Mark Scheme (Results)

Summer 2018

Pearson Edexcel International A Level
In Mathematics
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 2018
Publications Code WST03_01_1806_MS

June 2018
WST03/01 Statistics 3
Mark Scheme

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.
- 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 $\sqrt{}$ 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.

June 2018 IAL - WST03/01 Statistics 3

Question Number	Scheme	Marks
1. (a)	Footballer: A, B, C, D, E, F, G, H, I Rank x: 1, 2, 3, 4, 5, 6, 7, 8, 9 Rank y: 6, 9, 8, 2, 5, 4, 7, 3, 1 Rank x: 9, 8, 7, 6, 5, 4, 3, 2, 1 Rank y: 4, 1, 2, 8, 5, 6, 3, 7, 9	M1
	$\sum d^2 = 25+49+25+4+0+4+0+25+64 = 196$	M1 A1
	$r_s = 1 - \dfrac{6(196)}{9(9^2-1)}; = -0.63333333...$ or $-\dfrac{19}{30}$	dM1; A1
		[5]
(b)	$H_0: \rho_s = 0, \quad H_1: \rho_s < 0$	B1
	Critical Value = -0.6000 or -0.6 or CR: $r_s \leqslant -0.6000$	B1
	Since $r_s = -0.6333...$ lies in the CR (or $-0.6333 < -0.6$), reject H_0	M1
	Either conclude that • <u>Russell's claim</u> is <u>true</u> • Footballers with <u>lower BMI are slower</u> Conclusion in context	A1
		[4]
(c)	**Both** Critical Value $r = -0.5822$ /CR: $r \leqslant -0.5822$ **and** <u>does not lie in the CR</u> /Result is <u>not significant</u>/<u>Do not reject H_0</u> (or accept H_0)	M1
	Conclude that there is <u>no negative correlation</u> oe Context not required here.	A1
		[2]
(d)	The relationship (between BMI and time taken to complete the obstacle course) is non-linear oe	B1
		[1] **12**

Question 1 Notes

1. (a)	**1st M1**	Attempt to rank data for x and y at least 5 correct for each (allow reverse rankings)
	2nd M1	For finding the difference between each of the ranks and evaluating $\sum d^2$
	1st A1	$\sum d^2 = 196$ or from reverse rankings $\sum d^2 = 9+1+1+16+0+0+16+1+0 = 44$
	3rd dM1	***is dependent on 1st M1*** for use of $1 - \dfrac{6(\text{"196"})}{9(9^2-1)}$ with their $\sum d^2$.
	2nd A1	awrt -0.633 or $-\dfrac{19}{30}$ or from reverse rankings $\dfrac{19}{30}$
(b)	**1st B1**	Both hypotheses stated in terms of r or ρ_s.
	Note	One tail H_1 must be compatible with their ranking.
	2nd B1	Critical value of ± 0.6
	M1	For a correct statement relating their r_S $(\lvert r_S \rvert < 1)$ with their c.v. where $\lvert \text{their c.v.} \rvert < 1$
	A1	For a contextualised comment which is rejecting H_0, which must mention either "<u>negative correlation</u>", "<u>BMI</u>" and "<u>time</u>" **or** "<u>lower BMI are slower</u>" o.e.
	Note	Follow through their r_S with their c.v. (provided $\lvert \text{their c.v.} \rvert < 1$)
(c)	**M1**	Allow ± 0.5822 Ignore hypotheses in this part for the M1 Use of -0.633 here is M0

Question Number	Scheme	Marks
2. (a)	$\hat{p} = \dfrac{7(3)+8(5)+9(18)+10(28)+11(17)+12(4)}{12(3+5+18+28+17+4) \text{ or } 12(75)}$ $\left\{= \dfrac{738}{900}\right\}$ $= 0.82\,(*)$ Answer is given. See notes.	M1 A1cso
		[2]
(b)	$r = 75 \times {}^{12}C_9(0.82)^9(0.18)^3\ \{= 16.1296941...\}$ (formula) $s = 75 - (2.80 + 7.97 + \text{their } r + 22.04 + 18.26 + 6.93)$	M1
	$r = 16.1296941...$; $s = 0.87...$ $r = \text{awrt } 16.13$; $s = \text{awrt } 0.87$	A1; A1
		[3]
(c)	H_0: Binomial distribution is a suitable (or good) model (or fit) H_1: Binomial distribution is not a suitable model	B1

#	O_i	E_i	Comb O_i	Comb E_i	$\dfrac{(O_i - E_i)^2}{E_i}$	$\dfrac{O_i^2}{E_i}$
$\leqslant 6$	0	0.87	8	11.64	1.1383…	5.4983…
7	3	2.80				
8	5	7.97				
9	18	16.13	18	16.13	0.2168…	20.0868…
10	28	22.04	28	22.04	1.6117…	35.5717…
11	17	18.26	17	18.26	0.0869…	15.8269…
12	4	6.93	4	6.93	1.2388...	2.3088...
				Totals	4.2925...	79.2925...

