Mark Scheme (Results)

Summer 2019

Pearson Edexcel International Advanced level In Statistics S3 (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 2019
Publications Code WST03_01_1906_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
- * The answer is printed on the paper
- ⌊ 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.

Question Number	Scheme	Marks
1(a)	Number all of the students in **each** year group.	B1
	Use random numbers to select/take a (simple) random sample of ...	B1
	24 students in Group 1 (year 7-9) 20 students in Group 2 (year 10-11) 6 students in Group 3 (year 12-13)	B1
		(3)
(b)	'It is **more** representative when there might be systematic differences between age groups.'	B1
		(1)
		Total 4

	Notes	
1(a)	**B1**	For numbering/labelling/ordering (o.e.) students in **each** group (Condone poor numbering but if just one numbered list, the Groups must be distinguishable)
	B1	For use of random sample/numbers/selection
	B1	All 3 numbers correct with associated groups
(b)	**B1**	Any suitable advantage of using stratified sampling versus simple random sampling e.g. 'it gives more accurate estimates for each strata' or 'each year group is fairly represented' or 'reflects the population structure' more accurate on its own is B0

Question Number	Scheme	Marks
2(a)	H_0: There is no association between gender and favourite activity (independent) H_1 : There is an association between gender and favourite activity (dependent)	B1
		M1
	Expd | **Water sports** | **Bushcraft** | **Mountain activities** | **Total** **Boys** | 147.6 | 83.17… | 97.228… | (328) **Girls** | 104.4 | 58.828… | 68.771… | (232) **Total** | (252) | (142) | (166) | (560)	A1
	Observed | Expected | $\frac{(O-E)^2}{E}$ | $\frac{O^2}{E}$ 142 | 147.6 | 0.212… | 136.612… 110 | 104.4 | 0.3003... | 115.9003… 96 | 83.17 | 1.978… | 110.807… 46 | 58.83 | 2.797… | 35.968… 90 | 97.23 | 0.537… | 83.308… 76 | 68.77 | 0.759… | 83.988… Totals | 6.5863… | 566.5862613	dM1
	$\chi^2 = \sum \frac{(O-E)^2}{E}$ or $\sum \frac{O^2}{E} - 560$;	dM1
	$= $ awrt 6.6	A1
	$\nu = (2-1)(3-1) = 2$	B1
	$\chi^2_2(0.05) = 5.991 \Rightarrow$ CR: $\chi^2 \geq 5.991$	B1
	[in the CR/significant/Reject H_0]	
	conclude that there is evidence of an **association**(o.e.) between **gender** and favourite **activity**	A1
		(9)
(b)	Bushcraft as these contribute the most to the χ^2 value	B1
		(1)
		Total 10

Notes

(a)	**B1**	For both hypotheses. Must mention "gender" ***and*** "activity" o.e. at least once. Use of "relationship" or "correlation" or "connection" is B0.
	M1	Some attempt at $\frac{(\text{Row Total})(\text{Column Total})}{(\text{Grand Total})}$, can be implied by at least one correct E_i to 1dp
	A1	All expected frequencies are correct to awrt 1dp.
	dM1	At least 2 correct terms for $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$ or correct expressions with their E_i .
	dM1	Dep on previous M being awarded. For applying either $\sum \frac{(O-E)^2}{E}$ or $\sum \frac{O^2}{E} - 560$
	A1	awrt **<u>6.6</u>**
	B1	$\nu = 2$ (This mark can be implied by a correct critical value of 5.991)
	B1	5.991 condone 5.99 or p value 0.037…
	A1	Dependent upon 3rd M1 and 3rd B1, correct contextualised conclusion rejecting H_0 Condone "relationship" or "connection" here but **not** "correlation".
(b)	**B1**	Bushcraft and reason – allow comparisons of **proportions**

