

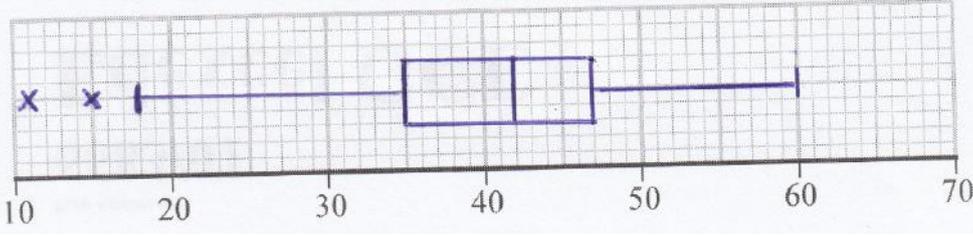


Mark Scheme (Results)

Summer 2019

Pearson Edexcel International Advanced Level
In Statistics S1 (WST01/01)

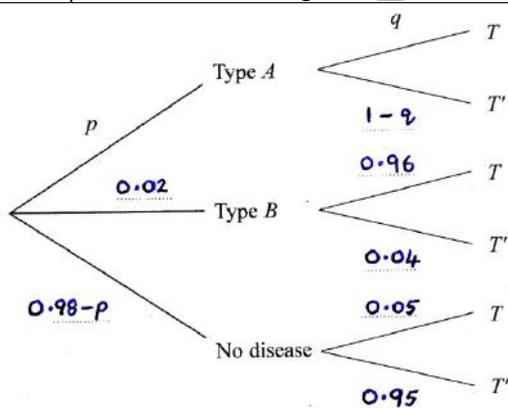
Question Number	Scheme	Marks
<p>1 (a)</p> <p>(b)</p> <p>(c)(i)</p> <p>(ii)</p>	$\bar{x} = \frac{58}{40} = \underline{\underline{1.45}}$ $\sigma^2 = \frac{84.829}{40} - 1.45^2$ $= 0.018225 \quad = \text{awrt } \underline{\underline{0.0182}}$ <p>New mean = <u>145</u></p> <p>New $\sigma = \underline{\underline{13.5}}$</p> <p>Reason e.g. mean of two extra children is the same as the original mean Conclusion the mean is therefore unchanged or = <u>145</u></p> <p>Reason e.g. extra children <u>more than 1 sd</u> from mean so increased spread Conclusion therefore standard deviation will increase</p>	<p>B1</p> <p>M1</p> <p>A1</p> <p>(3)</p> <p>B1ft</p> <p>B1</p> <p>(2)</p> <p>M1</p> <p>A1</p> <p>M1</p> <p>A1</p> <p>(4)</p> <p>[9]</p>
Notes		
<p>(a)</p> <p>(b)</p> <p>(c)(i)</p> <p>(ii)</p>	<p>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 $s^2 = 0.0186923\dots$ awrt 0.0187)Correct ans only 2/2 [No fraction]</p> <p>1st B1ft for new mean = 145 <u>or</u> 100×their \bar{x} 2nd B1 for new s.d. = awrt 13.5 (accept $s = 13.6719\dots$ or awrt 13.7)</p> <p>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” <u>or</u> “both same distance from the mean” <u>or</u> “mean of new values is 145 <u>or</u> the same”</p> <p>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)]</p> <p>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 <u>greater</u> (o.e.) [M1 must be seen]</p> <p>Calculations (You may see) e.g. $\Sigma y^2 = 84.829 + 1.3^2 + 1.6^2 = 89.079$ leading to $\sigma = \sqrt{0.01842\dots} = 0.13575\dots$ or <u>13.6</u> (cm) <u>or</u> $\frac{89.079}{42} = 2.1209\dots > \frac{84.829}{40} = 2.1207\dots$ but $\frac{\Sigma x}{n}$ stays the same so σ greater BUT M0A0 unless we see mention of 15 (cm) or 1.5 (m) being more than 1 sd</p>	

