



Pearson
Edexcel

Mark Scheme (Results)

Summer 2021

Pearson Edexcel IAL In Statistics 3
Paper WST03/01

Question Number	Scheme	Notes	Marks																														
1(a)	<table border="1"> <tr> <td></td> <td><i>A</i></td> <td><i>B</i></td> <td><i>C</i></td> <td><i>D</i></td> <td><i>E</i></td> </tr> <tr> <td><i>C</i></td> <td>3</td> <td>2</td> <td>1</td> <td>5</td> <td>4</td> </tr> <tr> <td><i>M</i></td> <td>2</td> <td>4</td> <td>1</td> <td>5</td> <td>3</td> </tr> <tr> <td><i>d</i></td> <td>1</td> <td>-2</td> <td>0</td> <td>0</td> <td>1</td> </tr> <tr> <td><i>d</i>²</td> <td>1</td> <td>4</td> <td>0</td> <td>0</td> <td>1</td> </tr> </table>		<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>C</i>	3	2	1	5	4	<i>M</i>	2	4	1	5	3	<i>d</i>	1	-2	0	0	1	<i>d</i> ²	1	4	0	0	1	Attempt to rank at least 1 row with at least 3 correct. Allow reverse rankings.	M1
			<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>																										
<i>C</i>	3	2	1	5	4																												
<i>M</i>	2	4	1	5	3																												
<i>d</i>	1	-2	0	0	1																												
<i>d</i> ²	1	4	0	0	1																												
		Attempt at d^2 for their rankings , can be implied by $\sum d^2 = 6$	M1																														
	$\sum d^2 = 6$	Can be implied by correct answer. Must come from correct rankings.	A1																														
	$r_s = 1 - \frac{6(6)}{5(24)}$	(dep on 1 st M1) Use of correct formula with their $\sum d^2$	dM1																														
	$r_s = 0.7$	0.7 o.e. Must come from correct rankings.	A1																														
			(5)																														
1(b)	$H_0 : \rho = 0, H_1 : \rho > 0$	Both correct in terms of ρ or ρ_s . Must be compatible with their ranking.	B1																														
	cv 0.9 or cr $r_s \geq 0.9$	0.9 sign should match H_1 or their r_s	B1																														
	$r_s = 0.7$ does not lie in cr so do not reject H_0	Correct non-contextual statement e.g. “do not reject H_0 ”, “not in critical region”, “not significant”, “no positive correlation”. $ \text{test value} $ or $ \text{cv} > 1$ award M0	M1																														
	Data does not support plant biologist’s claim .	Correct conclusion in context. Must mention “biologist’s claim” o.e. or moisture and plant coverage . All previous marks in (b) must have been scored.	A1ft																														
SC	For use of two-tailed test: May score B0B1M1A0 for cv = 1(.000) and ‘not significant’ oe																																
			(4)																														
			Total 9																														

