Mark Scheme (Results)

October 2021

Pearson Edexcel International A Level
In Statistics S3 (WST03) Paper 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

October 2021
Question Paper Log Number P71287A
Publications Code WST03_01_2110_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
- $\underline{}$ 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. Ignore wrong working or incorrect statements following a correct answer.

Special notes for marking Statistics exams (for AAs only)

- If a method leads to "probabilities" which are greater than 1 or less than 0 then M0 should be awarded unless the mark scheme specifies otherwise.
- Any correct method should gain credit. If you cannot see how to apply the mark scheme but believe the method to be correct then please send to review.
- For method marks, we generally allow or condone a slip or transcription error if these are seen in an expression. We do not, however, condone or allow these errors in accuracy marks.
- If a candidate is "hedging their bets" e.g. give Attempt 1…Attempt 2…etc then please send to review.

Question Number	Scheme		Marks
1.	H_0: $\mu = 30$ H_1: $\mu < 30$		B1
	$z = \dfrac{29.5 - 30}{\dfrac{2.5}{\sqrt{80}}}$		M1
	$z = -1.7888\ldots$	awrt –1.79	A1
	$-1.7888 < -1.6449$		B1
	Reject H_0 or significant result or in the critical region		
	There is evidence to support the manager's claim.		A1
			(5)
	Notes		**Total 5**

	B1	Both hypotheses correct in terms of μ
	M1	for attempting test statistic, allow $\pm$, Condone $\sqrt{\dfrac{2.5}{80}}$
	A1	awrt –1.79 allow $\lvert z\rvert = 1.7888\ldots$ Allow p value of 0.0367 or awrt 0.0368 or CR $\leqslant 29.54$
	B1	$\lvert CV\rvert = 1.6449$ or better (Ignore any comparisons) Allow CR $\leqslant 29.54$ SC If p value of 0.0367 or awrt 0.0368 award B1 if 2nd A1 is awarded
	A1	For correct conclusion. Allow the manager's claim in words if it includes screws and less (oe)

Question Number	Scheme		Marks
2	H_0: Potassium has no effect on the quality of apple H_1: Potassium has an effect on the quality of apple		B1
	<table><tr><td>Grade</td><td>*A*</td><td>*B*</td><td>*C*</td><td>*D*</td><td>*E*</td></tr><tr><td>Expected values</td><td>9.6</td><td>67.2</td><td>124.8</td><td>24.0</td><td>14.4</td></tr></table>		M1A1
	$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(9-"9.6")^2}{"9.6"} + ... + \frac{(3-"14.4")^2}{"14.4"}$ or $\chi^2 = \sum \frac{O^2}{E} - N = \frac{9^2}{"9.6"} + ... + \frac{3^2}{"14.4"} - 240$		M1
	= 10.657…	awrt 10.7	A1
	Degrees of freedom = 4 $\chi^2_{4,0.05} = 9.488$		B1 B1ft
	[Reject H_0] Data suggests that potassium may affect the distribution of the grades of apples **or** there is evidence that Andy's belief is incorrect		A1ft
			(8)
	Notes		**Total 8**

B1	Both hypotheses in context. May use other wording eg The grading of apples remains the same.
M1	A correct method to calculate expected values eg 0.04×240
A1	At least 3 expected values correct
M1	A correct method using their expected values to calculate χ^2 At least one correct, ft their expected values with an intention to add
A1	awrt 10.7
B1	Degrees of freedom = 4 (may by be implied by 9.488)
B1ft	9.488 ft their DoF. If no DoF stated then this must be correct for their working.
A1ft	ft their χ^2 value provided the 2nd M1 is awarded and CV. If no hypotheses or hypotheses wrong way round do not award. Must include the word 'apples' or ' belief' oe

Question Number	Scheme	Marks
3(a)	<table><tr><td>jam</td><td>*A*</td><td>*B*</td><td>*C*</td><td>*D*</td><td>*E*</td><td>*F*</td><td>*G*</td><td>*H*</td><td>*I*</td></tr><tr><td>Price</td><td>1</td><td>2</td><td>4</td><td>5</td><td>3</td><td>6</td><td>7</td><td>8</td><td>9</td></tr><tr><td>Taste</td><td>1</td><td>2</td><td>8</td><td>9</td><td>4</td><td>3</td><td>6</td><td>5</td><td>7</td></tr></table>	M1
	$\sum d^2 = [0+0+]16+16+1+9+1+9+4[=56]$	M1A1
	$r_s = 1-\frac{6(56)}{9(80)}; = \frac{8}{15} = 0.5333..$ awrt 0.533	dM1A1
		(5)
(b)	$H_0 : \rho = 0$, $H_1 : \rho \neq 0$	B1
	Critical Value = 0.7	B1
	There is no evidence of a relationship between price and taste of strawberry jam	B1ft
		(3)
(c)	$r = \frac{16.4943}{\sqrt{2.0455 \times 243.5556}}$	M1
	= 0.7389... awrt 0.739	A1
		(2)
(d)	$H_0 : \rho = 0$, $H_1 : \rho > 0$	B1
	CV = 0.5822	B1
	There is evidence of a positive correlation between price and taste of strawberry jam	B1ft
		(3)
(e)	Spearman's rank as it is unlikely that a joint normal distribution applies. or the marks are a judgement or the marks are not a meaningful scale.	B1
		(1)
	Notes	**Total 14**

