Pearson Edexcel

Mark Scheme (Provisional)

Summer 2021

Pearson Edexcel International Advanced Level
In Statistics S3 Paper WST03/01

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications are awarded by Pearson, the UK's largest awarding body. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information visit our qualifications websites at www.edexcel.com or www.btec.co.uk. Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

Pearson: helping people progress, everywhere

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk

Summer 2021
Question Paper Log number P63153A
Publications Code WST03_01_2106_MS

General Marking Guidance

- All candidates must receive the same treatment. Examiners must mark the first candidate in exactly the same way as they mark the last.
- Mark schemes should be applied positively. Candidates must be rewarded for what they have shown they can do rather than penalised for omissions.
- Examiners should mark according to the mark scheme not according to their perception of where the grade boundaries may lie.
- There is no ceiling on achievement. All marks on the mark scheme should be used appropriately.
- All the marks on the mark scheme are designed to be awarded. Examiners should always award full marks if deserved, i.e. if the answer matches the mark scheme. Examiners should also be prepared to award zero marks if the candidate's response is not worthy of credit according to the mark scheme.
- Where some judgement is required, mark schemes will provide the principles by which marks will be awarded and exemplification may be limited.
- When examiners are in doubt regarding the application of the mark scheme to a candidate's response, the team leader must be consulted.
- Crossed out work should be marked UNLESS the candidate has replaced it with an alternative response.

EDEXCEL IAL MATHEMATICS

General Instructions for Marking

1. The total number of marks for the paper is 75.

2. The Edexcel Mathematics mark schemes use the following types of marks:

- **M** marks: method marks are awarded for 'knowing a method and attempting to apply it', unless otherwise indicated.
- **A** marks: Accuracy marks can only be awarded if the relevant method (M) marks have been earned.
- **B** marks are unconditional accuracy marks (independent of M marks)
- Marks should not be subdivided.

3. Abbreviations

These are some of the traditional marking abbreviations that will appear in the mark schemes.

- bod – benefit of doubt
- ft – follow through
- the symbol $\surd$ will be used for correct ft
- cao – correct answer only
- cso - correct solution only. There must be no errors in this part of the question to obtain this mark
- isw – ignore subsequent working
- awrt – answers which round to
- SC: special case
- oe – or equivalent (and appropriate)
- dep – dependent
- indep – independent
- dp decimal places
- sf significant figures
- $\ast$ The answer is printed on the paper
- $\lfloor$ The second mark is dependent on gaining the first mark

4. All A marks are 'correct answer only' (cao.), unless shown, for example, as A1 ft to indicate that previous wrong working is to be followed through. After a misread however, the subsequent A marks affected are treated as A ft, but manifestly absurd answers should never be awarded A marks.

5. For misreading which does not alter the character of a question or materially simplify it, deduct two from any A or B marks gained, in that part of the question affected.

6. If a candidate makes more than one attempt at any question:
   - If all but one attempt is crossed out, mark the attempt which is NOT crossed out.
   - If either all attempts are crossed out or none are crossed out, mark all the attempts and score the highest single attempt.

