National Qualifications 2023

2023 Statistics

Advanced Higher - Paper 1

Finalised Marking Instructions

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.

(a) 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.

(b) 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.

(c) One mark is available for each •. There are no half marks.

(d) 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.

(e) 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.

(f) Candidates may use any mathematically correct method to answer questions, except in cases where a particular method is specified or excluded.

(g) If an error is trivial, casual or insignificant, for example 6 × 6 = 12, candidates lose the opportunity to gain a mark, except for instances such as the second example in point (h) below.

(h) 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:

	•5	•6
•5	$x = 2$	$x = -4$
•6	$y = 5$	$y = -7$

Horizontal: •5 $x = 2$ and $x = -4$ •6 $y = 5$ and $y = -7$

Vertical: •5 $x = 2$ and $y = 5$ •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

$\frac{15}{0{\cdot}3}$ must be simplified to 50 $\frac{4\!\!\not/5}{3}$ 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

$\left(x^3 + 2x^2 + 3x + 2\right)(2x + 1)$ *written as*

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

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

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)		$\bullet^1$ appropriate comment $\bullet^2$ appropriate comment	$\bullet^1$ median (fat content) of bakery items is greater than non-bakery items $\bullet^2$ interquartile ranges of fat content are similar	**2**
Notes: 1. For $\bullet^1$, must mention at least the median 2. For $\bullet^1$, do not accept 'on average' 3. For $\bullet^2$, also accept 'similar ranges' 4. For $\bullet^2$, do not accept 'spread', or 'variation'. Technical names of either range or interquartile range are required. 5. Candidates who comment on the calories boxplot gain no marks for those comments.					
Commonly Observed Responses:					
	(b)		$\bullet^3$ state criteria	$\bullet^3$ any data point that is either above the upper fence or below the lower fence	**1**
Notes: 1. For $\bullet^3$, this can also be expressed using mathematical notation: eg values < Q1 − 1.5 IQR **or** > Q3 + 1.5 IQR 2. Mark $\bullet^3$ can be awarded if in part (c) candidates calculate and check both upper and lower fences.					
Commonly Observed Responses: 1. Candidates who simply state what a fence is, gain 0 marks. 2. Candidates who refer to 'inner fences' and 'outer fences' gain 0 marks.					
	(c)		$\bullet^4$ calculate fences $\bullet^5$ show that values are outliers	$\bullet^4$ from Output 1 calorie/bakery group: 1.5 × IQR = 1.5 × (430 – 320) = 165 Q1 – 1.5 × IQR = 320 – 165 = 155 $\bullet^5$ from Figure 1 the two values labelled as • in the calories/bakery group have values <150<155 and are therefore outliers.	**2**
Notes:					
Commonly Observed Responses:					

Question			Generic scheme	Illustrative scheme	Max mark
1.	(d)		$\bullet^6$ appropriate reason	$\bullet^6$ the values came from data entry errors	**1**

Notes:

1. For $\bullet^6$, other reasons are acceptable. eg the bakery items corresponding to these values are qualitatively different from the other bakery items and thus form a third class
2. For $\bullet^6$, also accept 'data collection errors' and 'misclassification'
3. For $\bullet^6$, do not accept reasons based on the effects outliers may have on subsequent analysis of the data

Commonly Observed Responses:

1. Candidates who describe 'anomalous data' gain 0 marks.
2. Candidates who make reference to outliers 'skewing the data' gain 0 marks.

	(e)		$\bullet^7$ appropriate information	$\bullet^7$ the (four) sample variances	**1**

Notes:

1. For $\bullet^7$, also accept 'the (four) sample standard deviations'
2. For $\bullet^7$, do not accept 'population variance' or 'population standard deviation'

Commonly Observed Responses:

1. Candidates who make reference to assumptions about normality gain 0 marks.

	(f)		$\bullet^8$ state both null hypotheses	$\bullet^8$ $H_0 : \mu_{\text{bakery fat}} = \mu_{\text{non-bakery fat}}$ $H_0 : \mu_{\text{bakery calories}} = \mu_{\text{non-bakery calories}}$	**1**

Notes:

1. For $\bullet^8$, must mention 'fat' and 'calories', or distinguish between Output 2 and Output 3
2. For $\bullet^8$, must mention 'bakery' and 'non-bakery'
3. For $\bullet^8$, also accept $\begin{cases} H_0 : \mu_{\text{bakery fat}} - \mu_{\text{non-bakery fat}} = 0 \\ H_0 : \mu_{\text{bakery calories}} - \mu_{\text{non-bakery calories}} = 0 \end{cases}$
4. For $\bullet^8$, do not accept H_0 referring to the mean of differences ie $H_0 : \mu_{\text{difference}} = 0$

Commonly Observed Responses:

1. Candidates who state "true difference in means equal to zero" gain 0 marks (see note 2)

Question			Generic scheme	Illustrative scheme	Max mark
1.	(g)		$\bullet^9$ approximate probability $\bullet^{10}$ double to obtain p-value	$\bullet^9\ \bullet^{10}$ $\begin{cases} p\text{-value} = 2\times \text{P}(Z > 1.05) \\ = 2\times(1-0.8531) \\ = 0.2938 \end{cases}$	**2**