Question Number	Scheme	Marks
3(a)	<table><tr><td>singers</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>Rank Jamil</td><td>3</td><td>1</td><td>2</td><td>9</td><td>7</td><td>5</td><td>4</td><td>8</td><td>6</td></tr><tr><td>Rank age</td><td>1</td><td>4</td><td>5</td><td>7</td><td>6</td><td>3</td><td>2</td><td>9</td><td>8</td></tr></table>	M1
	$\sum d^2 = 4+9+9+4+1+4+4+1+4[=40]$	M1A1
	$r_s = 1-\frac{6(40)}{9(80)}; = 0.6666..$ awrt 0.667	dM1A1
		(5)
(b)	$H_0: \rho = 0$, $H_1: \rho > 0$	B1
	Critical Value $r_s = 0.6$	B1
	Either reject H_0 / accept H_1 / Result is significant / $r_s = 0.666...$ does lie in the CR	M1
	conclude that there is evidence that Jamil can rank singers in order of age by listening to them sing	A1ft
		(4)
(c)	Give each singer a rank of 3.5	B1
	Use the pmcc	B1
		(2)
	Notes	**Total 11**

(a)

M1 Attempt to rank for Jamil's estimate and actual ages. (At least 5 correct in either row) (Allow reverse rankings)

M1 For finding difference between each of their **ranks** and evaluating $\sum d^2$ (may be implied by A1)

A1 $\sum d^2 = 40$

dM1 Dependent on the 2 previous M being awarded. Using $1-\frac{6'\sum d^2{}'}{9(80)}$

A1 Awrt 0.667

May see:

singers	*B*	*C*	*A*	*G*	*F*	*I*	*E*	*H*	*D*
Rank Jamil	1	2	3	4	5	6	7	8	9
Rank age	4	5	1	2	3	8	6	9	7

(b)

B1 Both hypotheses stated in terms of r or ρ_s.

B1 0.6 for CV

M1 For a correct non-contradictory statement relating their r_S $(|r_S| < 1)$ with their c.v.
where $|\text{their c.v.}| < 0.6$ allow 'Do not reject H_0', 'not significant', 'not in critical region' for the method mark

A1ft Dependent on all previous marks in (b) scored. Ft their answer to part(a)
For a correct contextualised comment which has rank, singer/voice and age.
Follow through their r_s with 0.6 (provided |their r_s| < 1)

Note: **Two-tailed test**
Applying a two-tailed test scores a maximum of B0B1M1A0
So Award SC B0B1 for $H_0: \rho = 0$, $H_1: \rho \neq 0$ followed by critical value $r_s = (\pm)\, 0.7$
and allow access to the M1 mark only.

(c)

B1 For use of rank 3.5

B1 Use pmcc (independent of 1st B1)

Question Number	Scheme		Marks
4(a)	$\frac{10.84}{50} \pm 2.3263 \times \frac{0.008}{\sqrt{50}}$ **or** $\frac{10840}{50} \pm 2.3263 \times \frac{8}{\sqrt{50}}$	any z value	M1 A1
		2.3263	B1
	= (0.2141…,0.2194…,) = awrt **(0.214, 0.219) or (214, 219)**		A1
			(4)
(b)	Since the **weights** (of individual packets) are normally distributed (the sample means will also be normally distributed).		B1
			(1)
(c)	Bindy's belief is not supported as 0.22 kg is outside of the CI		B1ft
			(1)
			Total 6

Notes

4(a)

M1 $\frac{10.84}{50} \pm z \times \frac{0.008}{\sqrt{50}}$ or $\frac{10840}{50} \pm z \times \frac{8}{\sqrt{50}}$ where $|z| > 2$ and **condone mixed units**

A1 allow $\frac{10.84}{50} \pm z \times \frac{0.008}{\sqrt{50}}$ or $\frac{10840}{50} \pm z \times \frac{8}{\sqrt{50}}$ **units must be consistent**

B1 2.3263 (condone awrt 2.33)

