National
Qualifications
2022

2022 Statistics

Advanced Higher - Paper 2

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 × 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 $x = -4$
$\bullet^6$ $y = 5$ 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

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

(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

<table>
<tr><th colspan="3">Question</th><th>Generic scheme</th><th>Illustrative scheme</th><th>Max mark</th></tr>
<tr><td>1.</td><td></td><td></td><td>•1 appropriate hypothesis</td><td>•1 H_0: prevalence of infection is not associated with the sex of the fish
H_1: there is an association between prevalence of infection and sex</td><td>6</td></tr>
<tr><td></td><td></td><td></td><td>•2 calculate E_i</td><td>•2 Expected frequencies are<table><tr><th></th><th>Male</th><th>Female</th></tr><tr><td>Yes</td><td>104.03</td><td>100.97</td></tr><tr><td>No</td><td>370.97</td><td>360.03</td></tr></table></td><td></td></tr>
<tr><td></td><td></td><td></td><td>•3 calculate test statistic</td><td>•3 $X^2 = 19.6$</td><td></td></tr>
<tr><td></td><td></td><td></td><td>•4 correct critical value</td><td>•4 $\chi^2_{1,0.990} = 6.635$</td><td></td></tr>
<tr><td></td><td></td><td></td><td>•5 deal with H_0</td><td>•5 $19.6 > 6.635$ so we reject H_0 (at the 1% level of significance)</td><td></td></tr>
<tr><td></td><td></td><td></td><td>•6 appropriate conclusion</td><td>•6 and conclude that there is (strong) evidence of an association between the prevalence of infection and sex</td><td></td></tr>
</table>

Notes:

1. For •1, at least one of the hypotheses must include the context of problem, i.e. the prevalence of infection and the sex of the fish.
2. For •2, also accept values rounded to at least 3 significant figures.
3. Alternatively, •4 p-value = 0.000009 and •5 '$0.000009 < 0.01$ so we reject H_0'.
4. 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:

Question			Generic scheme	Illustrative scheme	Max mark
2.	(a)		$\bullet^1$ state distribution	$\bullet^1$ $X \sim \text{Po}(2.3)$	**1**

Notes:

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
	(b)		$\bullet^2$ calculate probability	$\bullet^2$ $\text{P}(X=0)=0.1003$	**1**

Notes:

1. For $\bullet^2$, also accept values equivalent to rounding to 3 significant figures. i.e. 0.1, 0.10, 0.100

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
	(c)		$\bullet^3$ correct strategy $\bullet^4$ calculate probability	$\bullet^3$ $\text{P}(X=2)\text{P}(Y=2)$ $\bullet^4$ $0.2652 \times 0.2640 = 0.0700$	**2**

Notes:

1. For $\bullet^4$, also accept values equivalent to rounding to 3 significant figures. i.e. 0.07, 0.070

Commonly Observed Responses:

1. If values for $\text{P}(X=2)$ and $\text{P}(Y=2)$ calculated separately, but then nothing done with them, then no marks awarded.
2. If Po(4) used and $\text{P}(X+Y=4)$ calculated, then no marks awarded.
3. If candidates calculate $\text{P}(X=2)+\text{P}(Y=2)=0.5292$, then only gain $\bullet^4$.

Question			Generic scheme	Illustrative scheme	Max mark
	(d)		$\bullet^5$ correct distribution $\bullet^6$ correct strategy $\bullet^7$ calculate probability	$\bullet^5$ $X+Y \sim \text{Po}(4)$ $\bullet^6$ $\begin{cases}\text{P}(X+Y>5)\\ =1-\text{P}(X+Y\le 5)\end{cases}$ $\bullet^7$ 0.2149	**3**

Notes:

Commonly Observed Responses:

1. If candidates use a normal approximation distribution, then no marks awarded.
2. $\text{P}(X+Y>5)=1-\text{P}(X+Y\le 4)=1-0.6288=0.3712$, gains marks $\bullet^5$ and $\bullet^7$ only.