Question Number	Scheme	Notes	Marks																								
2	H_0 : Diet and health are independent (or not associated). H_1 : Diet and health are not independent (or associated).	“diet” and “health” mentioned at least once. Use of correlation is B0.	B1																								
	<table border="1"> <thead> <tr> <th>EXPECTED</th> <th>Good health</th> <th>Poor health</th> <th>Totals</th> </tr> </thead> <tbody> <tr> <td>Good diet</td> <td>83.19</td> <td>10.81</td> <td>94</td> </tr> <tr> <td>Poor diet</td> <td>93.81</td> <td>12.19</td> <td>106</td> </tr> <tr> <td>Totals</td> <td>177</td> <td>23</td> <td>200</td> </tr> </tbody> </table>	EXPECTED	Good health	Poor health	Totals	Good diet	83.19	10.81	94	Poor diet	93.81	12.19	106	Totals	177	23	200	Attempt $\frac{RT \times CT}{GT}$ with at least one correct to 1dp; all correct to 1dp.	M1; A1								
EXPECTED	Good health	Poor health	Totals																								
Good diet	83.19	10.81	94																								
Poor diet	93.81	12.19	106																								
Totals	177	23	200																								
	<table border="1"> <thead> <tr> <th>Observed</th> <th>Expected</th> <th>$\frac{(O-E)^2}{E}$</th> <th>$\frac{O^2}{E}$</th> </tr> </thead> <tbody> <tr> <td>86</td> <td>83.19</td> <td>0.095</td> <td>88.905</td> </tr> <tr> <td>8</td> <td>10.81</td> <td>0.730</td> <td>5.920</td> </tr> <tr> <td>91</td> <td>93.81</td> <td>0.084</td> <td>88.274</td> </tr> <tr> <td>15</td> <td>12.19</td> <td>0.648</td> <td>18.458</td> </tr> <tr> <td colspan="2">Totals</td> <td>1.557</td> <td>201.557</td> </tr> </tbody> </table>	Observed	Expected	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$	86	83.19	0.095	88.905	8	10.81	0.730	5.920	91	93.81	0.084	88.274	15	12.19	0.648	18.458	Totals		1.557	201.557	Attempt at $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$ with their values with 2 correct or 2 correct f.t. Allow 2sf for this mark.	M1
Observed	Expected	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$																								
86	83.19	0.095	88.905																								
8	10.81	0.730	5.920																								
91	93.81	0.084	88.274																								
15	12.19	0.648	18.458																								
Totals		1.557	201.557																								
	$\sum \frac{(O-E)^2}{E}$ or $\sum \frac{O^2}{E} - 200 = 1.557$	awrt 1.6	A1																								
	$\nu = (2-1)(2-1) = 1$ $\chi_1^2(5\%) = 3.841$	1 (may be implied) 3.841 NB: may see $\chi_3^2(5\%) = 7.815$ for f.t. NB: p -value 0.212 but scores B0B0 on its own	B1 B1f.t.																								
	$\chi^2 = 1.557$ does not lie in cr so insufficient evidence to reject H_0	For correct non-contextual statement linking their test statistic and their cv.	M1																								
	E.g. Diet and health are independent or There is no association between diet and health or The doctor’s belief is not supported by this data.	Dependent on a cv of 3.841 and 3 rd M1 Correct conclusion in context with “diet” and “health” or “doctor”. Condone “connection” or “relationship” but not “correlation”.	A1ft																								
			(9)																								
			Total 9																								

Question Number	Scheme	Notes	Marks
3.			
3(a)	$\bar{x} = \frac{1}{2}(11.52 + 13.75) = 12.635$	12.635 (may be implied by correct CI)	B1
		Use of 1.96	B1
	$\left(\frac{\sigma}{\sqrt{n}} = \frac{13.75 - 12.635}{1.96} (= 0.56887\dots)\right)$ $\left(\frac{\sigma}{\sqrt{n}} = \frac{13.75 - 11.52}{2 \times 1.96} (= 0.56887\dots)\right)$	For attempt at standard error (may be implied by awrt 0.569)	M1
		Use of 1.6449 or better (1.644853... from calc) Use of 1.64 or 1.65 is B0	B1
	$12.635 \pm 1.6449 \times 0.56887\dots$	For $(\text{their } \bar{x}) \pm (\text{their } 1.6449) \left(\text{their } \frac{\sigma}{\sqrt{n}} \right)$ their $\frac{\sigma}{\sqrt{n}}$ must be numerical	M1
	90% CI is (11.699..., 13.5707...)	awrt (11.7, 13.6) from correct working Correct answer with no working scores B1B1M1B0M1A1	A1
			(6)
3(b)	$4 \times 0.9^3 \times 0.1$ $= 0.2916$	$4p^3(1-p)$ (where $0 < p < 1$) awrt 0.292	M1 A1
			(2)
			Total 8