Question Number	Scheme	Marks
<p>2. (a)(i)</p> <p>(ii)</p> <p>(b)</p> <p>(c)</p> <p>(d)</p>	<p>[IQR = 47 – 33 =] 14</p> <p>[Range = 54 – 11 =] 43</p> <p>e.g. $Q_2 - Q_1 (=9) > (5=) Q_3 - Q_2$ Therefore <u>negative</u> (skew)</p> <p>25 → 37 ⇒ new $Q_1 = 35$ (may be on box plot) [54 → 60 (implies upper whisker now at 60) but no change to Q_3] New IQR = 12 so need to re-calculate for outliers Outliers now [> 47 + 18 = 65 or] < 35 – 18 = 17</p> <p>Box Plot</p>  <p>Box and two whiskers with median still at 42 Lower quartile at their 35 ($\neq 33$) and upper quartile unchanged at 47 Two outliers at 11 and 15 Lower whisker at 18 (or 17) <u>and</u> upper whisker at 60</p> <p>The value of pmcc is small <u>or</u> weak correlation (o.e.) Therefore the complaint is <u>not</u> supported</p>	<p>B1 B1 (2)</p> <p>M1 A1 (2)</p> <p>B1 M1 A1</p> <p>M1 A1ft A1 A1 (7)</p> <p>M1 A1 (2)</p> <p>[13]</p>
Notes		
<p>(a)(i)</p> <p>(b)</p> <p>(c)</p> <p>(d)</p>	<p>1st B1 for 14 2nd B1 for 43</p> <p>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”]</p> <p>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) <u>and</u> upper whisker ending at 60 Correct box plot scores all except 1st M1A1 (i.e. 5/7) this M1A1 requires some working</p> <p>M1 for comment that pmcc is “small” so little correlation (just saying < 0 is not enough) Allow e.g. “not significant” <u>or</u> “not relevant” <u>or</u> $-0.5 < r < 0.5$ <u>or</u> “not close to – 1” but “no correlation” is M0</p> <p>A1 for suggesting the complaint is <u>not</u> supported e.g. “little evidence to support claim” Dep on M1 seen NB M1A0 is possible</p>	

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)	$P(T) = pq + 0.02 \times 0.96 + (0.98 - p) \times 0.05 = 0.169$ $\{pq - 0.05p = 0.1008\}$ $P(\text{do not have disease} T) = \frac{(0.98 - p) \times 0.05}{0.169} = \frac{41}{169}$ So $p = \underline{0.16}$ e.g. $0.16q - 0.16 \times 0.05 = 0.1008$ $q = \underline{0.68}$	M1; A1 M1A1ft A1 dM1 A1 (7)
(c)(i)	$P(\text{type } A T \text{ and not type } B) = \frac{pq}{pq + (0.98 - p) \times 0.05} = \frac{0.1088}{0.1088 + 0.041}$ $= 0.7263\dots$ awrt <u>0.726</u>	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
<p>4. (a)</p> <p>(b)</p> <p>(c)</p> <p>(d)</p>	<p>[$W =$ weight of a package delivered to Susie $W \sim N(510, 45^2)$]</p> <p>$P(W < 450) = P\left(Z < \frac{450 - 510}{45}\right)$ or $P(Z < -1.3333)$</p> <p style="text-align: right;">$= 1 - 0.9082$ $= 0.0918$ [0.0912~0.0918]</p> <p>[$P(W > d) = 0.05$ implies] $\frac{d - 510}{45} = 1.6449$</p> <p style="text-align: right;">$d = 584.0205\dots$ awrt 584</p> <p>[$P(W > 450 \mid W < \text{"584.02..."}) =] \frac{P(450 < W < \text{"584.02..."})}{P(W < \text{"584.02..."})}$</p> <p style="text-align: center;">$= \frac{0.95 - \text{"0.0918"}}{0.95}$ or $\frac{\text{"0.9082"} - 0.05}{0.95}$</p> <p style="text-align: right;">$= 0.903368\dots$ awrt 0.904 or 0.903</p> <p>$\left(\frac{19}{20}\right)^4 \times \frac{1}{20} \times 5$</p> <p style="text-align: right;">$= 0.203626\dots$ awrt 0.204</p>	<p>M1</p> <p>M1</p> <p>A1</p> <p style="text-align: right;">(3)</p> <p>M1B1</p> <p>A1</p> <p style="text-align: right;">(3)</p> <p>M1</p> <p>M1A1</p> <p>A1</p> <p style="text-align: right;">(4)</p> <p>M1dM1</p> <p>A1</p> <p style="text-align: right;">(3)</p> <p style="text-align: right;">[13]</p>
Notes		
<p>Ans only</p>	<p style="text-align: center;">Correct answer only in (a), (c) or (d) scores all the marks for that part</p> <p>(a) 1st M1 for standardising 450 with 510 and 45 (allow \pm) 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...)</p> <p>(b) M1 for standardising their letter d with 510 and 45 and setting equal to z value $1 < z < 2$ B1 for use of $z = \pm 1.6449$ or better (calc 1.644853626...) A1 for awrt 584 (calc 584.0184...) [awrt 584.02 scores 3/3 584 scores M1B0A1]</p> <p>(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</p> <p>(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</p>	

