Mark Scheme (Results)

October 2017

Pearson Edexcel International A Level
in Statistics S1 (WST01/01)

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 2017
Publications Code WST01_01_1710_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.

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: 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.

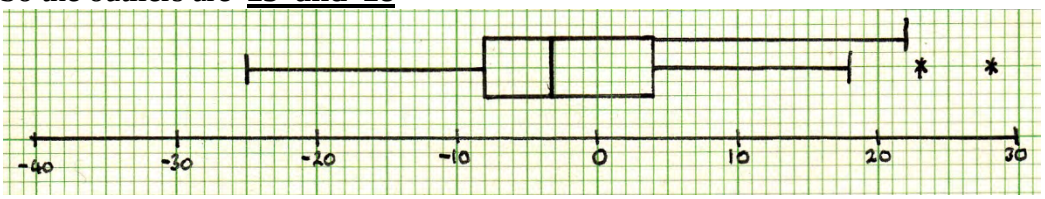
   - bod – benefit of doubt
   - ft – follow through
   - the symbol $\surd$ will be used for correct ft
   - cao – correct answer only
   - cso - correct solution only. There must be no errors in this part of the question to obtain this mark
   - isw – ignore subsequent working
   - awrt – answers which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

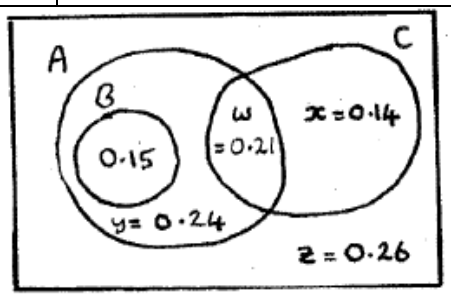
7. Ignore wrong working or incorrect statements following a correct answer.

Question Number	Scheme	Marks
1. (a)	[Range =] **63**	B1 (1)
(b)	[IQR =] **18**	B1 (1)
(c)	[Q_2 =] $(-8)+\frac{20}{33}\times 8$ or $(0)-\frac{13}{33}\times 8$ [NB (n +1) will have 20.5 or 12.5] = –3.1515... awrt **– 3.15**	M1 A1 (2)
(d)	[Q_3 =] mid-point of [0, 8] group so therefore = 4	B1cso (1)
(e)(i)	IQR = 4 – – 8 = 12 so upper limit is 4 + 1.5×12 = **22** lower limit is – 8 – 1.5×12 = **– 26** So the outliers are **23 and 28**	M1 A1 A1
(ii)		M1 A1 A1 (6)
(f)	Interquartile range is smaller (12 compared to 18) or range is smaller (53 v 63) Median is closer to zero (– 3.15 is closer than 5) So they have improved	B1 B1 dB1 (3)
		[Total 14]

	Notes
(c)	M1 for a correct fraction and ×8 (ignore end point) A1 for awrt – 3.15 (allow use of n + 1 leading to – 3.03) Accept $-\frac{104}{33}$ if box plot is OK or 3sf value is quoted in (f)
(d)	B1cso for a clear argument with no incorrect working seen. Allow 4.14… from 7.25 for (n + 1) case
(e)(i)	M1 for at least one correct calculation e.g. $4+1.5(8--4)$ (implied by one correct limit) 1st A1 for one correct limit 2nd A1 for both correct limits and the two correct outliers identified
(ii)	M1 for a box with 2 whiskers (one at each end) 1st A1 for – 8 and 4 and ft Q_2 between them and lower whisker ending at – 25 **no** outliers 2nd A1 for upper whisker ending at 18 **or** 22 and 2 outliers marked at 23 and 28
(e)(ii)SC	Two incorrect outliers in (e)(i), ft both A1s in (ii) using their outliers provided in [– 25, 28]
(f)	1st B1 for a statement about range or IQR saying that 2nd estimates are better Allow range or IQR" has decreased" or "is smaller" o.e. 2nd B1 for a statement about medians saying that 2nd one is closer to zero Don't allow "decreased" or "smaller" unless clearly using |median| or say e.g. 3.15 < 5 3rd dB1 dep on at least one other B1 for concluding that they have improved based on change in median or range/IQR. Must clearly state "improved" not just "yes"

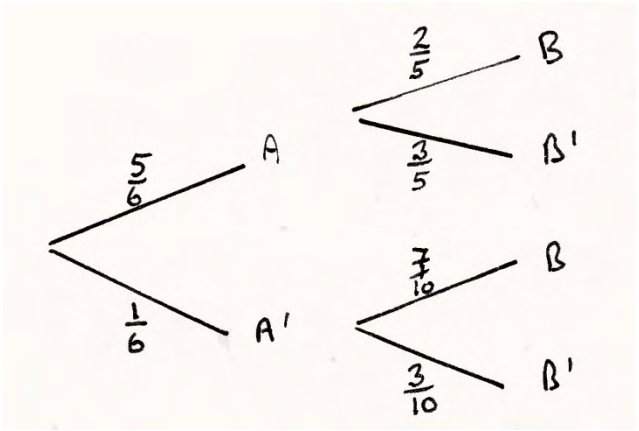
Question Number	Scheme	Marks
2. (a)	B and C (only)	B1 (1)
(b)	$\mathrm{P}(A \cap C) = 0.6 \times 0.35$ so [***w* =**] **0.21**	B1cso (1)
(c)	$x = \mathrm{P}(C) - w$ = **0.14** $y = \mathrm{P}(A) - w - \mathrm{P}(B)$,= **0.24** $z = 1 - \mathrm{P}(A \cup C)$ = **0.26**	B1 M1,A1 B1ft (4)
(d)	[$x + y$ =] **0.38**	B1ft (1)
(e)	[$\mathrm{P}(B \cup C) = 0.15 + 0.35$] = **0.5**	B1cao (1)
(f)	$\left[\mathrm{P}\left(A \mid [B \cup C]\right)\right] = \dfrac{\mathrm{P}(A \cap [B \cup C])}{\mathrm{P}(B \cup C)} = \dfrac{0.15 + 0.21}{\text{"0.5"}}$ = **0.72**	M1A1ft A1 (3) **[Total 11]**

	Notes
(a)	B1 for just B and C [NB Just writing $\mathrm{P}(B \cap C) = 0$ is B0]
(b)	B1cso for 0.21 clearly from $\mathrm{P}(A) \times \mathrm{P}(C)$ or 0.6×0.35 and no incorrect statements seen
(c)	1st B1 for $x = 0.14$ M1 for a correct expression for y A1 for $y = 0.24$ 2nd B1ft for $z = 0.26$ or correct ft of their values to make sum = 1 (provided all probs) These values may be seen in correct regions in the Venn diagram
(d)	B1ft for their $x + y$ or 0.38
(e)	B1 for 0.5 or exact equivalent
(f)	M1 for a correct ratio of probabilities formula num of: $\mathrm{P}(B \cup C \cap A)$ or $\mathrm{P}(A \cap [B \cup C])$ with brackets and some correct probability, ft their (e) May be implied by correct ratio. 1st A1ft for a numerator of 0.15 + 0.21 and a denominator of their (e) Can award M1A1ft for $\dfrac{0.15 + 0.21}{\text{"their 0.5"}}$ even if their formula is incorrect 2nd A1 for 0.72 or exact equivalent e.g. $\frac{18}{25}$



Question Number	Scheme	Marks
3. (a)	$\{P(L < 45) = 0.4\} \Rightarrow \frac{45-\mu}{\sigma} = -0.2533$ <u>or</u> $\Rightarrow 45 - \mu = -0.2533\sigma$ (o.e.) $45 + 0.2533\sigma = \mu$ **(*)**	M1 A1cso (2)
(b)	$P(L < 35) = 0.15 \Rightarrow \frac{35-\mu}{\sigma} = -1.0364$ e.g. $35 + 1.0364\sigma = \mu$	M1 A1 (2)
(c)	Solving: $10 - 0.7831\sigma = 0$ $\sigma = 12.7697\ldots$ awrt **<u>12.8</u>** μ = awrt **<u>48.2</u>**	M1 A1 A1 (3)
(d)(i)	$P(L > 35 \mid L < 45) = \frac{P(35 < L < 45)}{P(L < 45)} = \frac{0.25}{0.15 + 0.25} = \frac{5}{\underline{8}}$ (o.e.)	A1
(ii)	$P(L < 45 \mid L > 35) = \frac{P(35 < L < 45)}{P(L > 35)} = \frac{0.25}{0.60 + 0.25} = \frac{5}{\underline{17}}$ or awrt 0.294	M1 A1 (3)
(e)	Prob. of a yellow stick from Hei is $\frac{5}{8}$ which is > prob. of $\frac{5}{17}$ for Tang **So more likely to be Hei**	B1ft dB1ft (2) **[Total 12]**

	Notes
	Mark parts (a), (b) and (c) as one part but must see the "show that" for (a) somewhere
(a)	M1 for attempting to standardise with 45, μ and σ Allow $\pm$ and allow z = awrt 0.25 A1cso for sight of $P(L < 45) = 0.4$ (o.e.) and 0.2533 leading to given ans. [0.2533471… from calc]
(b)	M1 for standardising with 35 μ and σ and setting equal to a z value with $1 < \lvert z \rvert < 1.05$ A1 for any correct equation, $z = 1.04$ or better and correct signs
(c)	M1 for solving their 2 linear equations in μ and σ – reducing to an equation in 1 variable 1st A1 for σ = awrt 12.8 (NB use of 1.04 gives 12.7113… so we penalise that here) 2nd A1 for μ = awrt 48.2 [allow 48.3 if 12.8 used in a correct eqn e.g. $35+1.04\times12.8$ or better]
(d)	M1 for a correct expression [num = $P(35 < L < 45)$] with <u>some</u> correct values substituted This M1 may be implied by one of the correct probabilities for (i) or (ii)
(i)	1st A1 for $\frac{5}{8}$ or an exact equivalent e.g. 0.625
(ii)	2nd A1 for $\frac{5}{17}$ or awrt 0.294
(e)	1st B1ft for a correct comparison of their <u>probabilities</u> from (d) "probs" $\notin [0, 1]$ is B0 2nd dB1ft for choosing Hei (dependent on a suitable reason that it is more likely to be hers) Allow e.g. "Hei, because her prob is greater" to score B1B1 provided (d)(i) > (d)(ii)
ft	Allow "Tang" if their (d)(i) < their (d)(ii) and a correct comparison stated.

Question Number	Scheme	Marks
4. (a)	[Let P(A) = p]	
	$0.4p + 0.7(1-p) = 0.45$	M1A1
	$0.25 = 0.3p$	M1
	$\boldsymbol{p} = \frac{5}{6}$	A1
		B1ft B1
		(6)
(b)	$\left[\mathrm{P}(A' \mid B')\right] = \frac{\frac{1}{6} \times 0.3}{0.55}$	M1
	$= \frac{1}{11}$	A1
		(2) **[Total 8]**

	Notes
(a)	1st M1 for $0.4p$ or $0.7(1-p)$ seen in an equation for p 1st A1 for a fully correct equation for p
ALT	1st M1 for attempt at 2 sim' eq'ns in p and q Allow one error. $0.4p + 0.7q = \frac{9}{20}$ and $0.6p + 0.3q = \frac{11}{20}$ 1st A1 for any correct equation in p or q
	2nd M1 for simplifying their linear equation with at least 2 terms in p or q to $a = bp$ or bq 2nd A1 for P(A) = $\frac{5}{6}$ or exact equiv e.g. $0.8\dot{3}$ (may be seen on their tree diagram) 1st B1ft for 1st 2 branches i.e. $\frac{5}{6}$ and $\frac{1}{6}$ (follow through their P(A)) 2nd B1 for 2nd 4 branches i.e. $\frac{3}{5}$ and $\frac{3}{10}$
(b)	M1 for a ratio of probabilities ft their numerator from their tree diagram but denom = 0.55 A1 for $\frac{1}{11}$ or exact equivalent e.g. $0.\dot{0}\dot{9}$
SC	**[P(A)** $\neq \frac{5}{6}$ **]** award M1A0 for $\frac{\mathrm{P}(A') \times \frac{3}{10}}{\mathrm{P}(A) \times \frac{3}{5} + \mathrm{P}(A') \times \frac{3}{10}}$ ft their P(A) and $\mathrm{P}(A') = 1 - \mathrm{P}(A)$

Question Number	Scheme	Marks
5. (a)	$\left[\overline{x} = \frac{283}{10}\right] = \mathbf{\underline{28.3}}$	B1
	$\sigma_x^{\,2} = \frac{9011}{10} - 28.3^2$	M1
	$= 100.21$ accept awrt **<u>100</u>**	A1
		(3)
(b)	$\overline{y} =$ **<u>30.61</u>** (allow 30.6)	B1
	$\sigma_y^{\,2} =$ **<u>54.63</u>** (allow 54.6)	B1
		(2)
(c)	$0.659 = \frac{S_{xy}}{S_{xx}}$ and $S_{xx} = 10\sigma_x^{\,2}$ [= 1002.1]	M1M1
	$S_{xy} = 0.659 \times 1002.1$ **{=** 660.3839 (659~660.4]}	A1
	$r = \frac{660.3839}{\sqrt{1002.1 \times 546.3}}$	M1
	= awrt **<u>0.892~0.893</u>**	A1
		(5)
(d)	Value of r is close to 1 so it <u>does</u> support the use of a linear regression model	B1
		(1)
(e)	$y = 12.0 + 0.659 \times 35$ **=** [35.065]	M1
	Proposed salary is **<u>$ 35 065</u>** (awrt $35 100)	A1
		(2)
		[Total 13]

