

## Topic Y3 Correlation regression & chi-squared tests (Pre-TT) [48] MS

1.

$H_0$ : The data can be modelled by the theory $H_1$ : The data can't be modelled by the theory. Expected values 918, 306, 306, 102 $TS = \frac{(865 \cdot 918)^2}{918} + \dots$ $= 7.43$ $TS < 7.815$ , do not reject $H_0$ There is insufficient evidence to conclude that the data can't be modelled by the theory	B1  B1 M1  A1 M1 A1  [6]	For both  Can be implied by 7.43  fit TS fit TS	Allow compatible.   $p > 0.05$ do not reject $H_0$ $p = 0.05939$ and conclusion
---	---	--	---

2.

(i)	$S_{lm} = 0.2412$ $S_{hh} = 0.10992$ $S_{mm} = 27.212$ $r = \frac{S_{lm}}{\sqrt{(S_{hh}S_{mm})}}$ $= 0.139 \text{ (3 sfs)}$	B1 M1  A1 3	Allow x or ÷ 5  any one S correct fit their Ss
(ii)	Small, low or not close to 1 or close to 0 oe pts not close to line oe	B1 ft  B1	$1^{st}$ B1 about value of r $2^{nd}$ B1 about diag
(iii)	none or unchanged or "0.139" oe	B1 1	
(iv)	Larger oe	B1 1	
<b>Total</b>		[7]	

3.

(a)	<b>Shot Ranks</b>	10	1	8	6	3	5	7	4	1	9	B1 B1	1.1b	
	<b>Goals Rank</b>	5	6	4	3	2	10	7	9	1	8		1.1b	
	$\sum d^2 = 128$												B1	1.1b
	$1 - \frac{6 \times 128}{10(10^2 - 1)} = 0.224$												M1	2.1
												A1	1.b	
(b)	<b><math>H_0: \rho = 0</math>   <math>H_1: \rho &gt; 0</math></b>											B1	2.5	
	Critical value at the 5% significance level is 0.5636											B1	1.1b	
	0.224 is not significant so do not reject $H_0$											M1	2.4	
	The journalist claim is not supported.											A1	2.2b	
(c)	No change in ranks so no change in $r_s$											B1	2.2a	
<b>(10 marks)</b>														

4.

States both hypotheses using correct notation	AO2.5	B1	<p><math>H_0</math>: Language choice is independent of sex  <math>H_1</math>: Language choice is not independent of sex</p> <table border="1"> <thead> <tr> <th></th> <th>German</th> <th>Mandarin</th> <th></th> </tr> </thead> <tbody> <tr> <th>M</th> <td>7</td> <td>12</td> <td>19</td> </tr> <tr> <th>F</th> <td>23</td> <td>8</td> <td>31</td> </tr> <tr> <td></td> <td>30</td> <td>20</td> <td>50</td> </tr> </tbody> </table> <table border="1"> <thead> <tr> <th></th> <th>German</th> <th>Mandarin</th> </tr> </thead> <tbody> <tr> <th>M</th> <td>11.4</td> <td>7.6</td> </tr> <tr> <th>F</th> <td>18.6</td> <td>12.4</td> </tr> </tbody> </table> $\chi^2 = \sum \frac{(O - E)^2}{E}$ <p>Allow no Yates correction used            FT their observed and expected frequencies</p> $\chi^2 = \sum \frac{( O - E  - 0.5)^2}{E}$ $= \frac{(3.9)^2}{11.4} + \frac{(3.9)^2}{7.6} + \frac{(3.9)^2}{18.6} + \frac{(3.9)^2}{12.4}$ $= 1.33 + 2.00 + 0.82 + 1.23$ <p>ts <math>\chi^2 = 5.38</math>      <math>p = 0.0204</math>            1 tail 5%</p> <p>df = 1      cv = 3.84      ts &gt; 3.84            or <math>p &lt; 0.05</math></p> <p>Reject <math>H_0</math></p> <p>Significant evidence to suggest that language choice is not independent of sex</p>		German	Mandarin		M	7	12	19	F	23	8	31		30	20	50		German	Mandarin	M	11.4	7.6	F	18.6	12.4
	German	Mandarin																										
M	7	12		19																								
F	23	8		31																								
	30	20		50																								
	German	Mandarin																										
M	11.4	7.6																										
F	18.6	12.4																										
Constructs 'correct' contingency table with frequencies shown (PI)	AO3.1b	M1																										
Finds expected frequencies (PI)	AO3.4	B1																										
Evaluates $\chi^2 = \sum \frac{(O - E)^2}{E}$ Allow no Yates correction used FT their observed and expected frequencies	AO1.1a	M1F																										
Obtains $\chi^2$ test statistic or quotes $p$ -value	AO1.1b	A1																										
Evaluates $\chi^2$ model by comparing test statistic and correct critical value Or by comparing $p$ -value with 0.05	AO3.5a	M1																										
Infers that $H_0$ should be rejected	AO2.2b	E1																										
Completes test correctly and gives conclusion in context. (conclusion should not be definite)	AO3.2a	E1																										

5.