Question Number	Scheme	Marks
<p>5. (a)</p> <p>(b)</p> <p>(c)</p> <p>(d)</p> <p>(e)</p>	$E(X) = -2p - p + 0 + \frac{1}{2} + 3p ; = \frac{1}{2}$ $E(X^2) = 4p + p + 0 + 1 + 9p = [14p + 1]$ $[\text{Var}(X) =] E(X^2) - [E(X)]^2 = 14p + 1 - \left(\frac{1}{2}\right)^2$ <p>So $14p + 0.75 = 2.5$</p> $p = \frac{1}{8}$ <p>Sum of probabilities = 1 implies $q = \frac{3}{8}$</p> <p>P(Amar wins) = e.g. $P(X_1 > 0) + P(X_1 < 0) \times P([X_1 + X_2] > 0 \{ X_1 < 0 \})$ <u>or</u> $P(X_1 = 2 \text{ or } 3) + P(X_1 = -2) \times P(X_2 = 3) + P(X_1 = -1) \times P(X_2 = 2 \text{ or } 3)$</p> <p><u>Cases</u> $X_1 = -2$ and $X_2 = 3$ so probability = p^2 $X_1 = -1$ and $X_2 \geq 2$ so probability = $p(p + \frac{1}{4})$</p> $\text{Total probability} = p + 0.25 + p^2 + p(p + 0.25) = \frac{1}{8} + \frac{1}{4} + \frac{1}{64} + \frac{1}{64} + \frac{1}{32}$ $= \frac{7}{16}$ <p>[Although $E(X) > 0$ since] $P(\text{win}) < 0.5$ Amar should not play the game <u>or</u> "disagree"</p>	<p>M1 ; A1</p> <p>(2)</p> <p>M1A1 dM1</p> <p>M1</p> <p>A1</p> <p>(5)</p> <p>B1ft</p> <p>(1)</p> <p>M1</p> <p>M1</p> <p>A1ft</p> <p>A1</p> <p>(4)</p> <p>M1</p> <p>A1</p> <p>(2)</p> <p>[14]</p>
Notes		
<p>(a)</p> <p>(b)</p> <p>(c)</p> <p>(d)</p> <p>ALT</p> <p>(e)</p>	<p>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)</p> <p>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 $[\text{Var}(X) =] E(X^2) - [E(X)]^2$ 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</p> <p>B1ft for $q = \frac{3}{8}$ or exact equivalent e.g. 0.375 <u>or</u> $\frac{3}{4} - 3p$ $0 < p < 1$</p> <p>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 <u>and</u> 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]</p> <p>Allow $P(\text{loses}) = q + p(1 - p) + p(0.75 - p)$ only if $1 - P(\text{loses})$ is seen</p> <p>M1 for identifying that the important feature is that $P(\text{win}) < 0.5$ (o.e.) [ft their $\frac{7}{16} < 0.5$] A1cao for concluding that he shouldn't play the game (dep on M1 seen & $0.375 < (d) < 0.5$)</p>	<p>If they think $E(X^2) = \text{Var}(X)$ 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)</p>

Question Number	Scheme	Marks
6. (a)	$\left[\sum y = 16 \times 20.5 = 328 \right] S_{yy} = 8266 - \frac{328^2}{16}$ $= 1542 \quad (\text{allow awrt } 1540)$ $[r =] \frac{-630.9}{\sqrt{368.16 \times 1542}}$ $= -0.837336... \quad \text{awrt } \underline{\underline{-0.837}}$	M1 A1 M1 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 <u>does</u> support belief	B1 (1)
(d)	$b = \frac{-630.9}{368.16} [= -1.7136...]$ $a = 20.5 - (-1.7136... \times 8.1) [= 34.3806...]$ $y = 34.38... - 1.7136...x \quad \underline{\underline{y = 34.4 - 1.71x}}$	M1 M1 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 $r = -0.83788...$ 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 in context (just “strong neg correlation” B0)
- (c) B1 for “points close to a straight line” and stating does support manager’s belief
or allow “ r is close to -1 ” or “strong (negative) correlation” and supports manager’s claim
or for a curve drawn on scatter diagram 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)