Marks for table: M1 M1

Question Number	Scheme	Marks
	$X^2 = \text{awrt } 4.3$	A1
	$v = 5 - 1 - 1 = 3$	B1 ft
	$\chi^2_3(0.10) = 6.251 \Rightarrow$ CR: $X^2 \geqslant 6.251$	B1 ft
	[does not lie in the CR/not significant/Do not reject H_0/Accept H_0]	
	Binomial distribution is a suitable model. A correct conclusion (context not required here) which is based on ***their*** X^2-value and ***their*** χ^2-critical value.	A1
		[7] **12**

	Question 2 Notes	
2. (a)	**M1**	At least 2 non zero products on the numerator and correct division for their method
	A1 cso	Correct answer $p = 0.82$ with no incorrect working seen
(b)	**M1**	For any correct method (or a correct expression) for finding either r or s.
	A1; A1	$r = \text{awrt } 16.13$; $s = \text{awrt } 0.87$
(c)	**1st B1**	Must have both hypotheses and mention Binomial at least once. Inclusion of 0.82 for p in hypotheses is B0 but condone in conclusion.
	1st M1	For an attempt to pool 8, 7 and $\leqslant 6$ germinating seeds ONLY.
	2nd M1	For an attempt at the test statistic, at least 2 correct expressions/values (to awrt 2 d.p. or truncated 2 d.p.)
	1st A1	awrt 4.3
	2nd B1ft	For their evaluated $n - 1 - 1$. i.e. realising that they must subtract 2 from their n.
	3rd B1ft	For a correct ft for their $\chi^2_k(0.10)$, from their degrees of freedom
	Note	For 0.10 significance: $\chi^2_6 = 10.645$ $\chi^2_5 = 9.236$ $\chi^2_4 = 7.779$ $\chi^2_2 = 4.605$
	Final A1	***Dependent on the 2nd Method mark only.*** A correct conclusion (context not required) which is accepting H_0
	Note	No follow through on their hypotheses if they are stated the wrong way round.
	Note	Contradictory statements score A0. E.g. "significant, do not reject H_0".
	Note	Condone mentioning of $B(12, 0.82)$ in conclusion.

Question Number	Scheme		Marks
3. (a)	$\left\{\hat{\mu}_x = \bar{x} = \dfrac{92.0}{20} \Rightarrow\right\}\ \bar{x} = 4.6$ (cm)	4.6	B1
	$\left\{\hat{\sigma}_x^{\,2} =\right\}\ s_x^{\,2} = \dfrac{433.4974 - 20(4.6)^2}{20-1} = 0.541968...\ (\text{cm})^2$	Applies $\dfrac{\sum x^2 - 20(\text{their } \bar{x})^2}{20-1}$	M1
		awrt **<u>0.542</u>**	A1
			[3]
(b)	Combined Sample: Mean $= \dfrac{92.0+142.5}{20+30} = 4.69$	4.69 Can be implied.	B1
	$s^2 = \dfrac{433.4974 + 689.5078 - 50(4.69)^2}{20+30-1}$; $= 0.4734734694$		M1;
		awrt 0.473 or 0.4735 (can be implied)	A1
	$\dfrac{s}{\sqrt{n}} = \dfrac{\sqrt{0.4734734694...}}{\sqrt{50}}$; $= 0.09731119868...$	For use of $s/\sqrt{50}$	M1;
		awrt **<u>0.0973</u>**	A1
			[5]
(c)	$H_0 : \mu = 4.5 \quad H_1 : \mu > 4.5$	Correct hypotheses	B1
	$z = \dfrac{\text{"4.69"} - 4.5}{\left(\frac{0.71}{\sqrt{50}}\right)}$; $= 1.892257583...$	$\pm \dfrac{\text{their } 4.69 - 4.5}{\left(\frac{0.71}{\sqrt{50}}\right)}$ or equivalent.	M1;
		awrt **<u>1.89</u>**	A1
	One tailed c.v. $Z = 1.6449$ or CR: $Z \geqslant 1.6449$ or p-value = awrt 0.029 or awrt $0.029 < 0.05$	Critical value of 1.6449 or a correct probability comparison.	B1
	[in the CR/significant/Reject H_0/ $0.029 < 0.05$]		
	Conclude either • there is evidence to <u>support</u> the <u>farmer's claim</u> • that the <u>mean width</u> of duck <u>eggs</u> is <u>greater than 4.5</u> cm.	A correct conclusion which is rejecting H_0 in context and is based on ***their*** *z*-value and ***their*** critical value, where $\lvert \text{c.v.} \rvert > 1$.	A1
			[5]
			13