A1 Both values correct to 3sf

(b) **B1** Idea that we are told the underlying variable is normally distributed

(c) **B1ft** Ft their CI

Do not allow ft if their interval includes 0

Question Number	Scheme		Marks
5			
	x | 1 | 3 | 6 $P(X=x)$ | $\frac{1}{2}$ | $\frac{1}{3}$ | $\frac{1}{6}$		M1
	$E(X) = 2.5$		A1
	$Var(X) = 9.5 - 2.5^2 = 3.25$		M1
	$\overline{X}$ = mean of 60 rolls		
	$P\left(\overline{X} > 2.75\right) \approx P\left(Z > \frac{2.75-2.5}{\sqrt{3.25/60}}\right)$ [by CLT]		M1A1
	$= 1 - P(Z < 1.074..)$		
	$= 1 - 0.8586$ [= 0.1414] [0.1423 from tables]		M1
	$10 \times ("0.1414")^2 (1 - "0.1414")^3$		M1
	= 0.1265….	awrt 0.127/0.128	A1
			(8)
			Total 8

Notes

M1 Setting up the distribution – may be implied by correct $E(X)$ or $Var(X)$

A1 For $E(X)$

M1 If distribution given allow 9.5 – "(their $E(X)$)^2 or "their $E(X^2)$– "(their $E(X)$)^2. If no working shown it must be correct.
Sight of N(2.5, $\frac{13}{240}$) implies M1A1M1

M1 Standardising using their $E(X)$ and $Var(X)$ $\frac{2.75-"2.5"}{\sqrt{"3.25"/60}}\left[= \frac{0.25}{\text{awrt } 0.233}\right]$

A1 $\frac{2.75-2.5}{\sqrt{3.25/60}}$ or awrt 1.07

M1 Finding the correct area $p < 0.5$ (may be implied by a correct answer)

M1 $10 \times ("p")^2 (1 - "p")^3$ with their p

A1 awrt 0.127 / 0.128

Question Number	Scheme		Marks
6(a)	$H_0 : \mu = 280$ $\quad H_1 : \mu > 280$		B1
	$z = \dfrac{290-280}{70/\sqrt{125}}$ or $\dfrac{c-280}{70/\sqrt{125}} = 1.6449$		M1
	$=1.597\ldots$ $\quad c = 290.299$		
	CV = 1.6449		B1
	Not significant / accept H_0 / not in the critical region		M1
	There is insufficient evidence to support Baako's claim		A1cso
			(5)
(b)	$H_0 : \mu_f = \mu_w + 100$ $\quad H_1 : \mu_f > \mu_w + 100$		B1
	$z = \dfrac{410-290-100}{\sqrt{\dfrac{90^2}{300}+\dfrac{60^2}{200}}}$		M1 M1
	$= 2.9814\ldots$		A1
	CV = 2.5758		B1
	Significant / reject H_0 / accept H_1 / in the critical region		M1
	There is sufficient evidence to support Ayodele's claim		A1cso
			(7)
(c)	(i) The plants to receive fertiliser are allocated **randomly**		B1
	(ii) [n is sufficiently large that the CLT holds so that the] standard deviation of sample $\approx$ standard deviation of population		B1
			(2)
			Total 14

(a)	**B1**	Both hypotheses correct in terms of μ
	M1	Standardising with 290, 280 and $\dfrac{70}{\sqrt{125}}$ (or awrt 6.26)
	B1	CV 1.64(49) or better (or p = awrt 0.055)
	M1	For a correct non-contradictory non-contextual statement (may be implied)
	A1cso	Fully correct solution with a correct comment in context dep upon all previous marks in (a)
(b)	**B1**	Both hypotheses correct – must be clear which is μ_f and μ_w NB: If use t-test send to review
	M1	for $\pm\dfrac{410-290}{\sqrt{\dfrac{90^2}{300}+\dfrac{60^2}{200}}}$ ie ignore incorrect or missing 100
	M1	for $\pm\dfrac{410-290\pm100}{\sqrt{\dfrac{90^2}{300}+\dfrac{60^2}{200}}}$
	A1	Awrt $\pm$ 2.98
	B1	CV ± 2.5758 (condone awrt 2.58) and compatible sign with their test statistic (or p = awrt 0.0014)
	M1	Correct non-conflicting comment following from their test statistic and their critical value
	A1cso	Fully correct solution with a correct comment in context dep upon all previous marks in (b)
(c)	**B1**	Idea that plants allocated at random to be given fertiliser
	B1	Realising that they need the sd of the population

