



# Mark Scheme (Results)

January 2024

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

Question Number	Scheme	Marks																																												
1.	<p><math>H_0</math> : There is no association between treatment and presence of fungus  <math>H_1</math> : There is association between treatment and presence of fungus</p> <table border="1" data-bbox="288 344 1225 562"> <thead> <tr> <th>Expected</th> <th>No treatment</th> <th>Sulphur</th> <th>Copper sulphate</th> </tr> </thead> <tbody> <tr> <td>No Fungus</td> <td><math>\frac{123 \times 30}{150}</math> [=24.6]</td> <td><math>\frac{123 \times 63}{150}</math> [=51.66]</td> <td><math>\frac{123 \times 57}{150}</math> [=46.74]</td> </tr> <tr> <td>Fungus</td> <td><math>\frac{27 \times 30}{150}</math> [=5.4]</td> <td><math>\frac{27 \times 63}{150}</math> [=11.34]</td> <td><math>\frac{27 \times 57}{150}</math> [=10.26]</td> </tr> </tbody> </table> <table border="1" data-bbox="288 595 1305 958"> <thead> <tr> <th>Observed</th> <th>Expected</th> <th><math>\frac{(O-E)^2}{E}</math></th> <th><math>\frac{O^2}{E}</math></th> </tr> </thead> <tbody> <tr> <td>20</td> <td>24.6</td> <td>0.86016...</td> <td>16.2601...</td> </tr> <tr> <td>55</td> <td>51.66</td> <td>0.21594...</td> <td>58.5559...</td> </tr> <tr> <td>48</td> <td>46.74</td> <td>0.03396...</td> <td>49.2939...</td> </tr> <tr> <td>10</td> <td>5.4</td> <td>3.91851...</td> <td>18.5185...</td> </tr> <tr> <td>8</td> <td>11.34</td> <td>0.98373...</td> <td>5.6437...</td> </tr> <tr> <td>9</td> <td>10.26</td> <td>0.15473...</td> <td>7.8947...</td> </tr> <tr> <td colspan="2">Totals:</td> <td>6.167...</td> <td>156.167...</td> </tr> </tbody> </table> $X^2 = \sum \frac{(O-E)^2}{E} \quad \text{or} \quad \sum \frac{O^2}{E} - 150$ <p>= awrt 6.17  <math>\nu = (3-1)(2-1) = 2</math>  <math>\chi^2_2(0.05) = 5.991</math>  [Reject <math>H_0</math>/significant/in the CR] There is sufficient evidence to suggest there is an association between <u>treatment</u> and presence of <u>fungus</u>.</p>	Expected	No treatment	Sulphur	Copper sulphate	No Fungus	$\frac{123 \times 30}{150}$ [=24.6]	$\frac{123 \times 63}{150}$ [=51.66]	$\frac{123 \times 57}{150}$ [=46.74]	Fungus	$\frac{27 \times 30}{150}$ [=5.4]	$\frac{27 \times 63}{150}$ [=11.34]	$\frac{27 \times 57}{150}$ [=10.26]	Observed	Expected	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$	20	24.6	0.86016...	16.2601...	55	51.66	0.21594...	58.5559...	48	46.74	0.03396...	49.2939...	10	5.4	3.91851...	18.5185...	8	11.34	0.98373...	5.6437...	9	10.26	0.15473...	7.8947...	Totals:		6.167...	156.167...	<p>B1 M1 dM1 dM1 A1 B1 B1ft A1ft</p> <p>[8]</p>
Expected	No treatment	Sulphur	Copper sulphate																																											
No Fungus	$\frac{123 \times 30}{150}$ [=24.6]	$\frac{123 \times 63}{150}$ [=51.66]	$\frac{123 \times 57}{150}$ [=46.74]																																											
Fungus	$\frac{27 \times 30}{150}$ [=5.4]	$\frac{27 \times 63}{150}$ [=11.34]	$\frac{27 \times 57}{150}$ [=10.26]																																											