Question 3 Notes

3. (a)	**M1**	Also allow M1 for applying $\dfrac{20}{(20-1)}\left(\dfrac{\sum x^2}{20} - (\text{their } \bar{x})^2\right)$
(b)	**1st M1**	Also allow 1st M1 for applying $\dfrac{50}{(50-1)}\left(\dfrac{\sum x^2 + \sum y^2}{20+30} - (\text{their } \bar{x}_{\text{comb}})^2\right)$
	Note	Award B1M1A1M1A1 for awrt 0.0973 which follows from no working.
(c)	**1st M1** **2nd A1**	Condone use of 4.6 for this M1 mark. Conclusion must refer to either "farmer's claim" oe or "mean width" and "eggs".

Question Number	Scheme		Marks
4. (a)	H_0 : Mean number of reported first-aid incidents per 1000 employees is the same at each warehouse. H_1 : Mean number of reported first-aid incidents per 1000 employees is not the same.		B1
	Warehouse / **Calculation** / **Expected**: *A*: $\frac{(2)(114)}{12}$ = 19 *B*: $\frac{(1)(114)}{12}$ = 9.5 *C*: $\frac{(3.8)(114)}{12}$ = 36.1 *D*: $\frac{(3)(114)}{12}$ = 28.5 *E*: $\frac{(2.2)(114)}{12}$ = 20.9	Some attempt at using the correct formula to find their 5 expected values (expected number of incidents). Can be implied by at least one correct E_i.	M1
		All expected frequencies are correct.	A1
	Observed / Expected / $\frac{(O-E)^2}{E}$ / $\frac{O^2}{E}$: 15 / 19 / 0.8421… / 11.8421… 10 / 9.5 / 0.0263... / 10.5263... 40 / 36.1 / 0.4213… / 44.3213… 26 / 28.5 / 0.2193... / 23.7193... 23 / 20.9 / 0.2110… / 25.3110… Totals / 1.7200… / 115.72…	Dependent upon previous M1 At least 3 correct terms for $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$ Accept 2 sf accuracy for the dM1 mark.	dM1
	$X^2 = \sum\frac{(O-E)^2}{E}$ or $\sum\frac{O^2}{E} - 114 =$ awrt 1.72	awrt **1.72**	A1
	$n = 5 - 1 = 4 \Rightarrow \chi^2_4(0.05) = 9.488 \Rightarrow$ CR: $X^2 \geqslant 9.488$	**9.488**	B1
	[not in the CR/not significant/Do not Reject H_0 /Accept H_0]		
	Conclude either: • manager's claim is supported • that the mean number of reported first-aid incidents per 1000 employees is the same at each warehouse.	A correct conclusion in context which is based on ***their*** X^2 value and ***their*** χ^2 -critical value.	A1 ft
			[7]
(b)	Select every 4th record from warehouse *C*.		B1
	{having chosen the first record by}		
	selecting a random number.		dB1
			[2]
			9
	Question 4 Notes		

(a)

SC 1

Expected values of 9.5 used

Observed	Expected	$\frac{(O-E)^2}{E}$
7.5	9.5	0.4210…
10	9.5	0.0263...
10.5…	9.5	0.1108…
8.6…	9.5	0.0730...
10.4…	9.5	0.0959…
	Totals	0.727…

Can score B1M1A0M1A0B1A1ft (5 out of 7)