(a)	**M1**	Attempt to rank each jar for taste and price. At least 4 pairs of ranks correct
	M1	For an attempt at d^2 row for their ranks (may be implied by $\sum d^2 = 56$)
	A1	$\sum d^2 = 56$
	dM1	Dependent on the previous M being awarded. Using $1-\frac{6\sum d^2}{9(80)}$
	A1	$\frac{8}{15}$ or awrt 0.533
(b)	**B1**	Both hypotheses stated in terms of ρ. Must be two- tail.
	B1	0.7 for CV. Allow 0.6 if a one tail test is used
	B1ft	For a correct contextualised comment which has price and taste Follow through their r_s with their 0.7 (provided |their r_s| < 1)
(c)	**M1**	Correct method used
	A1	awrt 0.739
(d)	**B1**	Both hypotheses stated in terms of ρ. Must be one-tail. If B0 awarded in part (b) then allow any letter instead of ρ that is consistent with part (b)
	B1	0.5822 Allow 0.6664 if a two-tail test is used.
	B1ft	Correct conclusion in context which has positive correlation (this may be implied by a correct description of positive correlation), price and taste. Follow through their 0.5822 and 0.739
(e)	**B1**	Selecting Spearman's with a suitable reason. Do not allow 'because it is ranked' as a suitable reason

Question Number	Scheme	Marks
4(a)	Label the houses in area A 1- 41, area B 1 – 164, area C 1 – 123 and area D 1 - 82	M1
	Use random numbers to select a ...	M1
	Simple random sample of 20 area A, 80 area B, 60 area C and 40 area D	A1
		(3)
(b)	$\frac{357\times260}{595}$ or $\frac{238\times260}{595}$	M1
	156 and 104	A1
		(2)
(c)		

Observed	Expected	$\frac{(O-E)^2}{E}$
162	"156"	$\frac{(162-"156")^2}{"156"}=\frac{3}{13}=0.2307\ldots$
98	"104"	$\frac{(98-"104")^2}{"104"}=\frac{9}{26}=0.3461\ldots$

		M1
$\chi^2 = 4.657 + "0.2307..." + "0.346..."$		M1
$= 5.234\ldots.$	awrt 5.23	A1
$\nu = (2-1)(3-1) = 2$		B1
$\chi^2_2(0.05) = 5.991 \Rightarrow$ CR: $\chi^2 > 5.991$		B1ft
There is no evidence to suggest that there is an association between age and listening to *LSB*		dA1
		(6)

Notes **Total 11**

(a)

M1 For suitable labelling of all four areas. E.g. for area A: 1 – 41 or 0 - 40

M1 For use of random numbers to select houses in each area.

A1 For 20 A, 80B, 60C and 40 D (dependent on 2nd M1 only)

NB A simple random sample of 20 A, 80B, 60C and 40 D scores M0M1A1.

Allow M1 : allocate random numbers to each house

M1 : arrange the numbers in order

A1 : select the 1st 20 for area A, 80 for area B, 60 for area C and 40 for area D

SC If M0M0 scored then award B1 for 20 area A, 80 area B, 60 area C and 40 area D

(b)

M1 A correct method for finding one expected value.

A1 Correct answer for both values

(c)

M1 A correct method for finding both contributions to the χ^2 value

M1 Adding the two values to 4.657 (may be implied by a full χ^2 calculation, do not ISW)

A1 awrt 5.23

B1 2

B1ft 5.991 or better ft their DoF

dA1 A correct contextual conclusion, which has the words age and listening dependent on both M marks being awarded.

NB if they give a *p* value of 0.0730… rather than the CV they can get M1M1B1B0A1

Question Number	Scheme	Marks
5(a)	$2.977 \pm 2.5758 \times \frac{0.015}{3}$	M1,B1
	= (2.9641…, 2.9898…) awrt (2.964, 2.990)	A1
		(3)
(b)	The CI does not contain the stated weight.	B1
		(1)
(c)	$2.995 - 1.96 \times \frac{0.015}{\sqrt{n}} < 2.991$	M1
	$\sqrt{n} < \frac{1.96 \times 0.015}{2.995 - 2.991}$	M1d
	$\sqrt{n} <$ awrt 7.35	A1
	$n = 54$	A1cao
		(4)
	Notes	**Total 8**

(a)	**M1**	$2.977 \pm (z \text{ value}) \times \frac{0.015}{3}$
	B1	awrt 2.5758
	A1	awrt (2.964, 2.990 (condone 2.99))
(b)	**B1**	cao this must be consistent with their confidence interval
(c)	**M1**	Setting up an inequality using z value > 1.5 Condone =
	M1d	Dep on previous M mark. Correct rearranging to get $\sqrt{n} <$ … or $n <$ … Condone = or >
	A1	awrt 7.35 may be implied by awrt 54
	A1cao	54