7. Ignore wrong working or incorrect statements following a correct answer.

<table>
<tr><th>Question Number</th><th>Scheme</th><th>Notes</th><th>Marks</th></tr>
<tr><td rowspan="2">1(a)</td><td rowspan="2"><table>
<tr><td></td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td></tr>
<tr><td>C</td><td>3</td><td>2</td><td>1</td><td>5</td><td>4</td></tr>
<tr><td>M</td><td>2</td><td>4</td><td>1</td><td>5</td><td>3</td></tr>
<tr><td>d</td><td>1</td><td>-2</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td>d^2</td><td>1</td><td>4</td><td>0</td><td>0</td><td>1</td></tr>
</table></td><td>Attempt to rank at least 1 row with at least 3 correct. Allow reverse rankings.</td><td>M1</td></tr>
<tr><td>Attempt at d^2 for their rankings, can be implied by $\sum d^2 = 6$</td><td>M1</td></tr>
<tr><td></td><td>$\sum d^2 = 6$</td><td>Can be implied by correct answer. Must come from correct rankings.</td><td>A1</td></tr>
<tr><td></td><td>$r_s = 1 - \frac{6(6)}{5(24)}$</td><td>(dep on 1st M1) Use of correct formula with their $\sum d^2$</td><td>dM1</td></tr>
<tr><td></td><td>$r_s = 0.7$</td><td>0.7 o.e.
Must come from correct rankings.</td><td>A1</td></tr>
<tr><td></td><td></td><td></td><td>(5)</td></tr>
<tr><td>(b)</td><td>$H_0 : \rho = 0,\ H_1 : \rho > 0$</td><td>Both correct in terms of ρ or ρ_s. Must be compatible with their ranking.</td><td>B1</td></tr>
<tr><td></td><td>cv 0.9 or cr $r_s \geq 0.9$</td><td>0.9
sign should match H_1 or their r_s</td><td>B1</td></tr>
<tr><td></td><td>$r_s = 0.7$ does not lie in cr so do not reject H_0</td><td>Correct non-contextual statement e.g. “do not reject H_0”, “not in critical region”, “not significant”, “no positive correlation”.
$\lvert\text{test value}\rvert$ or $\lvert\text{cv}\rvert > 1$ award M0</td><td>M1</td></tr>
<tr><td></td><td>Data does not support plant biologist’s claim.</td><td>Correct conclusion in context. Must mention “biologist’s claim” o.e. or moisture and plant coverage. All previous marks in (b) must have been scored.</td><td>A1ft</td></tr>
<tr><td>SC</td><td colspan="2">For use of two-tailed test:
May score B0B1M1A0 for cv = 1(.000) and ‘not significant’ oe</td><td></td></tr>
<tr><td></td><td></td><td></td><td>(4)</td></tr>
<tr><td></td><td></td><td></td><td>Total 9</td></tr>
</table>

<table>
<tr><th>Question Number</th><th>Scheme</th><th>Notes</th><th>Marks</th></tr>
<tr><td>2</td><td>H_0 : Diet and health are independent (or not associated).
H_1 : Diet and health are not independent (or associated).</td><td>“diet” and “health” mentioned at least once. Use of correlation is B0.</td><td>B1</td></tr>
<tr><td></td><td>
<table>
<tr><th>EXPECTED</th><th>Good health</th><th>Poor health</th><th>Totals</th></tr>
<tr><td>Good diet</td><td>83.19</td><td>10.81</td><td>94</td></tr>
<tr><td>Poor diet</td><td>93.81</td><td>12.19</td><td>106</td></tr>
<tr><td>Totals</td><td>177</td><td>23</td><td>200</td></tr>
</table>
</td><td>Attempt $\frac{\text{RT}\times\text{CT}}{\text{GT}}$
with at least one correct to 1dp;
all correct to 1dp.</td><td>M1;
A1</td></tr>
<tr><td></td><td>
<table>
<tr><th>Observed</th><th>Expected</th><th>$\frac{(O-E)^2}{E}$</th><th>$\frac{O^2}{E}$</th></tr>
<tr><td>86</td><td>83.19</td><td>0.095</td><td>88.905</td></tr>
<tr><td>8</td><td>10.81</td><td>0.730</td><td>5.920</td></tr>
<tr><td>91</td><td>93.81</td><td>0.084</td><td>88.274</td></tr>
<tr><td>15</td><td>12.19</td><td>0.648</td><td>18.458</td></tr>
<tr><td colspan="2">Totals</td><td>1.557</td><td>201.557</td></tr>
</table>
</td><td>Attempt at $\frac{(O-E)^2}{E}$
or $\frac{O^2}{E}$ with their values with 2 correct or 2 correct f.t.
Allow 2sf for this mark.</td><td>M1</td></tr>
<tr><td></td><td>$\sum\frac{(O-E)^2}{E}$ or $\sum\frac{O^2}{E}-200=1.557$</td><td>awrt 1.6</td><td>A1</td></tr>
<tr><td></td><td>$\nu=(2-1)(2-1)=1$
$\chi^2_1(5\%)=3.841$</td><td>1 (may be implied)
3.841
NB: may see $\chi^2_3(5\%)=7.815$ for f.t.
NB: p-value 0.212 but scores B0B0 on its own</td><td>B1
B1f.t.</td></tr>
<tr><td></td><td>$\chi^2=1.557$ does not lie in cr so insufficient evidence to reject H_0</td><td>For correct non-contextual statement linking their test statistic and their cv.</td><td>M1</td></tr>
<tr><td></td><td>E.g. Diet and health are independent or
There is no association between diet and health or
The doctor’s belief is not supported by this data.</td><td>Dependent on a cv of 3.841 and 3rd M1
Correct conclusion in context with “diet” and “health” or “doctor”. Condone “connection” or “relationship” but not “correlation”.</td><td>A1ft</td></tr>
<tr><td></td><td></td><td></td><td>(9)</td></tr>
<tr><td></td><td></td><td></td><td>Total 9</td></tr>
</table>