SC 2

Expected values of 9.43… used

Observed	Expected	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$
7.5	9.43	0.3948…	5.965…
10	9.43	0.0345...	10.6050...
10.5…	9.43	0.1275…	11.7507…
8.6…	9.43	0.0617...	7.9655...
10.4…	9.43	0.1114…	11.5910…
	Totals	0.729…	47.877…

Can score B1M1A0M1A0B1A0 (4 out of 7)

(b) Use of 3800 in part (b) is B0B0

Question Number	Scheme		Marks
5.	95% CI for μ is $(30.612, 31.788)$; c% CI for μ is $(30.66, 31.74)$		
(a)	$\frac{2(1.96)\sigma}{\sqrt{25}} = 31.788 - 30.612 \{= 1.176\}$	$\frac{2"z"\sigma}{\sqrt{25}} = 31.788 - 30.612$	M1 oe
		1.96	B1
	$\left\{\Rightarrow \sigma = \frac{(1.176)(5)}{2(1.96)} \Rightarrow\right\} \sigma = 1.5$	$\sigma = 1.5$	A1
			[3]
(b)	$\frac{2z(1.5)}{\sqrt{25}} = 31.74 - 30.66 \{= 1.08\}$	$\frac{2z("1.5")}{\sqrt{25}} = 31.74 - 30.66$	M1 oe
	$z = \frac{(1.08)(5)}{2("1.5")} \to z = 1.8$	$z = 1.8$	A1ft
	$\left[\frac{c}{100} =\right] 2(0.9641) - 1$	$2\Phi(\text{their "1.8"}) - 1$ oe	M1
	$\Rightarrow c = 92.8$ (3sf)	awrt **92.8**	A1
			[4]
			7

		Question 5 Notes
5. (a)	**M1**	Also allow M1 (oe) for $31.2 + \frac{\text{"their } z\text{"}\sigma}{\sqrt{25}} = 31.778$, where $31.2 = \frac{30.612 + 31.778}{2}$
(b)	**1st M1**	Also allow M1 (oe) for $31.2 + \frac{z(\text{their "1.5"})}{\sqrt{25}} = 31.74$, where $31.2 = \frac{30.66 + 31.74}{2}$
	1st A1ft	For a correct (ft) expression using their value of σ
	2nd M1	awrt 0.928 implies this mark
	Note	Use of 1.6449 gives $\sigma = 1.787...$ and leads to $z = 1.51...$ and $c = 86.9$ (3sf) (M1A1ftM1A0)

Question Number	Scheme	Marks
6.	Y has a continuous uniform distribution $\left[a - 3,\ a+6\right]$	
(a)	$\mathrm{E}(Y) = \dfrac{a+6+a-3}{2}$ $\left\{= \dfrac{(2a+3)}{2}\ \text{or}\ a+\dfrac{3}{2}\right\}$	M1
	$\mathrm{Var}(Y) = \dfrac{(a+6-a+3)^2}{12}$ $\left\{= \dfrac{81}{12}\ \text{or}\ \dfrac{27}{4}\ \text{or}\ 6.75\right\}$ May be implied	M1
	$\bar{Y} \sim \mathrm{N}\left(a+\dfrac{3}{2}, \dfrac{9}{80}\right)$ $\mathrm{N}\left(a+\dfrac{3}{2}, \dfrac{9}{80}\right)$	A1
		[3]
(b)	$13.4 - 2.3263\sqrt{\dfrac{9}{80}} < \mathrm{m} < 13.4 + 2.3263\sqrt{\dfrac{9}{80}}$ $13.4 \pm "z"(\text{their } SE_{\bar{Y}})$	M1
	2.3263	B1
	$13.4 - 2.3263\sqrt{\dfrac{9}{80}} < a+\dfrac{3}{2} < 13.4 + 2.3263\sqrt{\dfrac{9}{80}}$	
	$13.4 - 2.3263\sqrt{\dfrac{9}{80}} + 4.5 < a+6 < 13.4 + 2.3263\sqrt{\dfrac{9}{80}} + 4.5$ $13.4 \pm "z"(\text{their } SE_{\bar{Y}}) + 4.5$	M1
	$17.11973576... < a+6 < 18.68026474...$ awrt $(17.1, 18.7)$	A1
		[4]
	<u>**Alternative Method for part (b)**</u>	
(b)	$13.4 - 2.3263\sqrt{\dfrac{9}{80}} < \mathrm{m} < 13.4 + 2.3263\sqrt{\dfrac{9}{80}}$ $13.4 \pm "z"(\text{their } SE_{\bar{Y}})$	M1
	2.3263	B1
	$11.11973526... < a < 12.68026474...$	
	$11.11973526... + 6 < a+6 < 12.68026474... + 6$ $13.4 \pm "z"(\text{their } SE_{\bar{Y}}) - 1.5 + 6$	M1
	$17.11973576... < a+6 < 18.68026474...$ awrt $(17.1, 18.7)$	A1
		[4]
		7
	Question 6 Notes	
(b)	1st M1 The inequalities may be seen separately. For only considering 1-tail of confidence interval (usually the upper tail) allow access to 1st M1 only (so M1B1M0A0 is possible). A second division of their SE by 60 is 1st M0	

