National
Qualifications
2019

2019 Statistics

Advanced Higher

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.

For each question, the marking instructions 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 x 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

$$x^2 + 5x + 7 = 9x + 4$$
$$x - 4x + 3 = 0$$
$$x = 1$$

This is a transcription error and so the mark is not awarded. (refers to $x - 4x + 3 = 0$)

This is no longer a solution of a quadratic equation, so the mark is not awarded. (refers to $x = 1$)

The following example is an exception to the above

$$x^2 + 5x + 7 = 9x + 4$$
$$x - 4x + 3 = 0$$
$$(x-3)(x-1) = 0$$
$$x = 1 \text{ or } 3$$

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.

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

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

$\frac{15}{0\cdot3}$ must be simplified to 50

$\frac{4/5}{3}$ must be simplified to $\frac{4}{15}$

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

*The square root of perfect squares up to and including 100 must be known.

(k) Commonly Observed Responses (COR) are shown in the marking instructions to help mark common and/or non-routine solutions. CORs may also be used as a guide when marking similar non-routine candidate responses.

(l) 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

(m) 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.

(n) 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.

(o) 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.

(p) 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] correct strategy •[2] calculate probability	•[1] $P(X<3)=P(X\le 2)$ •[2] 0·4232	**2**
Notes:					
Commonly Observed Responses:					
	(b)		•[3] correct distribution •[4] calculate probability	•[3] $X \sim \text{Po}(21)$ •[4] $P(X=25)=0{\cdot}0555$	**2**
Notes:					
Commonly Observed Responses:					
2.	(a)		•[1] correct probability	•[1] $\frac{1}{9}$	**1**
Notes:					
Commonly Observed Responses:					
	(b)		•[2] correct strategy •[3] calculate probability •[4] appropriate reason •[5] appropriate conclusion	•[2] $P(CB).P(LH)=P(CB\cap LH)$ for independent events •[3] $\frac{90}{1000}.\frac{140}{1000}=\frac{63}{5000}$ •[4] $\frac{63}{5000}\ne\frac{10}{1000}$ •[5] so we may conclude that being colour blind is not independent of being left handed	**4**
Notes: One of several alternative methods would be $P(L\mid CB)\ne P(L)$: $\frac{1}{9}\ne\frac{140}{1000}$ etc					
Commonly Observed Responses:					

Question			Generic scheme	Illustrative scheme	Max mark
2.	(c)		•[6] appropriate hypotheses	•[6] H_0:There is no association between colour blindness and gender H_1:There is an association between colour blindness and gender	7
			•[7] correct E_i	•[7] 468·44 | 487·56 21·56 | 22·44	
			•[8] correct test statistic	•[8] $\chi^2 = 32 \cdot 3478$	
			•[9] correct cv	•[9] $\chi^2_{1,0\cdot950} = 3\cdot841$	
			•[10] deal with H_0	•[10] $32\cdot3478 > 3\cdot841$ we have evidence to reject H_0	
			•[11] appropriate conclusion	•[11] conclude that there is evidence of an association and	
			•[12] correct interpretation	•[12] colour blindness appears to be more common in males	

Notes:

The p-value approach (hereafter the PvA) would record that $P\left(\chi_1^2 < 32 \cdot 3478\right)$ $= 1\cdot29\times10^{-8} \ll 0\cdot05$ . . . . . .

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
3.	(a)		•¹ appropriate sampling strategy •² appropriate disadvantage •³ appropriate consequence in context	•¹ convenience sampling •² selection is almost certainly not representative of such books •³ and hence the proportion of female leads may be wrongly estimated	**3**
Notes:					
Commonly Observed Responses:					
	(b)		•⁴ appropriate sample frame •⁵ address randomness •⁶ appropriate sample	•⁴ obtain a list of all n current children's books and number each book consecutively from 1 to n •⁵ randomly select a number between 1 and 25 as the starting number •⁶ and then choose every 25^{th} number after that, selecting books corresponding to the chosen numbers	**3**
Notes: An alternative might be to start at any random point but the requirement to loop back must be acknowledged					
Commonly Observed Responses:					

Question			Generic scheme	Illustrative scheme	Max mark
4.			•[1] extract information	•[1] 64 matches and 37 won by scoring first	**5**
			•[2] appropriate strategy	•[2] $z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{\hat{p}\hat{q}}{n}}}$	
			•[3] calculate z	•[3] $= \dfrac{\dfrac{37}{64} - 0\cdot 5}{\sqrt{\dfrac{\dfrac{37}{64}\dfrac{27}{64}}{64}}} = 1\cdot 26$	
			•[4] appropriate conclusion	•[4] $1\cdot 48 < 1\cdot 64$ so we cannot reject H_0 at the 5% level	
			•[5] correct interpretation	•[5] conclude there is no evidence for the fan's claim	

Notes:

The PvA approach would record that $P(Z \geq 1\cdot 26) = 0\cdot 1038 > 0\cdot 05$