Question Number	Scheme	Notes	Marks
3.			
(a)	$\overline{x} = \frac{1}{2}(11.52 + 13.75) = 12.635$	12.635 (may be implied by correct CI)	B1
		Use of 1.96	B1
	$\left(\frac{\sigma}{\sqrt{n}} =\right) \frac{13.75 - 12.635}{1.96} (= 0.56887...)$ $\left(\frac{\sigma}{\sqrt{n}} =\right) \frac{13.75 - 11.52}{2 \times 1.96} (= 0.56887...)$	For attempt at standard error (may be implied by awrt 0.569)	M1
		Use of 1.6449 or better (1.644853…from calc) Use of 1.64 or 1.65 is B0	B1
	$12.635 \pm 1.6449 \times 0.56887...$	For $(\text{their } \overline{x}) \pm (\text{their } 1.6449)\left(\text{their } \frac{\sigma}{\sqrt{n}}\right)$	M1
	90% CI is $(11.699..., 13.5707...)$	awrt $(11.7, 13.6)$ from correct working Correct answer with no working scores B1B1M1B0M1A1	A1
			(6)
(b)	$4 \times 0.9^3 \times 0.1$ $= 0.2916$	$4p^3(1-p)$ (where $0 < p < 1$) awrt 0.292	M1 A1
			(2)
			Total 8

Question Number	Scheme	Notes	Marks
4. (a)	Label academic (1-1680) and vocational (1-2520)	For numbering/labelling/ordering (o.e.) students in **each** group	B1
	Use **random** numbers to select from each group.	For use of random sample/numbers/selection	B1
	28 academic and 42 vocational	Both numbers correct with the associated group	B1
			(3)
(b)	$H_0 : \mu_v - \mu_a = 0$ $H_1 : \mu_v - \mu_a > 0$	If the hypotheses are given in terms of $\mu_a - \mu_b$, a and b must be defined.	B1
	$\text{se} = \sqrt{\frac{70}{80} + \frac{60}{50}}$	Correct attempt at se – condone slip in sample sizes.	M1
	$z = \frac{62-57}{\sqrt{\frac{70}{80} + \frac{60}{50}}}$	Dep on previous M1 standardising with (62 – 57) and their se (Allow $\pm$)	dM1
	$z = 3.471...$ (or probability of 0.0003)	awrt ± 3.47 (or awrt 0.0003)	A1
	cv $z = 1.6449$	Allow $\pm$ but signs must be compatible Or allow comparison with probability of 0.05	B1
	Reject H_0 / significant	Dependent on 2nd M1. A correct non-contextual statement based on their normal cv and their test statistic.	dM1
	There is evidence that the **mean** basic skills score for **vocational** students is greater than the **mean** basic skills score for **academic** students.	Correct comment in context. Must mention "mean", "academic" and "vocational". Allow f.t. on their normal cv and their test statistic.	A1f.t.
			(7)
(c)	Mean / $\bar{X}_a$ (basic skills) score for academic students and mean/ $\bar{X}_v$ (basic skills) score for vocational students…	Must mention both means.	B1
	…have (approximately) a normal distribution as sample sizes are large.	Must mention normal.	B1
			(2)
(d)	Samples are (large enough) so that $s^2 = \sigma^2$	Must imply for both samples	B1
			(1)
(e)	Test no longer significant so insufficient evidence to reject H_0	Can be implied by correct comment in context.	M1
	Insufficient evidence that mean (basic skills) **score** for vocational students is greater than the mean (basic skills) **score** for academic students/There is no longer a difference in **scores**.	Must mention scores.	A1
			(2)
(f)	The course was **effective** (o.e.)	Dep on a significant result in (b) and a non-significant result in (e)	B1
			(1)
			Total 16