<table>
<tr><th colspan="3">Question</th><th>Generic scheme</th><th>Illustrative scheme</th><th>Max mark</th></tr>
<tr><td>3.</td><td></td><td></td><td>•¹ calculate probabilities

•² calculate E(T)

•³ correct strategy

•⁴ calculate V(T)</td><td>•¹
<table><tr><td>t</td><td>0</td><td>2</td><td>4</td><td>6</td><td>8</td></tr><tr><td>P(T = t)</td><td>0.1</td><td>0.2</td><td>0.4</td><td>0.2</td><td>0.1</td></tr></table>
•² $\text{E}(T) = 4$

•³ $\text{V}(T) = \text{E}(T^2) - \text{E}^2(T)$

•⁴ $\text{V}(T) = 4.8$</td><td>4</td></tr>
<tr><td colspan="6">Notes:</td></tr>
<tr><td colspan="6">Commonly Observed Responses:

1. If table used is:
<table><tr><td>t</td><td>0</td><td>2</td><td>4</td><td>6</td><td>8</td></tr><tr><td>P(T = t)</td><td>0.4</td><td>0.2</td><td>0.4</td><td>0</td><td>0</td></tr></table>which gives $\text{E}(T) = 2$ and $\text{V}(T) = 3.2$ then award •², •³ and •⁴

2. If table used is:
<table><tr><td>t</td><td>0</td><td>2</td><td>4</td><td>6</td><td>8</td></tr><tr><td>P(T = t)</td><td>0.2</td><td>0.2</td><td>0.2</td><td>0.2</td><td>0.2</td></tr></table>which gives $\text{E}(T) = 4$ and $\text{V}(T) = 8$ then award •², •³ and •⁴</td></tr>
</table>

<table>
<tr><th colspan="3">Question</th><th>Generic scheme</th><th>Illustrative scheme</th><th>Max mark</th></tr>
<tr><td>4.</td><td>(a)</td><td></td><td>•1 state distribution</td><td>•1 $X \sim \text{B}(104, 0.44)$</td><td>3</td></tr>
<tr><td></td><td></td><td></td><td>•2 use distribution</td><td>•2 $\text{P}(X = 52)$</td><td></td></tr>
<tr><td></td><td></td><td></td><td>•3 calculate probability</td><td>•3 0.0367</td><td></td></tr>
<tr><td colspan="6">Notes:
1. For •3, accept answer rounded to 3 decimal places</td></tr>
<tr><td colspan="6">Commonly Observed Responses:
1. If a Normal approximation to a Binomial distribution is used, with continuity correction, then probability = 0.0369, and marks •1, •2 and •3 can be awarded.
2. If a Poisson approximation to a Binomial distribution is used, then do not award •1, but marks •2 and •3 remain available.</td></tr>
<tr><td></td><td>(b)</td><td></td><td>•4 state approximate distribution</td><td>•4 $X \approx \text{N}(45.76, 25.6256)$</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td>•5 use continuity correction</td><td>•5 $\text{P}(39.5 < X < 50.5)$</td><td></td></tr>
<tr><td></td><td></td><td></td><td>•6 calculate z values</td><td>•6 $\text{P}(-1.24 < Z < 0.94)$</td><td></td></tr>
<tr><td></td><td></td><td></td><td>•7 calculate probability</td><td>•7 0.7189</td><td></td></tr>
<tr><td colspan="6">Notes:
1. For •4, omission of double-tilde notation is not penalised, as the question instructed candidates to use an approximation.</td></tr>
<tr><td colspan="6">Commonly Observed Responses:
1. If no continuity correction applied, then $\text{P}(-1.137 < Z < 0.837) = 0.6709$. Award •4 •6 •7 only.
2. If exact Binomial calculation performed, then probability = 0.7179. Award •7 only.
3. If Poisson approximation used, then probability = 0.5841. Award •7 only.</td></tr>
</table>

