National
Qualifications
RESOURCE

X803/77/11

Statistics
Paper 1

Marking Instructions

Please note that these marking instructions have not been standardised based on candidate responses. You may therefore need to agree within your centre how to consistently mark an item if a candidate response is not covered by the 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:

	$\bullet^5$	$\bullet^6$
$\bullet^5$	$x = 2$	$x = -4$
$\bullet^6$	$y = 5$	$y = -7$

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

$\frac{15}{0{\cdot}3}$ 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

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

Detailed Marking Instructions for each question

Question			Generic scheme	Illustrative scheme	Max mark
1.	(a)		•[1] describe improvement	•[1] add title for each side, and add a key eg $0 \mid 5 = 0 \cdot 5$	**2**
			•[2] describe improvement	•[2] reverse the order of the digits (leaves) on the left hand side	
	(b)		•[3] calculate upper fence	•[3] upper fence = $6{\cdot}5 + 1{\cdot}5\,(6{\cdot}5 - 3{\cdot}7) = 10{\cdot}7$	**2**
			•[4] state number of outliers	•[4] 5 outliers	
	(c)	(i)	•[5] state hypothesis test	•[5] z-test for a difference in population proportions	**1**
		(ii)	•[6] state hypotheses	•[6] $H_0 : p_{cat} = p_{dog}$ $H_1 : p_{cat} \neq p_{dog}$	**1**
	(d)		•[7] state assumption	•[7] populations of times to draw the doodles of both cats and dogs have the same shape…	**2**
			•[8] state assumption	•[8] … and variability	
	(e)		•[9] correct parameters	•[9] $\begin{cases} E(W) = 16153 \cdot 5 \\ V(W) = 390376 \cdot 25 \end{cases}$	**3**
			•[10] standardise W ….	•[10&11] $\begin{cases} \dfrac{12048 \cdot 5 - 16153 \cdot 5}{\sqrt{390376 \cdot 25}} \\ (= -6 \cdot 57009) \end{cases}$	
			•[11] … with correct continuity correction		
	(f)		•[12] state hypotheses	•[12] H_0: median time to draw a cat = median time to draw a dog H_1: median time to draw a cat $\neq$ median time to draw a dog	**3**
			•[13] state significance level	•[13] 1% significance level	
			•[14] state conclusion	•[14] we reject H_0 and we have evidence to suggest that the median time to draw a cat is different to the median time to draw a dog.	

Question			Generic scheme	Illustrative scheme	Max mark
1.	(g)		•[15] state assumption	•[15] assume that the standard deviations of the population doodle times for cats and dogs are equal	**2**
			•[16] correct information	•[16] the sample standard deviations of 2·307 and 2·655	

Question			Generic scheme	Illustrative scheme	Max mark
2.	(a)		$\bullet^1$ appropriate comment $\bullet^2$ appropriate comment	$\bullet^1$ there is a positive relationship $\bullet^2$ …but it appears to be non-linear	**2**
	(b)		$\bullet^3$ appropriate comment $\bullet^4$ appropriate comment	$\bullet^3$ both models may be suitable $\bullet^4$ both have small p-values (to reject H_0)	**2**
	(c)		$\bullet^5$ correct value $\bullet^6$ appropriate explanation	$\bullet^5$ $R^2 = 0{\cdot}805$ $\bullet^6$ Model A explains 80% of the variation in **square root** of cost dependent upon the length of the yacht	**2**
	(d)		$\bullet^7$ appropriate comment $\bullet^8$ appropriate comment $\bullet^9$ appropriate comment	$\bullet^7$ both plots appear to have random scatter centred on zero $\bullet^8$ both plots appear to have constant variance $\bullet^9$ both models appear valid	**3**
	(e)		$\bullet^{10}$ calculate sqrt(cost) $\bullet^{11}$ calculate estimate $\bullet^{12}$ calculate interval	$\bullet^{10}$ estimated sqrt(cost) = $\begin{cases} -204{\cdot}693 + 55{\cdot}437 \times 15 \\ = 626{\cdot}8627 \end{cases}$ $\bullet^{11}$ estimated cost = $626{\cdot}8627^2$ = £392956 $\bullet^{12}$ 95% Confidence interval ($592{\cdot}0915^2$, $661{\cdot}634^2$) = (£350572, £437760)	**3**
	(f)	(i)	$\bullet^{13}$ appropriate reason	$\bullet^{13}$ both models are set up regressing transformed cost on length.	**1**
		(ii)	$\bullet^{14}$ appropriate suggestion	$\bullet^{14}$ He needs to design a model that has regressed length on (transformed) cost	**1**

[END OF MARKING INSTRUCTIONS]