Observed	Expected	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$																																											
20	24.6	0.86016...	16.2601...																																											
55	51.66	0.21594...	58.5559...																																											
48	46.74	0.03396...	49.2939...																																											
10	5.4	3.91851...	18.5185...																																											
8	11.34	0.98373...	5.6437...																																											
9	10.26	0.15473...	7.8947...																																											
Totals:		6.167...	156.167...																																											
<b>Notes</b>																																														
<p>1<sup>st</sup> B1 both hypotheses correct with treatment (oe) and fungus (oe) (treatment and fungus need to only appear in either <math>H_0</math> or <math>H_1</math>). May be written in terms of independence.  1<sup>st</sup> M1 attempt at <math>\frac{\text{row total} \times \text{column total}}{\text{total}}</math> (can be implied by at least one correct <math>E_i</math> to 1dp)  2<sup>nd</sup> M1 (dep on 1<sup>st</sup> M1) at least 2 correct terms for <math>\frac{(O-E)^2}{E}</math> or <math>\frac{O^2}{E}</math> or correct expressions with their <math>E_i</math> (allow 2sf accuracy) (May be implied by awrt 6.17 or awrt 156.17)  3<sup>rd</sup> M1 (dep on 2<sup>nd</sup> M1) for using <math>\sum \frac{(O-E)^2}{E}</math> or <math>\sum \frac{O^2}{E} - 150</math> (May be implied by awrt 6.17)  1<sup>st</sup> A1 awrt 6.17  2<sup>nd</sup> B1 DoF/<math>\nu = 2</math> (May be implied by 5.991)  3<sup>rd</sup> B1ft 5.991 (or better) allow ft from their stated degrees of freedom)  2<sup>nd</sup> A1ft (dep on 3<sup>rd</sup> M1 and 3<sup>rd</sup> B1) for a correct ft contextualised conclusion. Must include “treatment” and “fungus”. Ignore any non-contextual statements. If hypotheses are the wrong way round then A0.</p>																																														

Question Number	Scheme	Marks
<p>2. (a)</p> <p>(b)</p> <p>(c)</p>	<p><b>Number</b> all employees [1-800] Use a <b>random</b> number to select the <b>first employee</b> oe Then select every <b>10<sup>th</sup></b> employee from the list of employees</p> <p><b>Number</b> all employees <b>by city/for each city</b> Use <b>random</b> numbers to select <b>54</b> employees from London, <b>31</b> employees from Edinburgh and <b>15</b> employees from Cardiff</p> <p>e.g Stratified sample reflects the population structure</p>	<p>B1 B1 B1 (3)</p> <p>B1 B1 B1 (3)</p> <p>B1 (1) [7]</p>
<b>Notes</b>		
<p>(a)</p> <p>(b)</p> <p>(c)</p>	<p>1<sup>st</sup> B1 idea of <b>numbering</b> all employees 2<sup>nd</sup> B1 idea of <b>randomly</b> selecting a starting point 3<sup>rd</sup> B1 selecting every <b>k<sup>th</sup></b> employee</p> <p>1<sup>st</sup> B1 idea of <b>numbering</b> employees for <b>each city</b> 2<sup>nd</sup> B1 use of <b>random</b> numbers (oe) 3<sup>rd</sup> B1 <b>54</b> from London, <b>31</b> from Edinburgh, <b>15</b> from Cardiff cao</p> <p>Any correct advantage e.g. Allows calculations [of statistics] for each city/group</p>	