<table>
<tr><th>Question Number</th><th>Scheme</th><th>Notes</th><th>Marks</th></tr>
<tr><td>5.(a)</td><td>Relief of symptoms is either a “success” or a “failure”.
The probability the medicine being a success is constant.
Samples from different medical practices are independent.</td><td>Any 2. Context required in one assumption.</td><td>B1 B1</td></tr>
<tr><td></td><td></td><td></td><td>(2)</td></tr>
<tr><td>(b)</td><td>$\text{Mean} = \dfrac{0\times4+1\times6+2\times3+...+8\times2}{50} = 3.54*$</td><td>At least two correct terms on the numerator and 50 on the denominator,
fully correct expression or $\dfrac{177}{50}$
dep on M1 scored cso.</td><td>M1
A1cso</td></tr>
<tr><td></td><td></td><td></td><td>(2)</td></tr>
<tr><td>(c)</td><td>$p = \dfrac{3.54}{8} = 0.4425$</td><td>Can be implied by at least 1 correct value for f or g.</td><td>B1</td></tr>
<tr><td></td><td>$f = 50\times C_4^8 \times 0.4425^4 \times 0.5575^4 = 12.96$
$g = 50\times 0.4425^8 = 0.07$</td><td>Use of Bin(50, p) for M1,
Allow awrt 12.96, awrt 0.07</td><td>M1A1A1</td></tr>
<tr><td></td><td></td><td></td><td>(4)</td></tr>
<tr><td>(d)</td><td>H_0 : Binomial distribution is a suitable model
H_1 : Binomial distribution is not a suitable model</td><td>Both hypotheses correct
If parameters used then B0.</td><td>B1</td></tr>
<tr><td></td><td rowspan="2"><table>
<tr><th>No of successes</th><th>O</th><th>E</th><th>$\dfrac{(O-E)^2}{E}$</th><th>$\dfrac{O^2}{E}$</th></tr>
<tr><td>0,1,2</td><td>13</td><td>11.66</td><td>0.154</td><td>14.494</td></tr>
<tr><td>3</td><td>12</td><td>13.07</td><td>0.088</td><td>11.018</td></tr>
<tr><td>4</td><td>10</td><td>12.96</td><td>0.676/7</td><td>7.716</td></tr>
<tr><td>5,6,7,8</td><td>15</td><td>12.31</td><td>0.588/7</td><td>18.278</td></tr>
<tr><td></td><td>50</td><td>50</td><td>1.506</td><td>51.506</td></tr>
</table></td><td>Combining 0,1,2 or 5,6,7,8.</td><td>M1</td></tr>
<tr><td></td><td>M1 for attempting $\dfrac{(O-E)^2}{E}$ or $\dfrac{O^2}{E}$
with at least 2 correct expressions or 2 correct values to 2sf.</td><td>M1</td></tr>
<tr><td></td><td>$\sum\dfrac{(O-E)^2}{E} = \sum\dfrac{O^2}{E} - 50 = 1.50...$</td><td>awrt 1.5 (calculator: 1.50498…)</td><td>A1</td></tr>
<tr><td></td><td>$\nu = 4-2 = 2, \chi^2_2(10\%) = 4.605$</td><td>2 can be implied by 4.605 seen</td><td>B1B1f.t.</td></tr>
<tr><td></td><td>Insufficient evidence to reject H_0</td><td>For correct non-contextual statement linking their test statistic and their cv.</td><td>M1</td></tr>
<tr><td></td><td>Data is consistent with a binomial distribution (oe)</td><td>A correct comment suggesting that binomial model is suitable / good fit.
Hypotheses wrong way around scores A0 here. Condone parameters here.</td><td>A1</td></tr>
<tr><td></td><td></td><td></td><td>(8)</td></tr>
<tr><td></td><td></td><td></td><td>Total 16</td></tr>
</table>