Question			Generic scheme	Illustrative scheme	Max mark
5.	(a)		•^{1} state reason	•^{1} we have paired data	**1**
	(b)		•^{2} show strategy •^{3} state hypotheses •^{4} calculate $\overline{x}_d$ and s_d •^{5} calculate t •^{6} state critical value •^{7} deal with H$_0$ •^{8} write conclusion •^{9} appropriate assumption	•^{2} $t_{n-1} = \frac{\overline{x}_d}{s_d / \sqrt{n}}$ •^{3} $\text{H}_0 : \mu_d = 0 \quad \text{H}_1 : \mu_d \neq 0$ •^{4} $\overline{x}_d = 2.44 \quad s_d = 2.92$ •^{5} $t = \frac{2.44}{2.92 / \sqrt{9}} = 2.507$ •^{6} $t_{8,0.975} = 2.306$ •^{7} $2.507 > 2.306$ so we reject H$_0$ at the 5% level of significance •^{8} and conclude that there **is** evidence that the mean of differences in performance between French and German is non-zero. •^{9} the differences are normally distributed	**8**

Question		Generic scheme	Illustrative scheme	Max mark
5.	(b)	**(continued)**		

Notes:

1. For $\bullet^1$, do not accept any references to independence or normality.
2. For $\bullet^3$, hypotheses must clearly reference ‘differences of means’.
3. For $\bullet^3$, also accept $H_0 : \mu_F = \mu_G \quad H_1 : \mu_F \neq \mu_G$ (as $\mu_F = \mu_G \Leftrightarrow \mu_F - \mu_G = 0 \Leftrightarrow \mu_{F-G} = 0$).
4. Alternatively, $\bullet^6$ p-value = 0.0363 and $\bullet^7$ $0.0363 < 0.05$ so we reject H_0 at the 5% level of significance.
5. To gain $\bullet^7$, there must be clear communication somewhere in the solution of the level of significance chosen, as it was not given in the question.
6. For $\bullet^7$ and $\bullet^8$, also accept other levels of significance, only if the logic is consistent.
7. For $\bullet^8$, response must reference the ‘mean of differences’.
8. For $\bullet^8$, do not accept conclusions that are too definite. Phrasing must include ‘evidence to conclude…’, or ‘evidence to suggest…’, or similar.
9. For $\bullet^9$, response must reference ‘differences’.

Question		Generic scheme	Illustrative scheme	Max mark
5.	(b)	**(continued)**		

Commonly Observed Responses:

Candidate A – performed two sample t-test for a difference in population means (non-paired data)
Mark $\bullet^2$ not available.
Mark $\bullet^4$ requires $\bar{x}_F = 65.78$ $s_F = 15.81$ $\bar{x}_G = 63.33$ $s_F = 16.73$
Mark $\bullet^5$ requires $t = 0.318$
Mark $\bullet^6$ requires $t_{16,0.975} = 2.120$
Marks $\bullet^7$, $\bullet^8$, $\bullet^9$ available only if consistent with previous workings

Candidate B – performed two sample z-test for a difference in population means (non-paired data)
Marks $\bullet^2$ and $\bullet^5$ not available.
Mark $\bullet^4$ requires $\bar{x}_F = 65.78$ $s_F = 15.81$ $\bar{x}_G = 63.33$ $s_F = 16.73$
Mark $\bullet^6$ requires $z_{0.975} = 1.96$
Marks $\bullet^7$, $\bullet^8$, $\bullet^9$ available only if consistent with previous workings

Candidate C – performed one-sample z-test on mean differences (paired data)
Marks $\bullet^2$ and $\bullet^5$ not available.
Mark $\bullet^4$ requires $\bar{x}_d = 2.44$ $s_d = 2.92$
Mark $\bullet^6$ requires $z_{0.975} = 1.96$
Marks $\bullet^7$, $\bullet^8$, $\bullet^9$ available only if consistent with previous workings

Candidate D – performed Mann-Whitney test (non-paired data)
Marks $\bullet^2$ and $\bullet^5$ not available.
Mark $\bullet^4$ requires rank sum, $W = 82$
Mark $\bullet^6$ requires critical value of 62 (for 5%)
Marks $\bullet^7$, $\bullet^8$, $\bullet^9$ available only if consistent with previous workings, and must mention ‘medians’ rather than ‘means’