Question Number	Scheme	Marks
<p><b>3. (a)</b></p> <p><b>(b)</b></p> <p><b>(c)</b></p>	<p><math>H_0 : \rho = 0</math>                      <math>H_1 : \rho \neq 0</math></p> <p><math>[r = ] \frac{83.634}{\sqrt{2.486 \times 3026.234}} = 0.9642.....</math>                      awrt 0.964</p> <p>CV = 0.7545</p> <p>[Reject <math>H_0</math>/Significant] There is evidence of correlation between annual <u>tea consumption</u> and <u>population</u>.</p> <p><b>Country</b>    A    B    C    D    E    F    G</p> <p><b>T Rank</b>    5    6    4    7    1    2    3</p> <p><b>P Rank</b>    7    6    4    3    1    2    5</p> <p>or</p> <p><b>Country</b>    A    B    C    D    E    F    G</p> <p><b>T Rank</b>    3    2    4    1    7    6    5</p> <p><b>P Rank</b>    1    2    4    5    7    6    3</p> <p><math>\sum d^2 = 4+0+0+16+0+4 [= 24]</math></p> <p><math>[r_s = ] 1 - \frac{6(24)}{7(48)} = 0.571428...</math>                      awrt 0.571</p> <p><math>H_0 : \rho_s = 0</math>                      <math>H_1 : \rho_s &gt; 0</math></p> <p>CV = 0.7143</p> <p>[Do not reject <math>H_0</math> / not significant] There is not enough evidence to suggest a <u>positive</u> correlation between annual <u>tea consumption</u> and <u>population</u>.</p>	<p>B1</p> <p>M1 A1</p> <p>B1ft</p> <p>A1</p> <p>(5)</p> <p>M1</p> <p>M1</p> <p>dM1A1</p> <p>(4)</p> <p>B1</p> <p>M1</p> <p>A1ft</p> <p>(3)</p> <p>[12]</p>
<b>Notes</b>		
<p><b>(a)</b></p> <p><b>(b)</b></p> <p><b>(c)</b></p>	<p>1<sup>st</sup> B1 both hypotheses correct in terms of <math>\rho</math> (must be two-tailed). Condone use of <math>p</math></p> <p>M1 use of formula for <math>r</math> (May be implied by awrt 0.964)</p> <p>1<sup>st</sup> A1 awrt 0.964</p> <p>2<sup>nd</sup> B1ft 0.7545 (or better) or ft 1-tailed alternative hypothesis (0.6694)</p> <p>2<sup>nd</sup> A1 correct contextual conclusion including tea consumption/<math>t</math> and population/<math>p</math>. Must be consistent with their <math>r</math> and their CV. (Ignore any non-contextual conclusion)</p> <p>Allow positive correlation</p> <p>1<sup>st</sup> M1 attempt to rank each country for tea and population (at least 4 correct in each)</p> <p>2<sup>nd</sup> M1 for <math>\sum d^2</math> for their ranks (implied by <math>\sum d^2 = 24</math>)</p> <p>2<sup>nd</sup> M1 (dep on 1<sup>st</sup> M1) use of <math>1 - \frac{6(24)}{7(48)}</math></p> <p>A1 awrt 0.571 (or <math>\frac{4}{7}</math>)</p> <p>B1 both hypotheses correct in terms of <math>\rho</math> or <math>\rho_s</math>. Condone use of <math>p</math></p> <p>M1 0.7143 (or better)</p> <p>A1ft correct contextual conclusion including positive, tea consumption/<math>t</math> and population/<math>p</math>. (Ignore any non-contextual conclusion) ft their part (b)</p>	

Question Number	Scheme	Marks
<p><b>4. (a)</b></p> <p><b>(b)</b></p> <p>or</p> <p><b>(c)</b></p> <p><b>(d)</b></p>	$\frac{[0 \times 24] + 1 \times 34 + 2 \times 28 + 3 \times 21 + 4 \times 8 + 5 \times 5}{120} [= 1.75]^*$ $[s = ] 120 \times \frac{e^{-1.75} 1.75^4}{4!} [= 8.15]^*$ $[s = ] 120 - \left( 20.85 + 36.49 + 31.93 + 120 \times \frac{e^{-1.75} 1.75^3}{3!} + 3.95 \right) [= 8.15]^*$ <p><math>[r = ] 18.63</math></p> <p><math>H_0</math> : Poisson distribution is a good fit.  <math>H_1</math> : Poisson distribution is not a good fit</p> $\sum \frac{(O_i - E_i)^2}{E_i} = 1.43 + \frac{(8 + 5 - (8.15 + 3.95))^2}{8.15 + 3.95}$ <p style="text-align: right;"><math>= 1.49694\dots</math> awrt 1.5(0)</p> <p><math>\nu = 5 - 1 - 1 = 3</math>  <math>\chi^2_3(0.05) = 7.815</math></p> <p>[Do not reject <math>H_0</math>/not significant]  There is insufficient evidence to reject the office manager's belief or the number of jobs sent to the printer are consistent with a Poisson distribution.</p>	<p>B1*cso (1)</p> <p>B1*cso (1)</p> <p>B1 (1)</p> <p>B1</p> <p>M1 M1</p> <p>A1 B1 B1ft</p> <p>A1 (7)</p> <p><b>[10]</b></p>
<b>Notes</b>		
<p><b>(a)</b></p> <p><b>(b)</b></p> <p><b>(c)</b></p> <p><b>(d)</b></p>	<p>B1cso correct calculation, minimum working <math>\frac{34 + 56 + 63 + 32 + 25}{120} = 1.75^*</math></p> <p>B1cso fully correct calculation (may be seen in stages) leading to 8.15*</p> <p>For 18.63 (This may be seen in part (b) if labelled as <math>r</math>)</p> <p>1<sup>st</sup> B1 both hypotheses correct (mention of 1.75 is B0)  1<sup>st</sup> M1 evidence of combining last 2 cells e.g. 8 + 5 and 8.15 + 3.95</p> <p>2<sup>nd</sup> M1 use of <math>1.43 + \sum \frac{(O_i - E_i)^2}{E_i}</math> for remaining cells (Condone cells not combined. May be implied by <math>1.43 + 0.00276\dots + 0.279\dots</math> or awrt 1.71)</p> <p>1<sup>st</sup> A1 awrt 1.50 (allow 1.5 from correct working)</p> <p>2<sup>nd</sup> B1 Dof/ <math>\nu = 3</math> implied by a correct critical value of 7.815</p> <p>3<sup>rd</sup> B1ft 7.815 (allow ft on the <math>\nu</math> so may see 9.488 or 11.070 etc)</p> <p>2<sup>nd</sup> A1 (dep on 2<sup>nd</sup> M1) a correct conclusion which states that the office manager's belief is correct/the data are consistent with a Poisson distribution which must be consistent with the test statistic and CV. Condone Po(1.75) is a suitable model. This mark is independent of the hypotheses</p>	



Question Number	Scheme	Marks						
<p><b>6. (a)</b></p> <p><b>(b)</b></p> <p><b>(c)</b></p>	<p style="text-align: center;"><math>[\bar{x} = 49.8]</math></p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; vertical-align: top;"> <math display="block">2 \times 1.96 \left( \frac{\sigma}{\sqrt{8}} \right) = 53.88 - 45.72 = 8.16</math> <math display="block">2 \times 2.5758 \left( \frac{\sigma}{\sqrt{8}} \right) = \frac{8.16 \times 2.5758}{1.96} = 10.7238\dots</math> <math display="block">99\% \text{ CI} = 49.8 \pm \frac{10.7238}{2}</math> </td> <td style="width: 5%; border-left: 1px solid black; border-right: 