Question Number	Scheme		Marks
7. (i)	$A \sim N(21, 2^2)$, $B \sim N(32, 7^2)$ and $C \sim N(45, 9^2)$ A, B, C are independent.		
(a)	$T = A + B + C$		
	$E(T) = 21 + 32 + 45$ or $Var(T) = 2^2 + 7^2 + 9^2$	A fully correct method of finding $E(T)$ or $Var(T)$	M1
	$E(T) = 98$ and $Var(T) = 134$	Both $E(T) = 98$ and $Var(T) = 134$	A1
	$\{So\ T \sim N(98, 134)\}$		
	$\{P(T > 90) = \}\ P\left(Z > \frac{90 - 98}{\sqrt{134}}\right)$	Standardising $(\pm)$ with their mean and their standard deviation	M1
	$= P(Z > -0.69109...)$		
	$= 0.7549$ (or $0.75525...$)	awrt **<u>0.755</u>**	A1
			[4]
(b)	$\{P(A > B) = P(A - B > 0)\}$		
	$E(A - B) = 21 - 32$ or $Var(T) = 2^2 + 7^2$	A fully correct method of finding $E(A - B)$ or $Var(A - B)$	M1
	$E(A - B) = -11$ and $Var(A - B) = 53$	Both $E(A - B) = -11$ and $Var(A - B) = 53$	A1
	$\{So\ A - B \sim N(-11, 53)\}$		
	$\{P(A - B > 0)\} \Rightarrow P\left(Z > \frac{0 - -11}{\sqrt{53}}\right)$	Standardising $(\pm)$ with their mean and their standard deviation	M1
	$= P(Z > 1.510966...)$		
	$= 0.06539855...$ (or 0.0655)	**<u>0.0655</u>** or awrt **<u>0.0654</u>**	A1
			[4]
(ii)	$\{P(X_1 > \bar{X} + k\sigma) = 0.1 \Rightarrow P(X_1 - \bar{X} > k\sigma) = 0.1\}$		
	$X_1 - \bar{X};\left\{= X_1 - \frac{(X_1 + X_2 + X_3 + X_4)}{4} = \frac{3X_1 - (X_2 + X_3 + X_4)}{4}\right\}$	For attempting to find the distribution of $X_1 - \bar{X}$	M1
	$E(X_1 - \bar{X}) = 0$	Correct mean	A1
	$Var(X_1 - \bar{X}) = \frac{9\sigma^2 + 3\sigma^2}{4^2};\ \Rightarrow\ X_1 - \bar{X} \sim N(0, 0.75\sigma^2)$	Correct expression for $Var(X_1 - \bar{X})$	dM1
		$X_1 - \bar{X} \sim N(0, 0.75\sigma^2)$	A1
	$\left\{P(X_1 - \bar{X} > k\sigma) = 0.1 \Rightarrow P\left(Z > \frac{k\sigma - 0}{\sqrt{0.75\sigma^2}}\right) = 0.1\right\}$		
	So, $\frac{k}{\sqrt{0.75}} = 1.2816$	Standardising using their $\sqrt{Var(X_1 - \bar{X})}$. Note that σ must cancel and equating to a z-value, $\lvert z \rvert > 1$.	M1
		1.2816	B1
	$\{k = \sqrt{0.75}\,(1.2816)\} \Rightarrow k = 1.109898157...$	awrt **<u>1.11</u>**	A1
			[7]
			15

	Question 7 Notes	
7. (i) (a)	**1st M1**	Can be implied by either a correct $E(T)$ or $Var(T)$
(i) (b)		Allow equivalent method using $B - A < 0$
(ii)	**Final A1**	Dependent upon all previous M marks in (ii)

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