Candidate E – performed Wilcoxon Signed-Rank Test (paired data)
Mark $\bullet^2$ not available.
Mark $\bullet^4$ requires rank sum, $W = 4$
Mark $\bullet^5$ requires correct omission of zero difference
Mark $\bullet^6$ requires critical value of 3 (for 5% with n=8)
Marks $\bullet^7$, $\bullet^8$, $\bullet^9$ available only if consistent with previous workings, and must mention ‘medians’ rather than ‘means’

Question			Generic scheme	Illustrative scheme	Max mark
6.	(a)		$\bullet^1$ appropriate description	$\bullet^1$ there is a distinct U-shaped pattern to the residuals	**2**
			$\bullet^2$ appropriate description	$\bullet^2$ non-constant variance	

Notes:

1. For $\bullet^1$, also accept ‘non-random scatter’.

Commonly Observed Responses:

1. Do not accept ‘mean of residuals being zero’.

Question			Generic scheme	Illustrative scheme	Max mark
	(b)		$\bullet^3$ calculate b-value	$\bullet^3$ $b=\frac{S_{xw}}{S_{xx}}$ $=\frac{-715.456}{51170}$ $=-0.01398$	**4**
			$\bullet^4$ calculate mean values and a value	$\bullet^4$ $\bar{w}=\frac{101.2529}{85}=1.1912$ $\bar{x}=\frac{3740}{85}=44$ $a=\bar{w}-b\bar{x}$ $=1.1912-44(-0.01398)$ $=1.806$	
			$\bullet^5$ appropriate regression line	$\bullet^5$ $w=1.81-0.014x$	
			$\bullet^6$ predicted percentage	$\bullet^6$ $y=10^{1.81-0.014\times 87}=3.9\%$	

Notes:

1. For $\bullet^3$, value must be stated to at least 2 significant figures.
2. For $\bullet^5$, also accept $\log y=1.81-0.014x$ and $\log_{10} y=1.81-0.014x$.

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
7.	(a)		$\bullet^1$ correct method	$\bullet^1$ $\mathrm{E}\left(\frac{X}{n}\right)=\frac{1}{n}\mathrm{E}(X)=\frac{1}{n}np=p$	**2**
			$\bullet^2$ correct method	$\bullet^2$ $\begin{cases}\mathrm{V}\left(\frac{X}{n}\right)=\frac{1}{n^2}\mathrm{V}(X)\\ \qquad =\frac{1}{n^2}npq\\ \qquad =\frac{pq}{n}\end{cases}$	

Notes:

1. For either mark, do not accept any methods that ‘work backwards from the answer’.
2. For $\bullet^1$, the minimum response must contain $\mathrm{E}\left(\frac{X}{n}\right)=\frac{1}{n}np$.
3. For $\bullet^2$, the minimum response must contain $\mathrm{V}\left(\frac{X}{n}\right)=\frac{1}{n^2}npq$.

Commonly Observed Responses:

1. For $\bullet^2$, $\mathrm{V}\left(\frac{X}{n}\right)=\frac{1}{n}\mathrm{V}(X)=\frac{1}{n^2}npq$ gains no marks.

Question			Generic scheme	Illustrative scheme	Max mark
	(b)		$\bullet^3$ correct verification	$\bullet^3$ $np=14$ and $nq=36$, both >5	**4**
			$\bullet^4$ appropriate strategy	$\bullet^4$ $\hat{p}\pm z\sqrt{\frac{\hat{p}\hat{q}}{n}}$	
			$\bullet^5$ substitute	$\bullet^5$ $0.28\pm2.58\sqrt{\frac{0.28\times0.72}{50}}$	
			$\bullet^6$ calculate interval	$\bullet^6$ $(0.116, 0.444)$	

Notes:

1. For $\bullet^3$, both np and nq must be evaluated and compared to 5.
2. For $\bullet^6$, omission of brackets not penalised (considered to be ‘bad form’).