Question Number	Scheme	Marks
6(a)	$\overline{h} = 65.4$	B1
	$s^2 = \dfrac{214676 - 50\times("65.4")^2}{49}$	M1
	$= 16.693...$ awrt 16.7	A1
		(3)
(b)	$H_0: \mu_{do} = \mu_{do\ not}$ $H_1: \mu_{do} < \mu_{do\ not}$	B1
	$z = \pm\dfrac{"65.4" - 70.8}{\sqrt{\dfrac{"16.693..."}{50} + \dfrac{29.6}{40}}}$	M1M1
	$= \pm 5.21...$ awrt 5.21	A1
	CV 1.6449	B1
	Amala's belief is supported	A1 ft
		(6)
(c)	CLT enables you to assume that (the sampling distribution of the sample mean of) resting heart rate is normally distributed for both groups	B1
		(1)
(d)	Each population/sample is independent **or** each male is independent of the other males.	B1
	Assume the $\sigma_{do}^2 = s_{do}^2$ and $\sigma_{do\ not}^2 = s_{do\ not}^2$	B1
		(2)
	Notes	**Total 12**

(a)	**B1**	65.4 only
	M1	Correct method to find s^2 using their $\overline{h}$
	A1	awrt 16.7
(b)	**B1**	Both hypotheses correct - must be clear which is exercise and which is not
	M1	For the denominator. Ft their 16.693…
	M1	Correct ft their 65.4 and 16.693…
	A1	awrt 5.21 allow $\|z\| = 5.21$…
	B1	$\|CV\| = 1.6449$ or better
	A1	ft their z value and CV if the hypotheses are the correct way round. Correct conclusion in context need belief. May be in words with heart and exercise e.g. resting heart rate is lower in men who exercise regularly
(c)	**B1**	For the idea both groups normally distributed
(d)	**B1**	For identifying the need for the groups **or** males to be independent.
	B1	Realising the $\sigma^2 = s^2$ Allow sample sizes big enough for CLT to hold

Question Number	Scheme	Marks
7(a)	$\mathrm{E}(B_1 - B_2) = 0$	B1
	$\mathrm{Var}(B_1 - B_2) = 0.006$	B1
	$\mathrm{P}(\lvert B_1 - B_2 \rvert > 0.1) = 2\mathrm{P}(B_1 - B_2 > 0.1)$	M1
	$= 2\times\mathrm{P}\left(Z > \dfrac{0.1}{\sqrt{"0.006"}}\right)\left[= 2\times\mathrm{P}(Z > 1.2909...)\right]$	M1
	$= 0.1967\ldots$ awrt 0.197	A1
		(5)
(b)	$\overline{B} \sim \mathrm{N}\left(1.96, \dfrac{0.003}{n}\right)$	B1
	$\mathrm{P}(\overline{B} > 2) = \mathrm{P}\left(Z > \dfrac{2-1.96}{\sqrt{0.003/n}}\right)$ [< 0.01]	M1
	$\dfrac{2-1.96}{\sqrt{\dfrac{0.003}{n}}} > 2.3263$	B1 dM1
	$n = 11$	A1
		(5)
(c)	$\mu_M = 21.8 + 500\times 1.96[= 1001.8]$; $\sigma_M{}^2 = 0.6 + 500\times 0.003[= 2.1]$	M1 ; M1
	Let $X = 4T - 3M$	M1
	$\mu_X = 4\times 774 - 3\times "1001.8"[= 90.6]$; $\sigma_X{}^2 = 16\times 1.8 + 9\times "2.1"[= 47.7]$	M1 ; M1
	$\mathrm{P}(4T - 3M > 100) = \mathrm{P}\left(Z > \dfrac{100 - "90.6"}{\sqrt{"47.7"}}\right)\left[= \mathrm{P}(Z > 1.361...)\right]$	M1
	= 0.0869 (table) or 0.08675… (calc)	A1
		(7)
		Total 17

(a)	**B1**	For expected value being 0 written or used
	B1	For 0.006 being written or used for Variance
	M1	Realising they need to consider both
	M1	Correct standardisation using their 0.1 and 0.006 If the expected value and/or standard deviation not stated then they must be correct
	A1	awrt 0.197
(b)	**B1**	The correct distribution written or used
	M1	Correct standardisation. Allow using their distribution if stated but must contain $\sqrt{n}$ for sd
	B1	Using awrt 2.3263
	dM1	Dep on previous M being awarded using a z value, $2 < z < 3$
	A1	11
(c)	**M1**	Correct method for finding the mean of M
	M1	Correct method for finding the var of M
	M1	Realising the need to find $4T - 3M$ or $4T - 3M - 100$ or $100 + 3M - 4T$
	M1	Correct method for finding the mean of X (using $4T - 3M - 100 = -9.4$ or $100 + 3M - 4T = 9.4$)
	M1	Correct method for finding the var of X
	M1	Correct standardisation using their mean of X and their standard deviation of X If these are not stated then they must be correct
	A1	awrt 0.0869 or 0.0868

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom