Pearson Edexcel

Mark Scheme (Results)

Summer 2019

Pearson Edexcel International Advanced 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

Summer 2019
Publications Code WST01_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.

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
   - $\sqsubset$ 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. 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 (a)	$\bar{x} = \frac{58}{40} =$ **1.45**	B1
	$\sigma^2 = \frac{84.829}{40} - 1.45^2$	M1
	$= 0.018225$ = awrt **0.0182**	A1
		(3)
(b)	New mean = **145**	B1ft
	New σ = **13.5**	B1
		(2)
(c)(i)	**Reason** e.g. mean of two extra children is the same as the original mean	M1
	Conclusion the mean is therefore unchanged or = **145**	A1
(ii)	**Reason** e.g. extra children more than 1 sd from mean so increased spread	M1
	Conclusion therefore standard deviation will increase	A1
		(4)
		[9]

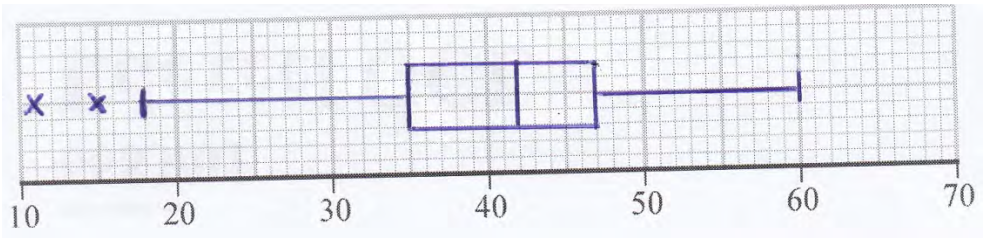
Notes

(a)
B1 for a correct mean (accept an exact fraction)
M1 for a correct expression for σ^2 (or s^2) (ft their mean and condone inside square root)
A1 for awrt 0.0182 (NB $s^2 = 0.0186923\ldots$ awrt 0.0187) **Correct ans only** 2/2 [No fraction]

(b)
1st B1ft for new mean = 145 or $100 \times$ their $\bar{x}$
2nd B1 for new s.d. = awrt 13.5 (accept $s = 13.6719\ldots$ or awrt 13.7)

(c)(i)
1st M1 for a suitable reason. May see recalculation e.g. $\frac{"145" \times 40 + 130 + 160}{42}$ (o.e.)
e.g. "both 15 away from the mean" or " both same distance from the mean" or "mean of new values is 145 or the same"
1st A1 for 145 or 1.45 or "no change" but M1 must be seen
[no further comment needed if answer matches their (b) or (a)]

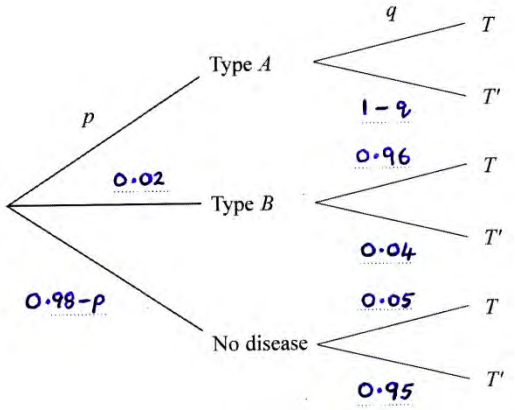
(ii)
2nd M1 for a suitable reason but must have idea that the "gap" (= 15) > 1 st. dev. [ft $\sigma < 15$]
2nd A1 for stating standard deviation will be greater (o.e.) [M1 must be seen]
Calculations (You may see)
e.g. $\Sigma y^2 = 84.829 + 1.3^2 + 1.6^2 = 89.079$ leading to $\sigma = \sqrt{0.01842...} = 0.13575\ldots$ or 13.6 (cm)
or $\frac{89.079}{42} = 2.1209... > \frac{84.829}{40} = 2.1207...$ but $\frac{\sum x}{n}$ stays the same so σ greater
BUT M0A0 unless we see mention of 15 (cm) or 1.5 (m) being more than 1 sd