Commonly Observed Responses:

1. If a z-interval is correctly calculated to be (12.84, 15.15) then do not award $\bullet^4$.

Question			**Generic scheme**	**Illustrative scheme**	**Max mark**
8.	(a)		•^{1} correct strat8egy	•^{1} $\frac{1}{5}\times\ldots$	**2**
			•^{2} calculate probability	•^{2} $\frac{1}{5}\cdot\frac{5}{8}=\frac{1}{8}$	
Notes:					
	(b)	(i)	•^{3} calculate probability	•^{3} $\frac{2}{5}$	**1**
		(ii)	•^{4} correct strategy	•^{4} $P(L)=P(1\cap S)+P(4\cap S)$	**3**
			•^{5} substitute probabilities	•^{5} $\frac{1}{5}\cdot\frac{2}{5}+\frac{1}{5}\cdot\frac{3}{8}$	
			•^{6} calculate probability	•^{6} $\frac{31}{200}$	
		(iii)	•^{7} correct strategy	•^{7} $P(1\mid L)=\frac{P(L\mid 1)P(1)}{P(L)}$	**3**
			•^{8} substitute probabilities	•^{8} $\dfrac{\frac{2}{5}\cdot\frac{1}{5}}{\frac{31}{200}}$	
			•^{9} calculate probability	•^{9} $\frac{16}{31}$	
Notes: 1. Other methods are acceptable in (ii) and (iii) eg tree diagrams. 2. Mark •^{4} can be awarded by implication from mark •^{5}. 3. Mark •^{7} can be awarded by implication from mark •^{8}.					
Commonly Observed Responses:					

Question			Generic scheme	Illustrative scheme	Max mark
9.	(a)		$\bullet^1$ appropriate description $\bullet^2$ appropriate description	$\bullet^1$ the distribution of the sample mean is approximately normal… $\bullet^2$ … irrespective of the distribution of the population.	**2**

Notes:

1. For $\bullet^1$, must reference 'distribution' and 'sample mean' and 'approximately normal'.
2. For $\bullet^1$, also accept $\bar{X} \approx \mathrm{N}(....)$.

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
	(b)		$\bullet^3$ correct test statistic $\bullet^4$ correct critical value $\bullet^5$ deal with H_0 $\bullet^6$ appropriate conclusion $\bullet^7$ appropriate assumption	$\bullet^3$ $z = \frac{52.6 - 50}{\sqrt{\frac{103.25}{45}}} = 1.71646$ $\bullet^4$ 5% cv is 1.64 $\bullet^5$ $1.72 > 1.64$ so we reject H_0 at the (5% level of significance) $\bullet^6$ conclude that there is evidence the mean width of the battens is more than 50mm $\bullet^7$ population variance is well approximated by the sample variance	**5**

Notes:

1. Alternatively for $\bullet^4$ *p*-value = 0.043 and $\bullet^5$ $0.043 < 0.05$ so we reject H_0.
2. For $\bullet^6$, do not accept conclusions that are too definite. Phrasing must include 'evidence to conclude…', or 'evidence to suggest…', or similar.
3. For $\bullet^7$, also accept 'population variance is the same as the sample variance'.

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
10.	(a)		•¹ correct hypothesis	•¹ $H_0 : \rho = 0 \;\; H_1 : \rho \neq 0$	**7**
			•² correct value of r	•² $r = \frac{46.29}{\sqrt{278.61 \times 10.95}} = \; 0.838$	
			•³ calculate test statistic	•³ $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = 3.071$	
			•⁴ correct cv	•⁴ $t_{4,\,0.975} = 2.776$	
			•⁵ deal with H_0	•⁵ $3.071 > 2.776$ so we reject H_0 (at the 5% level of significance)	
			•⁶ appropriate conclusion	•⁶ conclude that the evidence suggests a (positive) linear association between the exposure index and the number of related deaths per 10 000 of the population	
			•⁷ appropriate assumption	•⁷ assuming the pairs of observations are independent	

Notes:

1. For •¹, also accept hypothesis phrased in terms of correlation, in the context.
2. For •¹, do not accept ‘H_0: … are linearly associated’.
3. Alternatively, •⁴ p-value = 0.0372.
4. For •⁶, do not accept conclusions that are too definite. Phrasing must include ‘evidence to conclude…’, or ‘evidence to suggest…’, or similar.
5. For •⁷, also accept that variables are distributed with an (approximately) bivariate normal distribution.
6. For •⁷, also accept ‘all towns are independent’.