1px solid black;"></td> <td style="width: 45%; vertical-align: top;"> <math display="block">49.8 + 1.96 \left( \frac{\sigma}{\sqrt{8}} \right) = 53.88 \text{ or}</math> <math display="block">49.8 - 1.96 \left( \frac{\sigma}{\sqrt{8}} \right) = 45.72</math> <math display="block">2.5758 \left( \frac{\sigma}{\sqrt{8}} \right) = \frac{4.08 \times 2.5758}{1.96} = 5.3618\dots</math> <math display="block">99\% \text{ CI} = 49.8 \pm 5.3618</math> </td> </tr> <tr> <td colspan="2" style="text-align: center;"><math>= (44.438\dots, 55.1619\dots)</math></td> <td style="text-align: right;">(awrt 44.4, awrt 55.2)</td> </tr> </table> <p><math>\hat{\mu} = \bar{x} = \frac{91.2}{8} = 11.4</math></p> <p><math>\hat{\sigma}^2 = s^2 = \frac{1145.16 - 8 \times "11.4^2"}{7} = 15.06857\dots</math> awrt 15.1</p> <p>Combined <math>\Sigma x = 10.8 \times 24 + 91.2 = 350.4</math></p> <p>Combined <math>\Sigma x^2 = 1145.16 + 23 \times 17.64 + 24 \times 10.8^2 = 4350.24</math></p> <p>Combined <math>s^2 = \frac{"4350.24" - 32 \times \left( \frac{"350.4"}{32} \right)^2}{31} = 16.56</math></p> <p><math>\frac{s}{\sqrt{n}} = \frac{\sqrt{16.56}}{\sqrt{32}} = 0.719374\dots</math> awrt 0.719</p>	$2 \times 1.96 \left( \frac{\sigma}{\sqrt{8}} \right) = 53.88 - 45.72 = 8.16$ $2 \times 2.5758 \left( \frac{\sigma}{\sqrt{8}} \right) = \frac{8.16 \times 2.5758}{1.96} = 10.7238\dots$ $99\% \text{ CI} = 49.8 \pm \frac{10.7238}{2}$		$49.8 + 1.96 \left( \frac{\sigma}{\sqrt{8}} \right) = 53.88 \text{ or}$ $49.8 - 1.96 \left( \frac{\sigma}{\sqrt{8}} \right) = 45.72$ $2.5758 \left( \frac{\sigma}{\sqrt{8}} \right) = \frac{4.08 \times 2.5758}{1.96} = 5.3618\dots$ $99\% \text{ CI} = 49.8 \pm 5.3618$	$= (44.438\dots, 55.1619\dots)$		(awrt 44.4, awrt 55.2)	<p>M1</p> <p>B1 M1</p> <p>M1</p> <p>A1 (5)</p> <p>B1</p> <p>M1</p> <p>A1 (3)</p> <p>M1</p> <p>M1A1</p> <p>M1 A1</p> <p>M1 A1</p> <p>(7)</p> <p><b>[15]</b></p>
$2 \times 1.96 \left( \frac{\sigma}{\sqrt{8}} \right) = 53.88 - 45.72 = 8.16$ $2 \times 2.5758 \left( \frac{\sigma}{\sqrt{8}} \right) = \frac{8.16 \times 2.5758}{1.96} = 10.7238\dots$ $99\% \text{ CI} = 49.8 \pm \frac{10.7238}{2}$		$49.8 + 1.96 \left( \frac{\sigma}{\sqrt{8}} \right) = 53.88 \text{ or}$ $49.8 - 1.96 \left( \frac{\sigma}{\sqrt{8}} \right) = 45.72$ $2.5758 \left( \frac{\sigma}{\sqrt{8}} \right) = \frac{4.08 \times 2.5758}{1.96} = 5.3618\dots$ $99\% \text{ CI} = 49.8 \pm 5.3618$						
$= (44.438\dots, 55.1619\dots)$		(awrt 44.4, awrt 55.2)						
<b>Notes</b>								
<p><b>(a)</b></p> <p><b>(b)</b></p> <p><b>(c)</b></p>	<p>1<sup>st</sup> M1 use of <math>2z \frac{\sigma}{\sqrt{n}}</math> or <math>z \frac{\sigma}{\sqrt{n}}</math> with <math>1.5 &lt;  z  &lt; 2</math>. Allow <math>\sigma_m</math> for <math>\frac{\sigma}{\sqrt{n}}</math></p> <p>B1 1.96 (or better) and 2.5758 (or better)</p> <p>2<sup>nd</sup> M1 attempt to find width or semi-width of 99% CI with <math> z  &gt; 2</math> Allow <math>\sigma = \frac{4.08 \times \sqrt{8}}{1.96} [= 5.887\dots]</math></p> <p>3<sup>rd</sup> M1 Use of <math>49.8 \pm</math> awrt 5.36 or <math>49.8 \pm 2.5758 \left( \frac{"5.887\dots"}{\sqrt{8}} \right)</math> If <math>\sigma</math> is incorrect then working must be shown.