



# Mark Scheme (Results)

October 2021

Pearson Edexcel International A Level  
In Statistics S3 (WST03) Paper 01

Question Number	Scheme		Marks
1.	$H_0: \mu = 30 \quad H_1: \mu < 30$		B1
	$z = \frac{29.5 - 30}{\frac{2.5}{\sqrt{80}}}$		M1
	$z = -1.7888\dots$		awrt-1.79 A1
	$-1.7888 < -1.6449$		B1
	Reject $H_0$ <u>or</u> significant result <u>or</u> in the critical region		
	There is evidence to support the <u>manager's</u> claim.		A1
			(5)
<b>Notes</b>			<b>Total 5</b>
	<b>B1</b>	Both hypotheses correct in terms of $\mu$	
	<b>M1</b>	for attempting test statistic, allow $\pm$ , Condone $\sqrt{\frac{2.5}{80}}$	
	<b>A1</b>	awrt -1.79 allow $ z  = 1.7888\dots$ Allow $p$ value of 0.0367 or awrt 0.0368 or $CR \leq 29.54$	
	<b>B1</b>	$ CV  = 1.6449$ or better (Ignore any comparisons) Allow $CR \leq 29.54$ SC If $p$ value of 0.0367 or awrt 0.0368 award B1 if 2 <sup>nd</sup> A1 is awarded	
	<b>A1</b>	For correct conclusion. Allow the manager's claim in words if it includes screws and less (oe)	

Question Number	Scheme						Marks	
2	H <sub>0</sub> : Potassium has no effect on the quality of apple H <sub>1</sub> : Potassium has an effect on the quality of apple						B1	
	Grade	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	M1A1	
	Expected values	9.6	67.2	124.8	24.0	14.4		
	$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(9-"9.6")^2}{"9.6"} + \dots + \frac{(3-"14.4")^2}{"14.4"} \text{ or}$ $\chi^2 = \sum \frac{O^2}{E} - N = \frac{9^2}{"9.6"} + \dots + \frac{3^2}{"14.4"} - 240$						M1	
	= 10.657...						awrt 10.7	A1
	Degrees of freedom = 4 $\chi^2_{4,0.05} = 9.488$							B1 B1ft
	[Reject H <sub>0</sub> ] Data suggests that potassium may affect the distribution of the grades of apples <b>or</b> there is evidence that Andy's belief is incorrect							A1ft
							(8)	
<b>Notes</b>						<b>Total 8</b>		
<b>B1</b>	Both hypotheses in context. May use other wording eg The grading of apples remains the same.							
<b>M1</b>	A correct method to calculate expected values eg $0.04 \times 240$							
<b>A1</b>	At least 3 expected values correct							
<b>M1</b>	A correct method using their expected values to calculate $\chi^2$ At least one correct, ft their expected values with an intention to add							
<b>A1</b>	awrt 10.7							
<b>B1</b>	Degrees of freedom = 4 (may be implied by 9.488)							
<b>B1ft</b>	9.488 ft their DoF. If no DoF stated then this must be correct for their working.							
<b>A1ft</b>	ft their $\chi^2$ value provided the 2 <sup>nd</sup> M1 is awarded and CV. If no hypotheses or hypotheses wrong way round do not award. Must include the word 'apples' or 'belief' oe							