Notes:
1. For $\bullet^9$, the answer of 0.8531 gains 0 marks.
2. For $\bullet^9$, the standard normal is an appropriate approximation to the t_{75} distribution
3. For $\bullet^{10}$, if the approximate value of $2\times \text{P}(t_{75} > 1.0496) = 2\times \text{P}(Z > 1.0496) = 2\times 0.1470 = 0.2939$ is obtained from using a graphic calculator then award both $\bullet^9$ and $\bullet^{10}$
4. For $\bullet^{10}$, if the exact value of 2 × P(t_{75} >1.0496) = 2 × 0.1486 = 0.2973 is obtained from using a graphic calculator then award both $\bullet^9$ and $\bullet^{10}$

Commonly Observed Responses:

	(h)		$\bullet^{11}$ interpret the p-value $\bullet^{12}$ appropriate conclusion	$\bullet^{11}$ as p-value < 0.05 (we reject H_0) $\bullet^{12}$ (we have evidence that the mean calories for bakery is not equal to the mean calories for non-bakery) which suggests that there is an impact to your mean calorie intake.	**2**

Notes:
1. For $\bullet^{11}$, also accept references to the p-value being small
2. For $\bullet^{11}$, accept any appropriate significance level from 0.01 to 0.10
3. For $\bullet^{12}$, one tailed responses that say calories are higher/lower gain 0 marks.

Commonly Observed Responses:

	(i)		$\bullet^{13}$ state method $\bullet^{14}$ appropriate description $\bullet^{15}$ appropriate description	$\bullet^{13}$ simple random sampling $\bullet^{14}$ all coffee shop chains are listed and given a number $\bullet^{15}$ one coffee shop chain is selected at random	**3**

Notes:
1. For $\bullet^{13}$, also accept cluster sampling. In particular, two-stage cluster sampling since a random sample of bakery and non-bakery items is taken from the selected cluster.

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
2.	(a)	(i)	$\bullet^1$ appropriate comment for scatter plot	$\bullet^1$ the scatter plot indicates a (very strong) **linear** relationship between crop density and crop yield	**1**
Notes: 1. For $\bullet^1$, do not accept 'positive correlation' 2. For $\bullet^1$, candidates must mention the context of crop density and crop yield					
Commonly Observed Responses:					
		(ii)	$\bullet^2$ appropriate statement $\bullet^3$ appropriate statement	$\bullet^2$ subtract the fitted value from the observed value $\bullet^3$ a residual measures the error that is not explained by the regression line	**2**
Notes: 1. For $\bullet^2$, also accept 'a residual is the observed y-value subtract the y-value predicted by the model' 2. The awarding of mark $\bullet^2$ requires the direction of subtraction to be clearly stated, whereas the award of mark $\bullet^3$ does not.					
Commonly Observed Responses: 1. Candidates who reference the sums of the squares of residuals gain 0 marks.					
		(iii)	$\bullet^4$ appropriate comment $\bullet^5$ appropriate comment	$\bullet^4$ there is a pattern… $\bullet^5$ … that is quadratic/U-shaped	**2**
Notes: 1. For $\bullet^4$, also accept 'not randomly scattered'					
Commonly Observed Responses: 1. Candidates who make comments referring to $E(\varepsilon_i) \neq 0$ gain no marks. 2. Candidates who make comments referring to $V(\varepsilon_i) \neq$ constant gain no marks.					

Question			Generic scheme	Illustrative scheme	Max mark
2.	(b)	(i)	•[6] calculate b	•[6] $b=\dfrac{S_{x\sqrt{y}}}{S_{xx}}=1.7322$	**3**
			•[7] calculate a	•[7] $\begin{cases} a=\overline{\sqrt{y}}-b\bar{x} \\ =28.4813-b\times 5 \\ =19.8203 \end{cases}$	
			•[8] state equation	•[8] $\sqrt{y}=19.8203+1.7322x$	
Notes: 1. For •[8], equation must have $\sqrt{y}$ as the subject					
Commonly Observed Responses:					
		(ii)	•[9] calculate $\sqrt{y}$	•[9] $\sqrt{y}=1.7322\times 3.5+19.8203$ $=25.883$	**3**
			•[10] calculate residual	•[10] 26.0 – 25.883 = 0.117	
			•[11] circle correct point	•[11] point (25.883, 0.117) circled	
Notes:					
Commonly Observed Responses:					

Question			Generic scheme	Illustrative scheme	Max mark
2.	(c)	(i)	$\bullet^{12}$ appropriate comment	$\bullet^{12}$ you are extrapolating beyond the range of the data set	**1**
Notes: 1. For $\bullet^{12}$, candidates must communicate issues with going beyond 8 plants/m^2					
Commonly Observed Responses:					
		(ii)	$\bullet^{13}$ appropriate comment	$\bullet^{13}$ the least squares regression line will never show a decreasing yield for an increasing crop density (for this transformation)	**1**
Notes:					
Commonly Observed Responses:					
	(d)		$\bullet^{14}$ appropriate relationship $\bullet^{15}$ appropriate range of values	$\bullet^{14}$ there is a linear relationship between $\sqrt{\text{crop yield}}$ and crop density $\bullet^{15}$ for crop densities between 2 and 8 plants per square metre	**2**
Notes: 1. For $\bullet^{14}$, also accept reference to poorly fitting model of crop yield to crop density.					
Commonly Observed Responses:					

[END OF MARKING INSTRUCTIONS]