Mark Scheme (Results)

Summer 2022

Pearson Edexcel International Advanced 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

Summer 2022
Question Paper Log Number P72469A
Publications Code WST03_01_2206_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.
- Crossed out work should be marked UNLESS the candidate has replaced it with an alternative response.

PEARSON 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
These are marks given for a correct method or an attempt at a correct method. In Mechanics they are usually awarded for the application of some mechanical principle to produce an equation.
e.g. resolving in a particular direction, taking moments about a point, applying a suvat equation, applying the conservation of momentum principle etc.
The following criteria are usually applied to the equation.

To earn the M mark, the equation
(i) should have the correct number of terms
(ii) be dimensionally correct i.e. all the terms need to be dimensionally correct
e.g. in a moments equation, every term must be a 'force x distance' term or 'mass x distance', if we allow them to cancel 'g' s.
For a resolution, all terms that need to be resolved (multiplied by sin or cos) must be resolved to earn the M mark.

M marks are sometimes dependent (DM) on previous M marks having been earned.
e.g. when two simultaneous equations have been set up by, for example, resolving in two directions and there is then an M mark for solving the equations to find a particular quantity – this M mark is often dependent on the two previous M marks having been earned.

'A' marks
These are dependent accuracy (or sometimes answer) marks and can only be awarded if the previous M mark has been earned. E.g. M0 A1 is impossible.

'B' marks
These are independent accuracy marks where there is no method (e.g. often given for a comment or for a graph)

A few of the A and B marks may be f.t. – follow through – marks.

3. General 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 √ 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.

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.

Question Number	Scheme	Marks
1 (a)	You would assign an average rank between the tied ranks	B1
		(1)
(b)	Rank for total tournaments 1 3 4 6 8 9 2 5 10 7	M1
	$\sum d^2 = 0+1+1+4+9+9+25+9+1+9$ [= 68]	M1
	$r_s = 1-\dfrac{6\times'68'}{10\left(10^2-1\right)}$	dM1
	= 0.5878… awrt 0.588	A1
		(4)
(c)	$H_0: \rho = 0$, $H_1: \rho > 0$	B1
	Critical Value = 0.5636 or CR … 0.5636	B1
	Reject H_0 or significant or lies in the critical region	dM1
	There is sufficient evidence of a positive correlation between rank and total tournaments won	A1
		(4)
(d)	2.5% and r_s = 0.6485 or CR … 0.6485	B1
		(1)

	Notes	Total 10
(a)	**B1**	for an appropriate explanation of how to deal with tied ranks. Ignore any comments regarding PMCC Do not allow add 0.5 to both ranks
(b)	**M1**	attempt to rank total tournaments (at least four correct) Condone reversed ranks
	M1	finding the difference between players rank and each of their total tournaments ranks and evaluating $\sum d^2$ May be implied by 68
	dM1	dependent on 1st M1. Using $1-\dfrac{6\sum d^2}{10(99)}$ with their $\sum d^2$ (you will need to check their $\sum d^2$ if no value shown)
	A1	awrt 0.588 Allow $\frac{97}{165}$
(c)	**B1**	both hypotheses correct. Must be in terms of ρ. Must be attached to H_0 and H_1 If r_s is negative in part (b) then allow $H_1: \rho < 0$
	B1	critical value of 0.5636 If r_s is negative in part (b) then allow –0.5636
	dM1	dependent on 2nd B1. A correct statement ft their part (b) and their CV– no context needed but do not allow contradicting non contextual comments. This may be implied by a correct contextual conclusion.
	A1	correct conclusion which is rejecting H_0, which must mention **rank** and **total tournaments**. No hypotheses is A0.
		NB If they have used $H_1: \rho < 0$ then the maximum they can score is B1B1dM1A0
(d)	**B1**	for 2.5% and a correct critical value of 0.6485

Question Number	Scheme	Marks
2 (a)	$\bar{x} = \left[\frac{7690}{100}\right] = 76.9$	B1
	$s_x^2 = \frac{669.24}{99} = 6.76$	M1 A1
		(3)
(b)	$H_0 : \mu_x = \mu_y \qquad H_1 : \mu_x \neq \mu_y$	B1
	$Z = \frac{"76.9" - 75.9}{\sqrt{\frac{"6.76"}{100} + \frac{2.2^2}{80}}} = 2.793...$ awrt ± 2.79	M1 M1 A1
	2 tailed critical value $z = \pm 2.5758$	B1
	Reject H_0/Significant/In the critical region	M1
	There is sufficient evidence to suggest that the mean water temperature after 4 hours for brand *A* is different to brand B	A1ft
		(7)
(c)	(It is reasonable) since both samples are (reasonably) large	B1
		(1)
	Notes	**Total 11**