Commonly Observed Responses:

Candidate A – performed hypothesis test on β
Mark •¹ not available.
Mark •² requires b = 0.1661 and s = 0.9026
Mark •⁶ requires reference to the slope parameter

	(b)		•⁸ appropriate suggestion	•⁸ correlation is not causation	**1**

Notes:

1. For •⁸ accept other explanations such as deaths may be from other causes.

Commonly Observed Responses:

<table>
<tr><th colspan="3">Question</th><th>Generic scheme</th><th>Illustrative scheme</th><th>Max mark</th></tr>
<tr><td>11.</td><td>(a)</td><td>(i)</td><td>•¹ correct drawing, with appropriate key

•² write comment</td><td>•¹
<table>
<tr><th>Adults</th><th></th><th>Juveniles</th></tr>
<tr><td>3</td><td>0</td><td>7 9</td></tr>
<tr><td>5 3</td><td>1</td><td>1 1 3 7 9</td></tr>
<tr><td>9 7 3 2</td><td>2</td><td>8 9</td></tr>
<tr><td>5 5</td><td>3</td><td></td></tr>
<tr><td>0</td><td>4</td><td>1</td></tr>
</table>
Key $1|3 = 1.3$ seconds

•² it appears that juveniles react quicker to loud music</td><td>2</td></tr>
<tr><td></td><td></td><td>(ii)</td><td>•³ appropriate hypothesis

•⁴ correct critical value

•⁵ deal with H_0

•⁶ write conclusion</td><td>•³ $H_0: \eta_a = \eta_j \quad H_1: \eta_a \neq \eta_j$

•⁴ 5% cv is 78

•⁵ $89 > 78$ so we cannot reject H_0 at the 5% level of significance

•⁶ and conclude that there is no evidence of a difference in median reaction times (between adult and juvenile foxes)</td><td>4</td></tr>
</table>

Notes:

1. For •¹, do not penalise use of commas between digits (considered to be 'bad form').
2. For •¹, do not penalise lack of units in the key (considered to be 'bad form').
3. For •¹, leaves must be in descending order, away from the stem.
4. For •¹, also accept 'double stemmed' diagram, exemplified below:

Adults		Juveniles
3	0	
	0	7 9
3	1	1 1 3
5	1	7 9
3 2	2	
9 7	2	8 9
	3	
5 5	3	
0	4	1
	4	

5. Mark •² is only available for a contextual comment based on the stem-and-leaf diagram.
6. Mark •⁴ is not available if performed (an invalid) normal approximation to Mann-Whitney.
7. For •⁴, also accept 10% critical value is 82, or 2% critical value is 74.
7. To gain •⁵, there must be clear communication somewhere in the solution of the level of significance chosen, as it was not given in the question.
8. For •⁶, do not accept conclusions that are too definite. Phrasing must include 'evidence to conclude…', or 'evidence to suggest…', or similar.
9. For •⁶, response must reference '…different in median….'

Commonly Observed Responses:

Question			Generic scheme	Illustrative scheme	Max mark
	(b)		•⁷ find difference of random variables	•⁷ $D=A-J$	**5**
			•⁸ correct μ and σ^2	•⁸ $\text{E}(D)=0.5$, $\text{V}(D)=0.8$	
			•⁹ correct strategy	•⁹ $\text{P}(D>0)$	
			•¹⁰ continue strategy	•¹⁰ $\text{P}\left(Z>\frac{0-0.5}{\sqrt{0.8}}\right)=\text{P}(Z>-0.56)$	
			•¹¹ calculate probability	•¹¹ 0.7123	

Notes:

1. For •⁷, also accept calculating $D=J-A$ (which then impacts on direction of inequality in marks •⁹ and •¹⁰).
2. Mark •¹¹ can only be awarded (as a follow through) for the calculation of a probability from a normal distribution that does not have a variance to be either 0.3 or 0.5.

Commonly Observed Responses:

1. If V(D)=0.2, then z-value = ±1.118. Do not award mark •⁸ and other marks remain available as follow through.

[END OF MARKING INSTRUCTIONS]