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
5.	(a)		•[1] correct mean and variance	•[1] $\dfrac{(22+7)}{2} = 14\cdot 5 \quad \dfrac{(22-7)^2}{12} = 18\cdot 75$	**3**
			•[2] appropriate strategy	•[2] $P(\bar{X} > 16\cdot 7) \approx P\left(Z > \dfrac{16\cdot 7 - 14\cdot 5}{\sqrt{\dfrac{18\cdot 75}{25}}}\right)$	
			•[3] calculate the probability	•[3] $= 0\cdot 0055$	
	(b)		•[4] appropriate comment	•[4] for $n \geq 20$ by the CLT	**1**

Notes:

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
6.	(a)		•[1] appropriate comment •[2] appropriate graph	•[1] categorical data •[2] bar chart	2
Notes: Other alternatives may be acceptable eg qualitative					
Commonly Observed Responses:					
	(b)		•[3] appropriate assumption •[4] correct strategy •[5] correct mean and sd •[6] correct t •[7] calculate interval	•[3] energy produced is normally distributed •[4] a 95% CI is given by $\bar{x} \pm t_{n-1,0\cdot975}\frac{s}{\sqrt{n}}$ •[5] $\bar{x} = 2599\cdot5$ and $s = 313\cdot22$ •[6] $t_{9,0\cdot975} = 2\cdot262$ •[7] $(2376, 2824)$	5
Notes:					
Commonly Observed Responses:					
	(c)		•[8] appropriate reason •[9] appropriate comment	•[8] 2660 and 2820 lie within the CI and so •[9] there is evidence that the species is Birch or Maple	2
Notes:					
Commonly Observed Responses:					
	(d)		•[10] correct t •[11] calculate interval •[12] appropriate comment	•[10] $t_{9,0\cdot995} = 3\cdot250$ •[11] $(2278, 2921)$ •[12] an increase in confidence requires a wider interval so that the species could now be Elm, Maple, Birch or Pine.	3
Notes:					
Commonly Observed Responses:					

Question			Generic scheme	Illustrative scheme	Max mark
7.	(a)		•[1] correct target value •[2] orrect $\hat{\sigma}$ •[3] calculate limits	•[1] $\overline{\overline{X}}=\frac{263\cdot4}{8}=29\cdot55$ •[2] $\overline{R}=\frac{18}{8}=2\cdot25,\ \hat{\sigma}=\frac{2\cdot25}{2\cdot534}=0\cdot8879$ •[3] $3\sigma=29\cdot55\pm\frac{3\times0\cdot8879}{\sqrt{6}}$ $=[28\cdot46, 30\cdot64]$	3
Notes:					
Commonly Observed Responses:					
	(b)		•[4] correct strategy •[5] correct values •[6] calculate d	•[4] $UL=\overline{\overline{X}}+3\frac{\frac{\overline{R}}{d}}{\sqrt{n}}$ •[5] $30\cdot71=29\cdot92+3\frac{\frac{2\cdot35}{d}}{\sqrt{9}}$ •[6] $d=\frac{2\cdot35}{(30\cdot71-29\cdot92)}$ $=2\cdot975$	3
Notes:					
Commonly Observed Responses:					

Question			Generic scheme	Illustrative scheme	Max mark
8.	(a)		•1 appropriate comment	•1 positive correlation between number of weeks of gestation and IQ score at age 12.	**1**

Notes:

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
	(b)		•2 correct R^2 •3 appropriate comment	•2 $R^2 = \frac{555 \cdot 0811^2}{531 \cdot 5676 \times 1731 \cdot 2973} = 0 \cdot 3348$ •3 $33 \cdot 48\%$ of the total variation in IQ score is explained by the linear model	**2**

Notes:
Other comments may be acceptable eg the association does not appear to be strong

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
	(c)		•4 appropriate hypotheses •5 calculate test statistic •6 correct cv •7 deal with H_0 •8 appropriate conclusion •9 correct assumption	•4 $H_0 : \rho = 0$, $H_1 : \rho \neq 0$ •5 $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = 4 \cdot 197$ •6 $t_{35,0 \cdot 995} = 2 \cdot 724$ •7 since $4 \cdot 197 > 2 \cdot 724$ we reject H_0 at the 1% level of significance and •8 conclude that there is evidence of a linear association between gestation and IQ •9 the 37 children are independent	**6**

Notes:
For mark 9 mention of the bivariate normal distribution in context is acceptable
The PvA approach would record that $2 \times \mathrm{P}\left(t_{35} \geq 4 \cdot 197\right) = 2 \times 0 \cdot 000088 = 0 \cdot 000176 < 0 \cdot 01$

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
8.	(d)		•[10] appropriate suggestion •[11] appropriate comment	•[10] scatter plot of residuals against the fitted values •[11] the residual plot should show points randomly scattered, centred on zero with a constant variance	**2**

Notes:

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
9.			•1 appropriate assumption	•1 the distribution of exam marks is symmetrical	9
			•2 appropriate hypotheses	•2 $H_0 : \eta = 65 \;\; H_1 : \eta < 65$	
			•3 deal with the data	•3 ignore 65 $n = 21$	
			•4 correct value of *W*	•4 $W = 2{\cdot}5 + 6 + 8 + 10 + 13 = 39{\cdot}5$	
			•5 correct parameters	•5 $E(W) = 115{\cdot}5 \quad V(W) = 827{\cdot}75$	
			•6&7 calculate z with correct continuity correction	•6&7 $z = \frac{40 - 115{\cdot}5}{\sqrt{827{\cdot}75}} = -2{\cdot}62$	
			•8 deal with H_0	•8 $-2{\cdot}62 < -2{\cdot}33$ so we reject H_0 at the 1% level of significance and	
			•9 appropriate conclusion	•9 conclude that there is strong evidence of poorer performance the following year	

Notes:

The PvA would record that $P(Z \le -2{\cdot}62) = 0{\cdot}0043 < 0{\cdot}01$. . . . .

