frac{404 - 477}{\sqrt{2703}} = -1.40$ award •⁸ × •⁹ ✓.					
	(f)		•¹⁰ appropriate description	•¹⁰ one p-value is greater than 10%, whilst the other p-value is less than 10%	1
Notes:					
Commonly Observed Responses:					
	(g)		•¹¹ appropriate description •¹² appropriate description	•¹¹ should refer to median, not mean •¹² should be 'less than' not 'different to'	2
Notes:					
Commonly Observed Responses:					
	(h)	(i)	•¹³ appropriate description	•¹³ (at least one of) the distributions of samples is not symmetrical	1
Notes:					
1. For •¹³, do not accept 'boxplots are skewed', or 'dotplots are skewed'. Reference must be made to the distribution of data itself.					
Commonly Observed Responses: Candidates who referred to 'positive skew', 'negative skew', 'left skew', 'right skew' or similar, often gained no marks as their understanding of the meaning of these phrases was incorrect.					

Question			Generic scheme	Illustrative scheme	Max mark
2.	(h)	(ii)	•¹⁴ state part of analysis •¹⁵ state part of analysis	•¹⁴ confidence intervals •¹⁵ <i>t</i>-test	2

Notes:

- For •¹⁴, also accept ‘*t*-intervals’.

Commonly Observed Responses:

Candidate A provides more than two parts of the analysis for marks •¹⁴ and •¹⁵.

Treat each statement as using the same strategy, mark each possible combination of pairs of statements and then award the lowest mark from these combinations:

Number of statements	Number correct	Number incorrect	Marks Awarded
3	0	3	0
3	1	2	0
3	2	1	1
4	0	4	0
4	1	3	0
4	2	2	0

	(i)		•¹⁶ appropriate description	•¹⁶ national Chains sample was not randomly selected	1
--	-----	--	--	---	---

Notes:

- For •¹⁶, also accept ‘National Chains sample was convenience sampling, and not representative of population’ (ie a reason is given for it not being representative).
- For •¹⁶, do not accept responses that refer to sample sizes.

Commonly Observed Responses:

[END OF MARKING INSTRUCTIONS]