Question Number	Scheme	Notes	Marks
4. (a)	Label academic (1-1680) and vocational (1-2520)	For numbering/labelling/ordering (o.e.) students in each group	B1
	Use random numbers to select from each group.	For use of random sample/numbers/selection	B1
	28 academic and 42 vocational	Both numbers correct with the associated group	B1
			(3)
4(b)	$H_0 : \mu_v - \mu_a = 0$ $H_1 : \mu_v - \mu_a > 0$	If the hypotheses are given in terms of $\mu_a - \mu_b$, a and b must be defined.	B1
	$se = \sqrt{\frac{70}{80} + \frac{60}{50}}$	Correct attempt at se – condone slip in sample sizes.	M1
	$z = \frac{62 - 57}{\sqrt{\frac{70}{80} + \frac{60}{50}}}$	Dep on previous M1 standardising with (62 – 57) and their se (Allow \pm)	dM1
	$z = 3.471...$ (or probability of 0.0003)	awrt ± 3.47 (or awrt 0.0003)	A1
	cv $z = 1.6449$	Allow \pm but signs must be compatible Or allow comparison with probability of 0.05	B1
	Reject H_0 / significant	Dependent on 2 nd M1. A correct non-contextual statement based on their normal cv and their test statistic.	dM1
	There is evidence that the mean(o.e.) basic skills score for vocational students is greater than the mean basic skills score for academic students.	Correct comment in context. Must mention “mean”, “academic” and “vocational”. Allow f.t. on their normal cv and their test statistic.	A1f.t.
			(7)
4(c)	Mean / \bar{X}_a (basic skills) score for academic students and mean/ \bar{X}_v (basic skills) score for vocational students...	Must mention both means.	B1
	...have (approximately) a normal distribution (as sample sizes are large.)	Must mention normal.	B1
			(2)
4(d)	Samples are (large enough) so that $s^2 = \sigma^2$	Must imply for both samples	B1
			(1)
4(e)	Test no longer significant so insufficient evidence to reject H_0	Can be implied by correct comment in context.	M1
	Insufficient evidence that mean (basic skills) score for vocational students is greater than the mean (basic skills) score for academic students/There is no longer a difference in scores /Academic students have improved their mean (basic skills) score .	Must mention scores (o.e).	A1
			(2)
4(f)	The course was effective (o.e.)	Dep on a significant result in (b) and a non-significant result in (e)	B1
			(1)
			Total 16

Question Number	Scheme	Notes	Marks																														
5.(a)	Relief of symptoms is either a “success” or a “failure”. The probability the medicine being a success is constant. Samples from different medical practices are independent.	Any 2. Context required in one assumption.	B1 B1																														
			(2)																														
5(b)	Mean = $\frac{0 \times 4 + 1 \times 6 + 2 \times 3 + \dots + 8 \times 2}{50} = 3.54^*$	At least two correct terms on the numerator and 50 on the denominator, fully correct expression or $\frac{177}{50}$ dep on M1 scored cso.	M1, A1cso																														
			(2)																														
5(c)	$p = \frac{3.54}{8} = 0.4425$	Can be implied by at least 1 correct value for f or g .	B1																														
	$f = 50 \times C_4^8 \times 0.4425^4 \times 0.5575^4 = 12.96$ $g = 50 \times 0.4425^8 = 0.07$	Use of Bin(50, p) for M1, Allow awrt 12.96, awrt 0.07	M1A1A1																														
			(4)																														
5(d)	H_0 : Binomial distribution is a suitable model H_1 : Binomial distribution is not a suitable model	Both hypotheses correct If parameters used then B0.	B1																														
	<table border="1"> <thead> <tr> <th>No of successes</th> <th>O</th> <th>E</th> <th>$\frac{(O-E)^2}{E}$</th> <th>$\frac{O^2}{E}$</th> </tr> </thead> <tbody> <tr> <td>0,1,2</td> <td>13</td> <td>11.66</td> <td>0.154</td> <td>14.494</td> </tr> <tr> <td>3</td> <td>12</td> <td>13.07</td> <td>0.088</td> <td>11.018</td> </tr> <tr> <td>4</td> <td>10</td> <td>12.96</td> <td>0.676/7</td> <td>7.716</td> </tr> <tr> <td>5,6,7,8</td> <td>15</td> <td>12.31</td> <td>0.588/7</td> <td>18.278</td> </tr> <tr> <td></td> <td>50</td> <td>50</td> <td>1.506</td> <td>51.506</td> </tr> </tbody> </table>	No of successes	O	E	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$	0,1,2	13	11.66	0.154	14.494	3	12	13.07	0.088	11.018	4	10	12.96	0.676/7	7.716	5,6,7,8	15	12.31	0.588/7	18.278		50	50	1.506	51.506	Combining 0,1,2 or 5,6,7,8. M1 for attempting $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$ with at least 2 correct expressions or 2 correct values to 2sf.	M1 M1
No of successes	O	E	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$																													
0,1,2	13	11.66	0.154	14.494																													
3	12	13.07	0.088	11.018																													
4	10	12.96	0.676/7	7.716																													
5,6,7,8	15	12.31	0.588/7	18.278																													
	50	50	1.506	51.506																													
	$\sum \frac{(O-E)^2}{E} = \sum \frac{O^2}{E} - 50 = 1.50\dots$	awrt 1.5 (calculator: 1.50498...)	A1																														
	$\nu = 4 - 2 = 2, \chi^2(10\%) = 4.605$	2 can be implied by 4.605 seen Only f.t. $\nu = r - 2$	B1 B1f.t.																														
	Insufficient evidence to reject H_0	For correct non-contextual statement linking their test statistic and their cv.	M1																														
	Data is consistent with a binomial distribution (oe)	A correct comment suggesting that binomial model is suitable / good fit. Hypotheses wrong way around scores A0 here. Condone parameters here.	A1																														
			(8)																														
			Total 16																														