Question Number	Scheme		Marks
7(a)			
	$\frac{e^{-2.8}(2.8)^5}{5!}n = n - 87.63 \quad [n = 96]$		M1
	Observed 5 and Expected awrt 8.37…		A1
			(2)
(b)	H_0: Poisson(2.8) is a suitable model/ good fit H_1: Poisson(2.8) is not a suitable model/ good fit		B1
	$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{("5"-"8.37")^2}{8.37} + \frac{(3-6.25)^2}{6.25} + 9.86$		M1 M1
	= awrt **12.9**		A1
	Degrees of freedom = 6 $\chi^2_{6,0.05} = 12.592$		M1 A1ft
	[Reject H_0] Data is not consistent with random sampling from a Poisson (2.8) model.		A1cso
			(7)

Notes

(a)	**M1**	Setting up an equation leading to a value for *n*. eg $5.84 = e^{-2.8} \times n$ May be implied by correct observed and expected (2dp) .
	A1	5 and awrt 8.37
(b)	**B1**	Both hypotheses correct must mention Po(2.8) at least once
	M1	For combining last 2 cells to get observed 3 and expected 6.25 or $\frac{(3-6.25)^2}{6.25}$ seen.
	M1	For $\frac{("5"-"8.37")^2}{8.37} + 9.86 + ...$ or a fully correct calculation
	A1	awrt 12.9 (this is dependent upon M1M1)
	M1	Using df for their (number of expected values used – 1)
	A1ft	12.592 (ft 14.067 from df = 7)
	A1cso	Jeff's model/Poisson is not suitable (condone missing 2.8 here) all marks must have been scored.

Question Number	Scheme		Marks
8(a)	Let $X = L - M$	Let $X = L - M - 90$	
	$\mathrm{E}(X) = 100$	$\mathrm{E}(X) = 10$	M1
	$\mathrm{Var}(X) = 7^2 + 6^2$	$\mathrm{Var}(X) = 7^2 + 6^2$	M1
	$X \sim N(100, 85)$	$X \sim N(10, 85)$	A1
	$\mathrm{P}(X > 90) = \mathrm{P}\left(Z > \frac{90-100}{\sqrt{85}}\right)$	$\mathrm{P}(X > 0) = \mathrm{P}\left(Z > \frac{0-10}{\sqrt{85}}\right)$	dM1
	$= 0.8599\ldots$(calc 0.860…) awrt 0.86		A1
			(5)
(b)	(Let $Y = L - 4S$)		
	$\mathrm{E}(Y) = 20$		B1
	$\mathrm{Var}(Y) = 7^2 + 16 \times \sigma^2$		M1
	$Y \sim \mathrm{N}(20, 49 + 16\sigma^2)$		A1
	$\mathrm{P}(Y > 0) = \mathrm{P}\left(Z > \frac{0 - '20'}{\sqrt{'49 + 16\sigma^2'}}\right)$		M1
	$\frac{0 - '20'}{\sqrt{'49 + 16\sigma^2'}} = -1.96$		M1B1
	$20^2 = 1.96^2\left(49 + 16\sigma^2\right)$		dM1
	$\sigma = 1.8561...$		A1cso
			(8)
			Total 13

		Notes
8(a)	**M1**	Either 100 or 10
	M1	For addition of variances
	A1	N(100, 85) or N(10,85)
	dM1	Dependent on the first 2 M marks $\frac{90-100}{\sqrt{"85"}}$ or $\frac{0-10}{\sqrt{"85"}}$
(b)	**B1**	$\mathrm{E}(Y) = 20$
	M1	Use of $16 \times \mathrm{Var} + \ldots$
	A1ft	$\mathrm{N}('20', 49 + 16\sigma^2)$ (may be implied by its use)
	M1	M1 standardising using their mean, sd and zero
	M1	Putting standardisation equal to a z value where $\lvert z \rvert > 1.5$
	B1	–1.96 or better used in a standardisation equation with compatible signs
	dM1	Squaring (dependent upon 2nd and 3rd M1 marks)
	A1cso	awrt 1.86 (must come from correct working)

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