For •8, 5% level would be acceptable.

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
10.	(a)		•[1] correct hypothesis •[2] correct z •[3] correct critical values •[4] appropriate conclusion •[5] appropriate conclusion	•[1] $H_0: \mu = 50$ $H_1: \mu < 50$ •[2] $= \dfrac{48{\cdot}3 - 50}{\frac{4}{5}} = -2{\cdot}13$ •[3] 1·64 and 2·33 •[4] $2{\cdot}13 > 1{\cdot}64$, so at the 5% level of significance, there is evidence that the wingspan of the species has decreased •[5] but there is no such evidence at the 1% level since $2{\cdot}13 < 2{\cdot}33$	**5**

Notes:

The PvA approach would record that $P(Z \le 2{\cdot}13) = 0{\cdot}0166$

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
	(b)		•[6] appropriate strategy •[7] calculate b for 2 levels •[8&9] appropriate summary	•[6] $\dfrac{\overline{X} - 50}{\frac{4}{5}} = -1{\cdot}64$ •[7] 5% value of $b = 48{\cdot}688$ 1% value of $b = 48{\cdot}136$ •[8&9] Only if a sample mean is below 48·7 cm is there is some evidence of a reduction in wingspan, but if it is below 48·1 cm then the evidence is much stronger	**4**

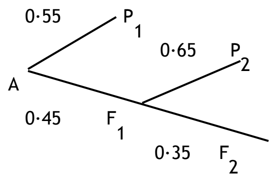
Notes:

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
	(c)		•[10] appropriate information •[11] appropriate justification •[12] appropriate choice of test	•[10] the sample variance would be calculated •[11] and used to estimate σ^2 •[12] in a t-test	**3**

Notes:

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
11.	(a)		•[1] appropriate strategy •[2] correct proportion for A	•[1] tree diagram annotated clearly  •[2] A: $0\cdot55+0\cdot45(0\cdot65)=0\cdot8425$	**2**
Notes:					
Commonly Observed Responses:					
	(b)	(i)	•[3] correct proportion for B •[4] appropriate conclusion	•[3] B: 52 + 35 = 87% •[4] Yes: 87 > 84·25%	**4**
		(ii)	•[5] appropriate strategy •[6] correct percentage	•[5] $P(P_2 \mid F_1)=\dfrac{P(P_2 \cap F_1)}{P(F_1)}$ •[6] $=\dfrac{0\cdot35}{0\cdot48}=72\cdot9\%$	
Notes:					
Commonly Observed Responses:					
	(c)		•[7] appropriate strategy •[8] correct calculation •[9] calculate probability	•[7] factors of $\frac{2}{3}$ and $\frac{1}{3}$ now apply •[8] $P(B \mid F_2)=\dfrac{0\cdot13\times\frac{2}{3}}{0\cdot13\times\frac{2}{3}+0\cdot1575\times\frac{1}{3}}$ •[9] $\dfrac{104}{167}=0\cdot6228$	**3**
Notes:					
Commonly Observed Responses:					

Question			Generic scheme	Illustrative scheme	Max mark
12.	(a)		•[1] correct values of X •[2&3] correct probabilities •[4] calculate $E(X)$ •[5] calculate $SD(X)$	•[1-3] X: 99, 9, 0, –1 $P(X)$: 0·0020, 0·0410, 0·2871, 0·6699 •[4] $E(X) = -0\cdot1029$ dollars •[5] $SD(X) = 4\cdot8562$ dollars	**5**

X	99	9	0	–1
P(X)	0·0020	0·0410	0·2871	0·6699

Notes:
Evidence of working required for •[4] and •[5]

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
	(b)	(i)	•[6] calculate expected value •[7] correct strategy •[8] calculate standard deviation •[9] appropriate assumption	•[6] $60(-0\cdot1029) + 45(-0\cdot06) = -8\cdot874$ •[7] $60\,V(X) + 45\,V(Y)$ •[8] $= 60(23\cdot58) + 45(400) = 19\ 415$ and $\sqrt{19\ 415} \approx 139$ •[9] assuming all games played are independent	**5**
	(b)	(ii)	•[10] appropriate comment	•[10] on average a gambler can expect to lose about 9 dollars but with very high variability anything is possible	

Notes:
Comment on both average and variability is required for •[10]
The use of exact values is acceptable $(-9\cdot145 \text{ and } 139\cdot239)$

Commonly Observed Responses:

[END OF MARKING INSTRUCTIONS]