Question Number	Scheme	Notes	Marks
6(a)	$W = B - 1.1R$	May be implied by correct mean or variance	M1
	$W \sim \mathrm{N}(55 - 1.1\times 51, 1.3^2 + 1.1^2 \times 1.2^2)$ or $W \sim \mathrm{N}(-1.1, 3.4324)$	$(\pm)1.1$, awrt 3.43 (may be seen in standardisation)	A1, A1
	$\mathrm{P}(W < 0) = \mathrm{P}\left(Z < \dfrac{0+1.1}{\sqrt{3.4324}}\right)$	Standardising with their mean and their sd. leading to a probability > 0.5	M1
	$= \mathrm{P}(Z < 0.5937...)$		
	$= 0.7224$ or 0.7237	**awrt 0.72**	A1
			(5)
(b)	$X = B_1 - B_2$	May be implied by correct mean or variance	M1
	$X \sim \mathrm{N}(55 - 55, 2\times 1.3^2)$ or $X \sim \mathrm{N}(0, 3.38)$	$0, 3.38$	A1, A1
	$\mathrm{P}\left(Z > \dfrac{1-0}{\sqrt{3.38}}\right)$ or $\mathrm{P}\left(Z < \dfrac{-1-0}{\sqrt{3.38}}\right)$	dep on 1st M1 for standardising with their mean and their sd.	dM1
	$\mathrm{P}(\lvert X\rvert > 1) = 2\times \mathrm{P}(X > 1)$	For $2\times$ seen or implied	M1
	$= 2\times \mathrm{P}(Z > 0.5439...)$ $= 2\times(1 - 0.7054)$	2×0.2946 or 2×0.2932 (calc)	
	$= 0.5892$	**awrt 0.59**	A1
			(6)
(c)	$V = B_1 + B_2 + B_3 + B_4 + B_5 + B_6 + B_7 + B_8 + B_9 + B_{10} + S$ $Y = R_1 + R_2 + R_3 + R_4 + R_5 + R_6 + R_7 + R_8 + R_9 + R_{10} + R_{11} + S$	May be implied by either correct distribution	M1
	$V \sim \mathrm{N}(553, 16.94)$ and $Y \sim \mathrm{N}(564, 15.88)$	Both correct	A1
	$D = Y - V$ so $D \sim \mathrm{N}(11\times 51 - 10\times 55,\ 11\times 1.2^2 + 10\times 1.3^2 + 2\times 0.2^2)$ or $D \sim \mathrm{N}(11,\ 32.82)$	Attempt at their difference for the mean, and their sum for the variance.	M1
		11 and awrt 32.8	A1
	$\mathrm{P}(D > 0) = \mathrm{P}\left(Z > \dfrac{0-11}{\sqrt{32.82}}\right)$	dep on 1st M1 for standardising using their mean and the standard deviation leading to a probability > 0.5	dM1
	$= \mathrm{P}(Z > -1.920...)$		
	$= 0.9726$	**awrt 0.973**	A1
			(6)
			Total 17

Pearson Education Limited. Registered company number 872828
with its registered office at Edinburgh Gate, Harlow, Essex CM20 2JE