(a)	**B1**	for 76.9
	M1	for use of $\frac{1}{n-1}\sum\left(x-\bar{x}\right)^2$ oe
	A1	for 6.76
(b)	**B1**	for both hypotheses correct. Must be attached to H_0 and H_1 Allow equivalent hypotheses. Must be in terms of μ Allow any letter for the subscripts
	M1	for a correct method to find the standard error. Follow through their values from (a)
	M1	an attempt at $\pm\frac{a-b}{\sqrt{\frac{c}{100}+\frac{d^2}{80}}}$ with at least 3 of *a*, *b*, *c* or *d* correct.
	A1	awrt ± 2.79
	B1	$z = \text{awrt} \pm 2.5758$ seen (Allow $z = \text{awrt} \pm 2.3263$ if a one tailed test is used)
	M1	a correct statement consistent with their CV and *Z* value – need not be contextual but do not allow contradicting non contextual comments. This may be implied by a correct contextual conclusion.
	A1ft	This mark is dependent on the 2nd M mark being awarded. A correct contextual statement ft their CV and their *Z* value
(c)	**B1**	a correct explanation, which makes reference to both samples. e.g. Do not allow the sample is large enough

Question Number	Scheme	Marks
3 (a)	$\left[\frac{26.624+28.976}{2}\right]=27.8$	B1
		(1)
(b)	$28.976-26.624=2\times1.96\times\frac{\sigma}{\sqrt{25}}$ or $26.624='27.8'-1.96\times\frac{\sigma}{\sqrt{25}}$ or $28.976='27.8'+1.96\times\frac{\sigma}{\sqrt{25}}$	M1 B1
	$\sigma=3$*	A1* cso
		(3)
(c)	$2\times z\times\frac{3}{\sqrt{25}}=2.1$ So $z=1.75$	M1 A1
	$\text{P}(Z>'1.75')=\text{P}(Z<-'1.75')=1-'0.9599'='0.0401'$	M1 A1ft
	Confidence level = $100\times(1-2\times'0.0401')=91.98\%$	M1 A1
		(6)
(d)	$2\times1.96\times\frac{3}{\sqrt{n}}<1.5$	M1
	$\sqrt{n}>\frac{6\times1.96}{1.5}$	dM1
	$\sqrt{n}>$ awrt 7.84 So $n=62$	A1 A1
		(4)
	Notes	**Total 14**

(a)	**B1**	for 27.8
(b)	**M1**	for $28.976-26.624=2\times z$ value $\times\frac{\sigma}{\sqrt{25}}$ or $26.624='27.8'-z$ value $\times\frac{\sigma}{\sqrt{25}}$ or $28.976='27.8'-z$ value $\times\frac{\sigma}{\sqrt{25}}$ where $1.5 < z < 2.4$
	B1	awrt 1.96
	A1* cso	answer is given so no incorrect working must be seen
(c)	**M1**	for $2\times z\times\frac{3}{\sqrt{25}}=2.1$
	A1	for $z=1.75$
	M1	for $1-p$, where p is a probability
	A1ft	for 0.0401 or ft their z value (Allow 0.04)
	M1	for $100\times(1-2\times0.0401)$ ft their $\text{P}(Z<-1.75)$
	A1	awrt 92.0 (allow 92)
(d)	**M1**	for $2\times z$ value $\times\frac{3}{\sqrt{n}}<1.5$ oe z value must either be correct or consistent with part (b) Allow $\leqslant$ or = Condone > or $\geqslant$
	dM1	Dependent on previous M mark. Correct rearrangement to get $\sqrt{n}>...$ or $n>\ldots$ oe Allow $\geqslant$ or = Condone < or $\leqslant$
	A1	awrt 7.84 may be implied by awrt 61.5
	A1	for $n=62$

Question Number	Scheme	Marks
4 (a)	[Continuous] uniform on the interval [0, 7]	B1
		(1)
(b)	mean = 3.5	B1
	standard deviation = $\sqrt{\frac{(7-0)^2}{12}}$	M1
	$= \frac{7}{\sqrt{12}} = 2.0207...$ awrt 2.02	A1
		(3)
(c)	By the CLT $\bar{T} \sim N\left(3.5, \frac{49}{552}\right)$	M1
	$P(3.4 < \bar{T} < 3.6) = P\left(\frac{3.4-"3.5"}{"\sqrt{\frac{49}{552}}"} < Z < \frac{3.6-"3.5"}{"\sqrt{\frac{49}{552}}"}\right) = \left[P(-0.34 < Z < 0.34)\right]$	M1 A1
	= 0.6331 – (1 – 0.6331) (Calculator gives 0.6314…)	M1
	= 0.2662 (Calculator gives 0.2628…) awrt 0.263 to 0.266	A1
		(5)
(d)	Large/ independent/ random sample allows use of CLT	B1
		(1)
	Notes	**Total 10**