	Notes
(a)	M1 for a correct expression for variance (no $\sqrt{}$) but ft their 28.3 A1 for awrt 100
(c)	1st M1 for a correct expression using b = 0.659 that connects this value with S_{xx} and S_{xy} 2nd M1 for a correct method for S_{xx} (value not required) 1st A1 for a value for S_{xy} in (659, 660.4) <u>or</u> 0.659×1002.1 3rd M1 for a correct expression for r (ft their S_{xx} and their S_{xy}) 2nd A1 for awrt 0.892 or 0.893
(d)	B1 if $0.5 \leqslant \lvert r\rvert \leqslant 1$ for saying that it <u>does</u> support **and** giving a suitable comment, e.g. "strong" correlation, about r being close to 1 (or -1) **$\lvert r\rvert > 1$ is B0**
SC	If $\lvert r\rvert < 0.5$ allow it <u>does not</u> support **with** supporting comment about r close to 0
(e)	M1 for substituting 35 into the given regression equation (may be implied by 35.065) A1 for 35 065 (i.e. a correct value <u>and</u> multiplying by 1000) **must be at least 3sf** Accept "35.065 thousand dollars"
NB	$ 35 000 is A0 (could have come from the 35 in the question)

Question Number	Scheme	Marks
6. (a)	[Discrete] **uniform** (BUT continuous uniform is B0)	B1 (1)
(b)	$P(D = 3) + P(D = 1) \times P(D = 2) = \frac{1}{4} + \frac{1}{4} \times \frac{1}{4} = \frac{5}{16}$ **(*)**	M1A1cso (2)
(c)	$[P(D = 1) \times P(D = 1) = \frac{1}{4} \times \frac{1}{4}$ or $1 - \left(\frac{1}{4} + \frac{5}{16} + \frac{5}{16} + \frac{1}{16}\right) =] \frac{1}{16}$	B1 (1)
(d)	$E(X) = 0 + 2 \times \frac{1}{4} + 3 \times \frac{5}{16} + 4 \times \frac{5}{16} + 5 \times \frac{1}{16} = \mathbf{3}$	M1A1 (2)
(e)	$E(X^2) = 0 + 2^2 \times \frac{1}{4} + 3^2 \times \frac{5}{16} + 4^2 \times \frac{5}{16} + 5^2 \times \frac{1}{16} = [\frac{166}{16}$ or $\frac{83}{8}$ or $10.375]$ $\text{Var}(X) = \frac{166}{16} - 3^2$ $\sigma_X{}^2 = 1.375$ or $\frac{11}{8}$	M1 dM1 A1 (3)
(f)	r: 1, 2 $P(R = r)$: $\frac{3}{4}$, $\frac{1}{4}$	M1 A1 (2)
(g)	$E(R) = 1 \times \frac{3}{4} + 2 \times \frac{1}{4}$ [= 1.25 o.e.] $E(Y) = 2E(R) + 0.5$ y: 2.5, 4.5 $P(Y = y)$: $\frac{3}{4}$, $\frac{1}{4}$ $E(Y) = 2.5 \times \frac{3}{4} + 4.5 \times \frac{1}{4}$ $= 2.5 + 0.5 = 3$ **(*)**	M1 M1 A1cso (3)
(h)	$R = 1$ so $Y = 2.5 \Rightarrow X = D = 2$ or 3 or 4 so $D = 3$ or 4 work and prob $= \frac{1}{4} + \frac{1}{4}$ May use $P(X > 2.5 \mid R = 1) = \frac{2}{3}$ then prob will be $\frac{2}{3} \times \frac{3}{4}$ $R = 2$ so $Y = 4.5 \Rightarrow D = 1$ then $X = 0$, 3 or 4 or 5 so $X = 5$ only prob $= \frac{1}{16}$ So $P(X > Y) = \frac{1}{4} + \frac{1}{4} + \frac{1}{16} = \frac{9}{16}$	M1 M1 A1 (3) **[Total 17]**

	Notes
(b)	M1 for a correct expression in terms of $P(D)$ or with $\frac{1}{4}$s for $P(X = 3)$ A1cso M1 scored and no incorrect working seen [$P(X = 0) + P(X = 3)$ is M0A0 if identified!]
(d)	M1 for an attempt i.e. an expression with at least 3 correct products seen A1 for 3 or an exact equivalent e.g. $\frac{48}{16}$
(e)	1st M1 for an attempt i.e. an expression with at least 3 correct products seen[implied by $\frac{166}{16}$] 2nd dM1 dep on 1st M1 for use of $\text{Var}(X) = E(X^2) - E(X)^2$ must see values but ft their values A1 for 1.375 or an exact equivalent
(f)	M1 for one correct value of r and it's associated probability A1 for a fully correct probability distribution – needn't be in a table
(g)	1st M1 for correct expression for $E(R)$ [ft their (f)] or 1.25 or correct distribution for Y[ft (f)] 2nd M1 for correct use of $E(Y) = 2E(R) + 0.5$ or correct expr'n for $E(Y)$ [ft (f)] [$\Rightarrow$ 1st M1] A1cso for 3 with no incorrect working seen provided both Ms are scored
(h)	1st M1 for cases where $R = 1$ **and** prob. 2nd M1 for cases where $R = 2$ **and** prob A1 for $\frac{9}{16}$ or exact equivalent

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