Question Number	Scheme										Marks	
3(a)	jam	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	M1	
	Price	1	2	4	5	3	6	7	8	9		
	Taste	1	2	8	9	4	3	6	5	7		
	$\sum d^2 = [0+0+]16+16+1+9+1+9+4 [= 56]$										M1A1	
$r_s = 1 - \frac{6(56)}{9(80)}; = \frac{8}{15} = 0.5333..$										awrt 0.533	dM1A1	
											(5)	
(b)	$H_0 : \rho = 0, H_1 : \rho \neq 0$										B1	
	Critical Value = 0.7										B1	
	There is no evidence of a relationship between <u>price</u> and <u>taste</u> of strawberry jam										B1ft	
												(3)
(c)	$r = \frac{16.4943}{\sqrt{2.0455 \times 243.5556}}$										M1	
	= 0.7389...										awrt 0.739	A1
												(2)
(d)	$H_0 : \rho = 0, H_1 : \rho > 0$										B1	
	CV = 0.5822										B1	
	There is evidence of a <u>positive correlation</u> between <u>price</u> and <u>taste</u> of strawberry jam										B1ft	
												(3)
(e)	Spearman's rank										B1	
	as it is unlikely that a joint normal distribution applies.											
	or the marks are a judgement or the marks are not a meaningful scale.										(1)	
<b>Notes</b>											<b>Total 14</b>	
(a)	<b>M1</b>	Attempt to rank each jar for taste and price. At least 4 pairs of ranks correct										
	<b>M1</b>	For an attempt at $d^2$ row for their ranks (may be implied by $\sum d^2 = 56$ )										
	<b>A1</b>	$\sum d^2 = 56$										
	<b>dM1</b>	Dependent on the previous M being awarded. Using $1 - \frac{6\sum d^2}{9(80)}$										
(b)	<b>A1</b>	$\frac{8}{15}$ or awrt 0.533										
	<b>B1</b>	Both hypotheses stated in terms of $\rho$ . Must be two- tail.										
	<b>B1</b>	0.7 for CV. Allow 0.6 if a one tail test is used										
	<b>B1ft</b>	For a correct contextualised comment which has price and taste Follow through their $r_s$ with their 0.7 (provided $ \text{their } r_s  < 1$ )										
(c)	<b>M1</b>	Correct method used										
	<b>A1</b>	awrt 0.739										
	(d)	<b>B1</b>	Both hypotheses stated in terms of $\rho$ . Must be one-tail. If B0 awarded in part (b) then allow any letter instead of $\rho$ that is consistent with part (b)									
		<b>B1</b>	0.5822 Allow 0.6664 if a two-tail test is used.									
<b>B1ft</b>		Correct conclusion in context which has positive correlation (this may be implied by a correct description of positive correlation), price and taste. Follow through their 0.5822 and 0.739										
(e)	<b>B1</b>	Selecting Spearman's with a suitable reason. Do not allow 'because it is ranked' as a suitable reason										

Question Number	Scheme		Marks									
4(a)	Label the houses in area A 1- 41, area B 1 – 164, area C 1 – 123 and area D 1 - 82		M1									
	<u>Use random numbers</u> to select a ...		M1									
	Simple random sample of <u>20</u> area <u>A</u> , <u>80</u> area <u>B</u> , <u>60</u> area <u>C</u> and <u>40</u> area <u>D</u>		A1									
			(3)									
(b)	$\frac{357 \times 260}{595}$ or $\frac{238 \times 260}{595}$		M1									
	156 and 104		A1									
			(2)									
(c)	<table border="1"> <thead> <tr> <th>Observed</th> <th>Expected</th> <th><math>\frac{(O-E)^2}{E}</math></th> </tr> </thead> <tbody> <tr> <td>162</td> <td>"156"</td> <td><math>\frac{(162 - "156")^2}{"156"} = \frac{3}{13} = 0.2307\dots</math></td> </tr> <tr> <td>98</td> <td>"104"</td> <td><math>\frac{(98 - "104")^2}{"104"} = \frac{9}{26} = 0.3461\dots</math></td> </tr> </tbody> </table>		Observed	Expected	$\frac{(O-E)^2}{E}$	162	"156"	$\frac{(162 - "156")^2}{"156"} = \frac{3}{13} = 0.2307\dots$	98	"104"	$\frac{(98 - "104")^2}{"104"} = \frac{9}{26} = 0.3461\dots$	M1
	Observed	Expected	$\frac{(O-E)^2}{E}$									
	162	"156"	$\frac{(162 - "156")^2}{"156"} = \frac{3}{13} = 0.2307\dots$									
	98	"104"	$\frac{(98 - "104")^2}{"104"} = \frac{9}{26} = 0.3461\dots$									
	$\chi^2 = 4.657 + "0.2307\dots" + "0.346\dots"$		M1									
	$= 5.234\dots$		awrt 5.23									
	$\nu = (2 - 1)(3 - 1) = 2$		B1									
$\chi^2(0.05) = 5.991 \Rightarrow \text{CR: } \chi^2 > 5.991$		B1ft										
There is no evidence to suggest that there is an association between <u>age</u> and <u>listening</u> to <i>LSB</i>		dA1										
		(6)										
<b>Notes</b>			<b>Total 11</b>									
(a)	<b>M1</b> <b>M1</b> <b>A1</b>	For suitable labelling of all four areas. E.g. for area A: 1 – 41 or 0 - 40 For use of random numbers to select houses in each area. For 20 A, 80B, 60C and 40 D (dependent on 2 <sup>nd</sup> M1 only) NB A simple random sample of 20 A, 80B, 60C and 40 D scores M0M1A1. Allow M1 : allocate random numbers to each house M1 : arrange the numbers in order A1 : select the 1 <sup>st</sup> 20 for area <u>A</u> , <u>80</u> for area <u>B</u> , <u>60</u> for area <u>C</u> and <u>40</u> for area <u>D</u> SC If M0M0 scored then award B1 for <u>20</u> area <u>A</u> , <u>80</u> area <u>B</u> , <u>60</u> area <u>C</u> and <u>40</u> area <u>D</u>										
(b)	<b>M1</b> <b>A1</b>	A correct method for finding one expected value. Correct answer for both values										
(c)	<b>M1</b>	A correct method for finding both contributions to the $\chi^2$ value										
	<b>M1</b>	Adding the two values to 4.657 (may be implied by a full $\chi^2$ calculation, do not ISW)										
	<b>A1</b>	awrt 5.23										
	<b>B1</b>	2										
	<b>B1ft</b>	5.991 or better ft their DoF										
	<b>dA1</b>	A correct contextual conclusion, which has the words age and listening dependent on both M marks being awarded. <b>NB</b> if they give a <i>p</i> value of 0.0730... rather than the CV they can get M1M1B1B0A1										