(a)	**B1**	For the correct distribution stated (need uniform and correct interval) Allow U[0, 7] A fully correct pdf implies B1 e.g. $f(x) = \begin{cases} \frac{1}{7} & 0 \leq x \leq 7 \\ 0 & \text{otherwise} \end{cases}$
(b)	**B1**	For 3.5
	M1	For a correct method for finding the standard deviation
	A1	awrt 2.02 (Allow $\frac{7}{\sqrt{12}}$ or $\frac{7\sqrt{3}}{6}$ oe)
(c)	**M1**	For writing or using $N\left(3.5, \frac{49}{552}\right)$ oe Allow $N\left(3.5, \frac{2.02^2}{46}\right)$ or ft from part (b) e.g. if Po(7) given in part (a) allow $N\left(7, \frac{7}{46}\right)$
	M1	For standardising using either 3.4 or 3.6 and their mean and standard deviation
	A1	For a fully correct expression for either 3.4 or 3.6. May be implied by ± awrt 0.34
	M1	For $p - (1 - p)$ or $2(p - 0.5)$ oe
	A1	awrt 0.263 to 0.266
(d)	**B1**	Any suitable assumption

Question Number	Scheme	Marks
5 (a)	It is not a statistic as it involves unknown [population parameters]	B1
		(1)
(b)	An estimator for μ is unbiased if its expected value is equal to μ	B1
		(1)
(c)	$\mathrm{E}(U_1) = 3\mathrm{E}(X_1) - 2\mathrm{E}(X_2)$ or $\mathrm{E}(U_2) = \frac{1}{4}\left(\mathrm{E}(X_1) + 3\mathrm{E}(X_2)\right)$	M1
	$\mathrm{E}(U_1) = 3\mu - 2\mu = \mu$ (therefore unbiased)	A1cso
	$\mathrm{E}(U_2) = \frac{1}{4}(\mu + 3\mu) = \mu$ (therefore unbiased)	A1cso
		(3)
(d)	$\mathrm{Var}(U_1) = 9\mathrm{Var}(X_1) + 4\mathrm{Var}(X_2)$ or $\mathrm{Var}(U_2) = \frac{1}{16}\mathrm{Var}(X_1) + \frac{9}{16}\mathrm{Var}(X_2)$	M1
	$\left[\mathrm{Var}(U_1) =\right] 13\sigma^2$	A1
	$\left[\mathrm{Var}(U_2) =\right] \frac{5}{8}\sigma^2$	A1
	As $\mathrm{Var}(U_1) > \mathrm{Var}(U_2)$ U_2 is the most efficient estimator for μ	A1
		(4)

	Notes	**Total 9**
(a)	**B1**	for a correct explanation, must include unknown
(b)	**B1**	for a correct explanation that refers to expected *X*. Allow $\mu - \mathrm{E}(X) = 0$, but bias = 0 is B0
(c)	**M1**	for use of $a\mathrm{E}(X_1) + b\mathrm{E}(X_2)$ May be implied by $3\mu - 2\mu$ or $\frac{1}{4}(\mu + 3\mu)$
	A1cso	for a correct solution for $\mathrm{E}(U_1)$ with no incorrect working Condone missing notation. Condone missing subscripts
	A1cso	for a correct solution for $\mathrm{E}(U_2)$ with no incorrect working seen Condone missing notation. Condone missing subscripts
(d)	**M1**	for use of $a^2\mathrm{Var}(X_1) + b^2\mathrm{Var}(X_2)$
	A1	Allow $9\sigma^2 + 4\sigma^2$
	A1	Allow $\frac{1}{16}\sigma^2 + \frac{9}{16}\sigma^2$ or $\frac{5}{8}\sigma^2$ oe
	A1	for U_2 with a correct reason
		NB It is possible to score M1 A0 A0 A1 if $\mathrm{Var}(U_1)$ and $\mathrm{Var}(U_2)$ are correct