</p> <p>A1 correct interval with (awrt 44.4, awrt 55.2)</p> <p>Correct answer from less accurate <math>z</math>-values scores M1B0M1M1A1</p> <p>B1 11.4 cao</p> <p>M1 full attempt at <math>s^2</math> ft their <math>\bar{x}</math></p> <p>A1 awrt 15.1</p> <p>M1 for correct combined sum (may be implied by combined mean of 10.95)</p> <p>2<sup>nd</sup> M1 for attempt at combined sum of squares <math>1145.16 + (n-1) \times 17.64 + n \times 10.8^2</math> (allow 1 error)</p> <p>1<sup>st</sup> A1 fully correct expression or awrt 4350</p> <p>3<sup>rd</sup> M1 using their values in a complete expression for combined <math>s^2</math> oe</p> <p>2<sup>nd</sup> A1 <math>s^2 = 16.56</math> or <math>s =</math> awrt 4.07 (either of these implies M1M1A1M1A1)</p> <p>4<sup>th</sup> M1 use of <math>\frac{s}{\sqrt{n}}</math> with combined values</p> <p>3<sup>rd</sup> A1 awrt 0.719</p>							

Question Number	Scheme	Marks
<p><b>7. (a)</b></p> <p><b>(b)</b></p> <p><b>(c)</b></p>	$a = 2 \times 180 - 330 = 30$ $b = 4.5^2 \times 2 + 6.7^2 = 85.39$ $X = L - 1.8S$ $E(X) = 330 - 1.8 \times 180 = 6$ $\text{Var}(X) = 6.7^2 + 1.8^2 \times 4.5^2 = 110.5$ $P(X > 0) = P\left(Z > \frac{0-6}{\sqrt{110.5}}\right)$ $P(Z > -0.57) = 0.7157$ $T = S_1 - \frac{S_1 + S_2 + S_3}{3} = \frac{2S_1 - S_2 - S_3}{3}$ $E(T) = 0$ $\text{Var}(T) = \frac{1}{9}(2^2 \times 4.5^2 + 4.5^2 + 4.5^2) = \frac{6}{9}(4.5^2) = 13.5$ $P(T > 5) = P\left(Z > \frac{5-0}{\sqrt{13.5}}\right)$ $P(Z > 1.36) = 1 - 0.9131 = 0.0869$	<p>B1 M1 A1 (3)</p> <p>M1 M1 A1</p> <p>M1 A1 (5)</p> <p>M1 A1 M1</p> <p>M1</p> <p>M1 A1 (6) <b>[14]</b></p>
<b>Notes</b>		
<p><b>(a)</b></p> <p><b>(b)</b></p> <p><b>(c)</b></p>	<p>B1 30 cao M1 <math>2 \times \text{Var}(S) + \text{Var}(L)</math> A1 85.39 (allow 85.4)</p> <p>1<sup>st</sup> M1 Seeing or using <math>E(X) = 6</math> or correct expression for mean 2<sup>nd</sup> M1 <math>\text{Var}(L) + 1.8^2 \text{Var}(S)</math> (condone mixing variances for M1) 1<sup>st</sup> A1 for 110.5 (allow 65.61 + 6.7<sup>2</sup>) 3<sup>rd</sup> M1 standardising with their mean and s.d. leading to a probability <math>p &gt; 0.5</math> 2<sup>nd</sup> A1 awrt 0.716 [calc: 0.7159262...]</p> <p>1<sup>st</sup> M1 realising the need to write as a single distribution using <math>\bar{S} = \frac{S_1 + S_2 + S_3}{3}</math> 1<sup>st</sup> A1 for <math>\frac{2S_1 - S_2 - S_3}{3}</math> 2<sup>nd</sup> M1 Using mean = 0 3<sup>rd</sup> M1 using <math>\text{Var}(aS) = a^2 \text{Var}(S)</math> 4<sup>th</sup> M1 standardising with their mean and sd 2<sup>nd</sup> A1 awrt 0.0868 to awrt 0.0869 [calc: 0.08678...] Note: Assuming <math>S_1</math> and <math>\bar{S}</math> are independent, leads to <math>E(T) = 0</math>, <math>\text{Var}(T) = 27</math>, <math>P(T &gt; 5) = 0.167...</math> scores M0A0M1M0M1A0</p>	