5i	$\frac{2685 - \frac{140 \times 106.8}{8}}{3500 - \frac{140^2}{8}}$ $= \frac{2685 - \frac{140 \times 106.8}{8}}{2500 - 2417.5}$ $= \frac{2685 - 1491.3}{2500 - 2417.5}$ $= \frac{1193.7}{82.5} = 14.47$ $= \frac{136}{175} \text{ or } 0.777 \text{ (3 sfs)}$ $y - \frac{106.8}{8} = 0.777(x - \frac{140}{8})$ $y = 0.78x - 0.25 \text{ or better or } y = \frac{136}{175}x - \frac{1}{4}$	M1 A1 M1 A1	4	Correct sub in any correct formula for $b$ (incl. $(x - \bar{x})$ etc)  or $a = \frac{106.8}{8} - 0.777x \frac{140}{8}$ ft $b$ for M1 $\geq 2$ sfs sufficient for coeffs
ii	$0.78 \times 12 - 0.25$ $= 9.1 \text{ (2 sfs)}$	M1 A1f	2	M1: ft their equn A1: dep const term in equn
iii a	Reliable	B1		Just "reliable" for both: B1
b	Unreliable because extrapolating oe	B1	2	
<b>Total</b>			<b>8</b>	

6.

	Marking Instructions	AO	Marks	Typical Solution																				
(a)	Obtains one missing expected value eg Low and Breed A = $\frac{21 \times 27}{105}$	AO1.1a	M1	<table border="1"> <thead> <tr> <th></th> <th>Low</th> <th>Med</th> <th>High</th> </tr> </thead> <tbody> <tr> <td><b>A</b></td> <td>5.4</td> <td>9.6</td> <td>6</td> </tr> <tr> <td><b>B</b></td> <td>5.14</td> <td>9.14</td> <td>5.71</td> </tr> <tr> <td><b>C</b></td> <td>8.23</td> <td>14.63</td> <td>9.14</td> </tr> <tr> <td><b>D</b></td> <td>8.23</td> <td>14.63</td> <td>9.14</td> </tr> </tbody> </table>		Low	Med	High	<b>A</b>	5.4	9.6	6	<b>B</b>	5.14	9.14	5.71	<b>C</b>	8.23	14.63	9.14	<b>D</b>	8.23	14.63	9.14
		Low	Med		High																			
<b>A</b>	5.4	9.6	6																					
<b>B</b>	5.14	9.14	5.71																					
<b>C</b>	8.23	14.63	9.14																					
<b>D</b>	8.23	14.63	9.14																					
	Obtains all expected values. AO if integers	AO1.1b	A1																					
(i)	States both hypotheses using correct language	AO2.5	B1	$H_0$ : Milk yield is independent of breed $H_1$ : Milk yield is not independent of breed 1 tail 1%  $ts = \sum \frac{(O-E)^2}{E} = 19.4$  $\chi^2$ cv for 6 df = 16.81 $(p = 0.00275)$  $19.4 > 16.81$ $(0.00275 < 0.01)$  Reject $H_0$ and conclude that there is evidence to suggest that milk yield is not independent of breed																				
	States critical value ( or $p$ -value)	AO1.1b	B1																					
	Evaluates the $\chi^2$ test statistic by comparing the cv with the ts (or $p$ -value with 0.01)	AO3.5a	R1																					
	Infers $H_0$ rejected	AO2.2b	E1																					
	Concludes correctly in context (conclusion should not be definite)	AO3.2a	E1																					
(ii)	Considers $\frac{(O-E)^2}{E}$ or $(O-E)$ to identify largest sources of association [Do not allow mark if no reference to why source selected]	AO2.4	E1	Largest sources of association Breed A/High and Breed B/ Low.  Far more than expected Breed A cows observed to have high milk yield or Far more than expected Breed B cows observed to have low milk yield																				
	Interprets main source(s) of association in context	AO3.2a	E1dep																					
	<b>Total</b>		<b>9</b>																					