Question Number	Scheme	Marks
6 (a)	$M \sim N(80,100)$ $W \sim N(69,25)$ $X = M_1 + M_2 + M_3 + M_4 + M_5 + M_6 + W_1 + W_2 + W_3$	
	$X \sim N(687,675)$	M1 A1
	$P(X > 700) = P\left(Z > \frac{700-687}{\sqrt{675}}\right) = P(Z > 0.500...)$	M1
	$(= 1 - 0.6915) = 0.3085$ (Calculator gives 0.3084)	A1
		(4)
(b)	Let Y = Number of men in the lift	
	$Y \sim N(80x, 100x)$	M1
	$P(Y > 700) = P\left(Z > \frac{700-80x}{10\sqrt{x}}\right) < 0.025$	M1
	$\frac{700-80x}{10\sqrt{x}} > 1.96$	B1
	$80x + 19.6\sqrt{x} - 700[< 0]$ | $6400x^2 - 112384.16x + 490000[> 0]$	M1
	Solving leading to $\sqrt{x} < 2.838...$ | Solving leading to $x < 8.05...$	M1
	So c = 8 (people)	A1
		(6)

	Notes		Total 10
(a)	**B1**	for setting up normal distribution with mean 687	
	B1	for a correct variance (675) or for standard deviation ($15\sqrt{3}$)	
	M1	for standardising with 700, 687 and their standard deviation	
	A1	for answer between 0.308 – 0.309	
(b)	**M1**	for setting up normal distribution with mean 80x and variance 100x (may be implied by use of sd = $10\sqrt{x}$) Allow any letter	
	M1	for standardising with 700, their mean and their standard deviation (if not stated then these must be correct)	
	B1	for an equation or inequality set = to 1.96 (Allow – 1.96)	
	M1	for a correct 3TQ ft their mean and standard deviation	
	M1	for an attempt to solve their 3TQ with either $\sqrt{x} < ...$ or $x < ...$ Allow = instead of < Condone > or $\geqslant$ If the answer is incorrect then we must see use of the quadratic formula/completing the square (Allow one error)	
	A1	cao	

Question Number	Scheme	Marks
7 (a)	H_0 : The observed distribution can be modelled by a discrete uniform distribution H_1 : The observed distribution cannot be modelled by a discrete uniform distribution	B1 (1)
(bi)	See table below	B1 M1
	$X^2 = \sum \frac{(O-E)^2}{E}$ or $\sum \frac{O^2}{E} - 6x$; $\frac{230}{x}$ or $\frac{6x^2+230}{x} - 6x$	M1 ; A1
	$\nu = 6-1=5$; $c\,{}^2_5(0.05) = 11.070 \Rightarrow$ CR: X^2 … 11.070	B1 ; B1
	Do not reject H_0 if $\frac{230}{x}$,, '11.070' or $\frac{6x^2+230}{x} - 6x$,, '11.070'	M1
	x … 20.7768... So $x = 21$	A1 (8)
(bii)	Hence the die was rolled "21" × 6 = 126 times	M1 A1 (2)

Observed	Expected	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$
$x + 6$	x	$\frac{36}{x}$	$\frac{(x+6)^2}{x}$
$x - 8$	x	$\frac{64}{x}$	$\frac{(x-8)^2}{x}$
$x + 8$	x	$\frac{64}{x}$	$\frac{(x+8)^2}{x}$
$x - 5$	x	$\frac{25}{x}$	$\frac{(x-5)^2}{x}$
$x + 4$	x	$\frac{16}{x}$	$\frac{(x+4)^2}{x}$
$x - 5$	x	$\frac{25}{x}$	$\frac{(x-5)^2}{x}$
Total = $6x$	Total = $6x$	Total = $\frac{230}{x}$	Total = $\frac{6x^2+230}{x}$

Notes **Total 11**

(a)	**B1**	for both hypotheses correct Allow H_0 : the die is not biased H_1 : the die is biased
(bi)	**B1**	for expected frequency = x
	M1	for one correct $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$ ft their expected frequency
	M1	for an attempt at X^2 ft their values (At least 4 of these need to be seen and added)
	A1	for either $\frac{230}{x}$ or $\frac{6x^2+230}{x} - 6x$
	B1	for $\nu = 6-1=5$ May be implied by a correct critical value
	B1	for a correct critical value ft their DOF (**NB** common error is $\nu = 4$ so $c\,{}^2_4(0.05) = 9.488$)
	M1	for either $\frac{230}{x}$,, their CV or $\frac{6x^2+230}{x} - 6x$,, their CV Allow < rather than ,,
	A1	for $x = 21$ provided the previous M mark has been awarded
(bii)	**M1**	for their 21 × 6 Allow 6 × x or the answer to 6 × their value for x
	A1	cao

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