Question Number	Scheme	Notes	Marks
6(a)	$W = B - 1.1R$	May be implied by correct mean or variance	M1
	$W \sim N(55 - 1.1 \times 51, 1.3^2 + 1.1^2 \times 1.2^2)$ or $W \sim N(-1.1, 3.4324)$	(\pm)1.1, awrt 3.43 (may be seen in standardisation)	A1, A1
	$P(W < 0) = P\left(Z < \frac{0 + 1.1}{\sqrt{3.4324}}\right)$	Standardising with their mean and their sd. leading to a probability > 0.5	M1
	$= P(Z < 0.5937\dots)$		
	$= 0.7224$ or 0.7237	awrt 0.72	A1
			(5)
6(b)	$X = B_1 - B_2$	May be implied by correct mean or variance	M1
	$X \sim N(55 - 55, 2 \times 1.3^2)$ or $X \sim N(0, 3.38)$	0, 3.38	A1, A1
	$P\left(Z > \frac{1 - 0}{\sqrt{3.38}}\right)$ or $P\left(Z < \frac{-1 - 0}{\sqrt{3.38}}\right)$	dep on 1 st M1 for standardising with their mean and their sd.	dM1
	$P(X > 1) = 2 \times P(X > 1)$	For 2 \times seen or implied	M1
	$= 2 \times P(Z > 0.5439\dots) = 2 \times (1 - 0.7054)$	2 \times 0.2946 or 2 \times 0.2932 (calc)	
	$= 0.5892$	awrt 0.59	A1
			(6)
6(c)	$V = B_1 + B_2 + B_3 + B_4 + B_5 + B_6 + B_7 + B_8 + B_9 + B_{10} + S$ $Y = R_1 + R_2 + R_3 + R_4 + R_5 + R_6 + R_7 + R_8 + R_9 + R_{10} + R_{11} + S$	May be implied by either correct distribution	M1
	$V \sim N(553, 16.94)$ and $Y \sim N(564, 15.88)$	Both correct	A1
	$D = Y - V$ so $D \sim N(11 \times 51 - 10 \times 55, 11 \times 1.2^2 + 10 \times 1.3^2 + 2 \times 0.2^2)$ or $D \sim N(11, 32.82)$	Attempt at their difference for the mean, and their sum for the variance.	M1
		(\pm)11 and awrt 32.8	A1
	$P(D > 0) = P\left(Z > \frac{0 - 11}{\sqrt{32.82}}\right)$	dep on 1 st M1 for standardising using their mean and their standard deviation leading to a probability > 0.5	dM1
	$= P(Z > -1.920\dots)$		
	$= 0.9726$	awrt 0.973	A1
			(6)
			Total 17