Question Number	Scheme		Marks
5(a)	$2.977 \pm 2.5758 \times \frac{0.015}{3}$		M1,B1
	$= (2.9641\dots, 2.9898\dots)$	awrt (2.964, 2.990)	A1
			<b>(3)</b>
(b)	The CI does not contain the stated weight.		B1
			<b>(1)</b>
(c)	$2.995 - 1.96 \times \frac{0.015}{\sqrt{n}} < 2.991$		M1
	$\sqrt{n} < \frac{1.96 \times 0.015}{2.995 - 2.991}$		M1d
	$\sqrt{n} < \text{awrt } 7.35$		A1
	$n = 54$		A1cao
			<b>(4)</b>
<b>Notes</b>			<b>Total 8</b>
(a)	<b>M1</b>	$2.977 \pm (z \text{ value}) \times \frac{0.015}{3}$	
	<b>B1</b>	awrt 2.5758	
	<b>A1</b>	awrt (2.964, 2.990 (condone 2.99))	
(b)	<b>B1</b>	cao this must be consistent with their confidence interval	
(c)	<b>M1</b>	Setting up an inequality using z value > 1.5 Condone =	
	<b>M1d</b>	Dep on previous M mark. Correct rearranging to get $\sqrt{n} < \dots$ or $n < \dots$ Condone = or >	
	<b>A1</b>	awrt 7.35 may be implied by awrt 54	
	<b>A1cao</b>	54	

Question Number	Scheme	Marks
6(a)	$\bar{h} = 65.4$	B1
	$s^2 = \frac{214676 - 50 \times ("65.4")^2}{49}$	M1
	$= 16.693\dots$	awrt 16.7 A1
		<b>(3)</b>
(b)	$H_0: \mu_{do} = \mu_{do\ not}$ $H_1: \mu_{do} < \mu_{do\ not}$	B1
	$z = \pm \frac{"65.4" - 70.8}{\sqrt{\frac{"16.693\dots"}{50} + \frac{29.6}{40}}}$	M1M1
	$= \pm 5.21\dots$	awrt 5.21 A1
	CV 1.6449	B1
	Amala's <u>belief</u> is supported	A1 ft <b>(6)</b>
(c)	CLT enables you to assume that (the sampling distribution of the sample mean of ) resting heart rate is normally distributed for <u>both</u> groups	B1
		<b>(1)</b>
(d)	Each population/sample is independent <b>or</b> each male is independent of the other males.	B1
	Assume the $\sigma_{do}^2 = s_{do}^2$ and $\sigma_{do\ not}^2 = s_{do\ not}^2$	B1
		<b>(2)</b>
	<b>Notes</b>	<b>Total 12</b>
(a)	<b>B1</b> 65.4 only <b>M1</b> Correct method to find $s^2$ using their $\bar{h}$ <b>A1</b> awrt 16.7	
(b)	<b>B1</b> Both hypotheses correct - must be clear which is exercise and which is not	
	<b>M1</b> For the denominator. Ft their 16.693...	
	<b>M1</b> Correct ft their 65.4 and 16.693...	
	<b>A1</b> awrt 5.21 allow $ z  = 5.21\dots$ <b>B1</b> $ CV  = 1.6449$ or better ft their $z$ value and CV if the hypotheses are the correct way round. Correct conclusion in context need belief. May be in words with heart and exercise e.g. resting heart rate is lower in men who exercise regularly	
(c)	<b>B1</b> For the idea both groups normally distributed	
(d)	<b>B1</b> For identifying the need for the groups <b>or</b> males to be independent.	
	<b>B1</b> Realising the $\sigma^2 = s^2$ Allow sample sizes big enough for CLT to hold	