Question Number	Scheme	Marks
2. (a)(i)	[IQR = 47 – 33 =] **14**	B1
(ii)	[Range = 54 – 11 =] **43**	B1 (2)
(b)	e.g. $Q_2 - Q_1 (=9) > (5=) Q_3 - Q_2$	M1
	Therefore negative (skew)	A1 (2)
(c)	$25 \rightarrow 37 \Rightarrow$ new $Q_1 = 35$ (may be on box plot)	B1
	[$54 \rightarrow 60$ (implies upper whisker now at 60) but no change to Q_3]	
	New IQR = 12 so need to re-calculate for outliers	M1
	Outliers now [> 47 + 18 = 65 or] < 35 – 18 = 17	A1
	Box Plot 	
	Box and two whiskers with median still at 42	M1
	Lower quartile at their 35 ($\neq 33$) and upper quartile unchanged at 47	A1ft
	Two outliers at 11 and 15	A1
	Lower whisker at 18 (or 17) and upper whisker at 60	A1 (7)
(d)	The value of pmcc is small or weak correlation (o.e.)	M1
	Therefore the complaint is not supported	A1 (2)
		[13]

	Notes
(a)(i)	1st B1 for 14 2nd B1 for 43
(b)	M1 for a suitable reason or calculation (allow longer whisker on left etc) A1 for negative skew (dep on M1 seen) "left skew" etc is A0 [Condone incorrect "9" or "5"]
(c)	B1 for new lower quartile at 35 (stated or on box plot) 1st M1 for finding the new IQR (< 14) and attempting to re-calculate for outliers 1st A1 for at least the correct lower limit of 17 seen 2nd M1 for drawing a box with only two whiskers and median at 42 (all points ± 0.5 square) 2nd A1ft for lower quartile of "35" (changed from 33) and upper quartile unchanged at 47 3rd A1 for only two outliers at 11 and 15 (no overlap with whisker) 4th A1 for lower whisker ending at 18 (or 17) and upper whisker ending at 60 Correct box plot scores all except 1st M1A1 (i.e. 5/7) this M1A1 requires some working
(d)	M1 for comment that pmcc is "small" so little correlation (just saying < 0 is not enough) Allow e.g. "not significant" or "not relevant" or $-0.5 < r < 0.5$ or "not close to – 1 " but "no correlation" is M0 A1 for suggesting the complaint is not supported e.g. "little evidence to support claim" Dep on M1 seen NB M1A0 is possible

Question Number	Scheme	Marks
3. (a)	0.02 and $0.98-p$ correctly placed [no mixing of % and probability] 0.96 and 0.05 plus $1-q$, 0.04, 0.95 correctly placed	B1 B1 (2)
(b)	$P(T) = pq + 0.02\times0.96 + (0.98-p)\times0.05 = 0.169$ $\{pq - 0.05p = 0.1008\}$ $P(\text{do not have disease} \mid T) = \dfrac{"(0.98-p)"\times0.05}{0.169} = \dfrac{41}{169}$ So $p =$ **0.16** e.g. $0.16q - 0.16\times0.05 = 0.1008$ $q =$ **0.68**	M1; A1 M1A1ft A1 dM1 A1 (7)
(c)(i)	$P(\text{type } A \mid T \text{ and not type } B) = \dfrac{pq}{pq+(0.98-p)\times0.05} = \dfrac{0.1088}{0.1088+0.041}$ $= 0.7263\ldots$ awrt **0.726**	M1A1ft A1 (3)
(ii)	Should find test useful, doctor knows there is a much greater chance that the person has type A (0.73 compared to 0.16 or 0.163…[from $\frac{0.16}{0.98}$])	B1 (1) **[13]**

	Notes
(a)	1st B1 for remainder of 1st column probabilities (allow use of correct p so 0.82) 2nd B1 for remainder of 2nd column probabilities (allow use of correct q so 0.68 and 0.32)
	In (b) or (c) if p or q are used as ft in M or A marks they must be probabilities
(b)	1st M1 for attempt to form eq'n in p and q using $P(T) = 0.169$ [at least 2 of 3 correct prod's] 1st A1 for a fully correct equation in p and q or possibly just q (using their p see 3rd M1) 2nd M1 for use of a conditional prob (ratio of probabilities with num or den correct, allow ft on num) and $\frac{41}{169}$ to form an equation in p 2nd A1ft for a correct equation using values from their tree diagram 3rd A1 for solving to get $p = 0.16$ (or exact equivalent) 3rd dM1 (dep on 1st M1) for substituting their p into an equation for q (ft their p value) 4th A1 for $q = 0.68$ (or exact equivalent)
(c)(i)	M1 for an attempt at a conditional prob with numerator of their pq (num < denom) 1st A1ft for a correct ratio of probs (ft their values for p or q with at least one correct) 2nd A1 for awrt 0.726 (or exact fraction $\frac{544}{749}$)
(ii)	B1 If (c)(i) < 0.7 then B0 for suggesting test should be useful (accept "yes") plus statement: about increased prob or "more likely to have type A than no disease" or "prob of A is high"



Question Number	Scheme	Marks
4. (a)	[W = weight of a package delivered to Susie $W \sim N(510, 45^2)$]	
	$P(W < 450) = P\left(Z < \frac{450-510}{45}\right)$ or $P(Z < -1.3333)$	M1
	$= 1 - 0.9082$	M1
	$= 0.0918$ **[0.0912~0.0918]**	A1
		(3)
(b)	[$P(W > d) = 0.05$ implies] $\frac{d-510}{45} = 1.6449$	M1B1
	$d = 584.0205\ldots$ awrt **584**	A1
		(3)
(c)	[$P(W > 450 \mid W <$ "584.02…") =] $\frac{P(450 < W < "584.02...")}{P(W < "584.02...")}$	M1
	$= \frac{0.95 - "0.0918"}{0.95}$ or $\frac{"0.9082" - 0.05}{0.95}$	M1A1
	$= 0.903368\ldots$ awrt **0.904** or **0.903**	A1
		(4)
(d)	$\left(\frac{19}{20}\right)^4 \times \frac{1}{20} \times 5$	M1dM1
	$= 0.203626\ldots$ awrt **0.204**	A1
		(3)
		[13]

	Notes
	Correct answer only in (a), (c) or (d) scores all the marks for that part
(a)	1st M1 for standardising 450 with 510 and 45 (allow ±) 2nd M1 for $1 - p$ (where $0.90 < p < 0.99$) A1 for answer in the range 0.0912 to 0.0918 inclusive (calc. 0.09121133…)
(b)	M1 for standardising their letter d with 510 and 45 and setting equal to z value $1 < \lvert z \rvert < 2$ B1 for use of $z = \pm 1.6449$ or better (calc 1.644853626…) A1 for awrt 584 (calc 584.0184…)
Ans only	[awrt 584.02 scores 3/3 584 scores M1B0A1]
(c)	1st M1 for a correct ratio of probability expressions ft their answer to (b) where (b) > 450 2nd M1 for numerator of awrt 0.95 – their answer to (a) 1st A1 for a correct denominator of awrt 0.95 (dep on M1M1) NB a correct ratio of probabilities will score the 1st 3 marks 2nd A1 for awrt 0.904 or awrt 0.903
(d)	1st M1 for $k\,p^4(1-p)$ for any positive integer k and any probability p (allow $k = 1$) 2nd dM1 for $k = 5$ A1 for awrt 0.204

Question Number	Scheme	Marks
5. (a)	$E(X) = -2p - p + 0 + \frac{1}{2} + 3p \;\; ; \;\; = \frac{1}{2}$	M1 ; A1 (2)
(b)	$E(X^2) = 4p + p + 0 + 1 + 9p = [14p + 1]$	M1A1
	[Var(X) =] $E(X^2) - [E(X)]^2 = 14p + 1 - ("\frac{1}{2}")^2$	dM1
	So $14p + 0.75 = 2.5$	M1
	$p = \frac{1}{8}$	A1 (5)
(c)	Sum of probabilities = 1 implies $q = \frac{3}{8}$	B1ft (1)
(d)	P(Amar wins) = e.g. $P(X_1 > 0) + P(X_1 < 0) \times P([X_1 + X_2] > 0 \{\mid X_1 < 0\})$ or $P(X_1 = 2 \text{ or } 3) + P(X_1 = -2) \times P(X_2 = 3) + P(X_1 = -1) \times P(X_2 = 2 \text{ or } 3)$	M1
	Cases $X_1 = -2$ and $X_2 = 3$ so probability = p^2	M1
	$X_1 = -1$ and $X_2 \geqslant 2$ so probability = $p\left(p + \frac{1}{4}\right)$	
	Total probability = $p + 0.25 + p^2 + p(p + 0.25) = \frac{1}{8} + \frac{1}{4} + \frac{1}{64} + \frac{1}{64} + \frac{1}{32}$	A1ft
	$= \frac{7}{16}$	A1 (4)
(e)	[Although E(X) > 0 since] P(win) < 0.5	M1
	Amar should not play the game or "disagree"	A1 (2)
		[14]

	Notes
(a)	M1 for a correct expr'n for E(X) in p (at least 3 non-zero terms seen). May be implied by A1 A1 for $\frac{1}{2}$ (or exact equivalent e.g. $\frac{2}{4}$ or 0.5)
(b)	1st M1 for a correct expression for E(X^2) (at least 3 non-zero terms). May be implied by A1 1st A1 for $14p + 1$ or any fully correct expression in terms of p 2nd dM1 dep on 1st M1 for use of [Var(X) =] $E(X^2) - [E(X)]^2$ 3rd M1 for forming a linear equation in p using the 2.5 2nd A1 for $p = \frac{1}{8}$ or exact equivalent e.g. 0.125
(c)	B1ft for $q = \frac{3}{8}$ or exact equivalent e.g. 0.375 or $\frac{3}{4} - 3p$ $0<p<1$
(d)	1st M1 for identifying only the correct cases (any correct list, adding not needed) 2nd M1 for identifying all the cases where a 2nd spin is required and probabilities (no extras) 1st A1ft for correct expression for total probability (allow their $0 < p < 1$ or letter p) 2nd A1 for $\frac{7}{16}$ (or exact equivalent e.g. 0.4375) [$\frac{7}{16}$ with no incorrect working seen gets 4/4]
ALT	Allow P(loses) = $q + p(1 - p) + p(0.75 - p)$ only if 1 – P(loses) is seen
(e)	M1 for identifying that the important feature is that P(win) < 0.5 (o.e.) [ft their $\frac{7}{16}$ < 0.5] A1cao for concluding that he shouldn't play the game (dep on M1 seen & 0.375 < (d) < 0.5)

If they think E(X^2) = Var(X) get $p = \frac{3}{28}, q = \frac{3}{7}$ and up to (b) M1A1M0M1A0 (c) B1ft and if they get $\frac{319}{784}$ in (d) it implies M1M1A1A0 there and access to (e)

Question Number		Scheme	Marks
6.	**(a)**	$\left[\sum y = 16 \times 20.5 = 328\right]$ $S_{yy} = 8266 - \frac{328^2}{16}$	M1
		$= 1542$ (allow awrt 1540)	A1
		$[r =] \frac{-630.9}{\sqrt{368.16 \times "1542"}}$	M1
		$= -0.837336\ldots$ awrt **$\underline{-0.837}$**	A1
			(4)
	(b)	As the distance from the hospital increases the percentage of referrals decreases (o.e.) e.g. smaller % of patients attend from clinics further away	B1
			(1)
	(c)	e.g. Points close to a straight line (of negative gradient) so $\underline{\text{does}}$ support belief	B1
			(1)
	(d)	$b = \frac{-630.9}{368.16}$ $[= -1.7136\ldots]$	M1
		$a = 20.5 - "-1.7136..." \times 8.1$ $[= 34.3806\ldots]$	M1
		$y = 34.38\ldots - 1.7136\ldots x$ **$\underline{y = 34.4 - 1.71x}$**	A1, A1
			(4)
	(e)	[On average] each km further from the hospital reduces the % attendance by 1.7%	B1
			(1)
	(f)	Correct line drawn on scatter diagram (use overlay within guidelines)	B1
			(1)
	(g)	Correct point circled (3.2,19) [Allow coords stated instead of point circled but if both, prioritise circled point]	B1 (1)
			[13]

	Notes
(a)	1st M1 for an attempt at a correct expression for S_{yy} (ft their 328 provided intention is Σy) 1st A1 for 1542 (allow awrt 1540 it leads to $r = -0.83788\ldots$ and scores 2nd A0) 2nd M1 for a correct expression for r (ft their S_{yy} but use of 8266 is M0 here) 2nd A1 for awrt -0.837 (ans only 4/4; awrt -0.838 M1A1M1A0; -0.84 M1A0M1A0)
(b)	B1 for an interpretation of negative correlation $\underline{\text{in context}}$ (just "strong neg correlation" B0)
(c)	B1 for " points close to a straight line" $\underline{\text{and}}$ stating does support manager's belief $\underline{\text{or}}$ allow "r is close to -1" $\underline{\text{or}}$ "strong (negative) correlation" $\underline{\text{and}}$ supports manager's claim $\underline{\text{or}}$ for a curve drawn on scatter diagram $\underline{\text{and}}$ comment that non-linear model may be better
(d)	1st M1 for a correct expression for b 2nd M1 for a correct expression for a (ft their value of b or even letter b in correct formula) 1st A1 (dep on 1st M1) for b = awrt -1.71 in an equation in y and x (no fractions) 2nd A1 (dep on 2nd M1) for a = awrt 34.4 in an equation in y and x
(e)	B1 for a comment with their b (<0) relating distance from hospital to % attendance/referrals Allow "as distance increases by 1 the % referrals decreases by 1.7" (o.e.)
(f)	B1 for drawing the line on scatter diagram (within guidelines of overlay-check both graphs)
(g)	B1 for correct point on scatter diagram circled (more than one point circled is B0)