Question Number	Scheme	Marks
7(a)	$E(B_1 - B_2) = 0$	B1
	$\text{Var}(B_1 - B_2) = 0.006$	B1
	$P( B_1 - B_2  > 0.1) = 2P(B_1 - B_2 > 0.1)$	M1
	$= 2 \times P\left(Z > \frac{0.1}{\sqrt{0.006}}\right) [= 2 \times P(Z > 1.2909...)]$	M1
	$= 0.1967...$ awrt 0.197	A1
		<b>(5)</b>
(b)	$\bar{B} \sim N\left(1.96, \frac{0.003}{n}\right)$	B1
	$P(\bar{B} > 2) = P\left(Z > \frac{2-1.96}{\sqrt{0.003/n}}\right) [< 0.01]$	M1
	$\frac{2-1.96}{\sqrt{\frac{0.003}{n}}} > 2.3263$	B1 dM1
	$n = 11$	A1
		<b>(5)</b>
(c)	$\mu_M = 21.8 + 500 \times 1.96 [= 1001.8]$ ; $\sigma_M^2 = 0.6 + 500 \times 0.003 [= 2.1]$	M1 ; M1
	Let $X = 4T - 3M$	M1
	$\mu_X = 4 \times 774 - 3 \times 1001.8 [= 90.6]$ ; $\sigma_X^2 = 16 \times 1.8 + 9 \times 2.1 [= 47.7]$	M1 ; M1
	$P(4T - 3M > 100) = P\left(Z > \frac{100 - 90.6}{\sqrt{47.7}}\right) [= P(Z > 1.361...)]$	M1
	$= 0.0869$ (table) or $0.08675...$ (calc)	A1
		<b>(7)</b>
		<b>Total 17</b>
(a)	<b>B1</b>	For expected value being 0 written or used
	<b>B1</b>	For 0.006 being written or used for Variance
(b)	<b>M1</b>	Realising they need to consider both
	<b>M1</b>	Correct standardisation using their 0.1 and 0.006 If the expected value and/or standard deviation not stated then they must be correct
	<b>A1</b>	awrt 0.197
(c)	<b>B1</b>	The correct distribution written or used
	<b>M1</b>	Correct standardisation. Allow using their distribution if stated but must contain $\sqrt{n}$ for sd
	<b>B1</b>	Using awrt 2.3263
(d)	<b>dM1</b>	Dep on previous M being awarded using a z value, $2 < z < 3$
	<b>A1</b>	11
	<b>M1</b>	Correct method for finding the mean of $M$
	<b>M1</b>	Correct method for finding the var of $M$
	<b>M1</b>	Realising the need to find $4T - 3M$ or $4T - 3M - 100$ or $100 + 3M - 4T$
	<b>M1</b>	Correct method for finding the mean of $X$ (using $4T - 3M - 100 = -9.4$ or $100 + 3M - 4T = 9.4$ )
	<b>M1</b>	Correct method for finding the var of $X$
<b>M1</b>	Correct standardisation using their mean of $X$ and their standard deviation of $X$ If these are not stated then they must be correct	
	<b>A1</b>	awrt 0.0869 or 0.0868