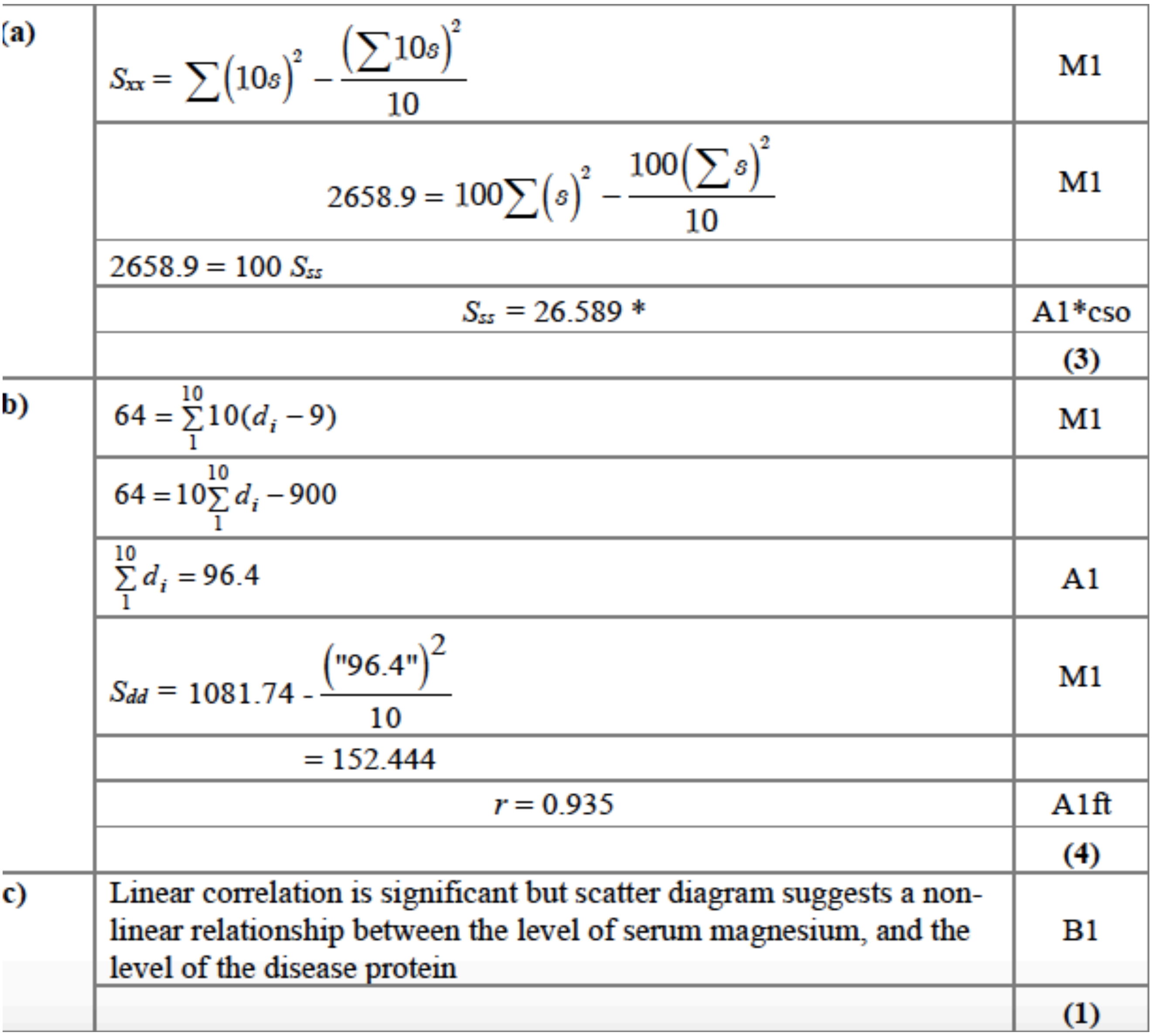
Topic Y3 Correlation regression & chi-squared tests (Pre-TT B) [47] MS

1.

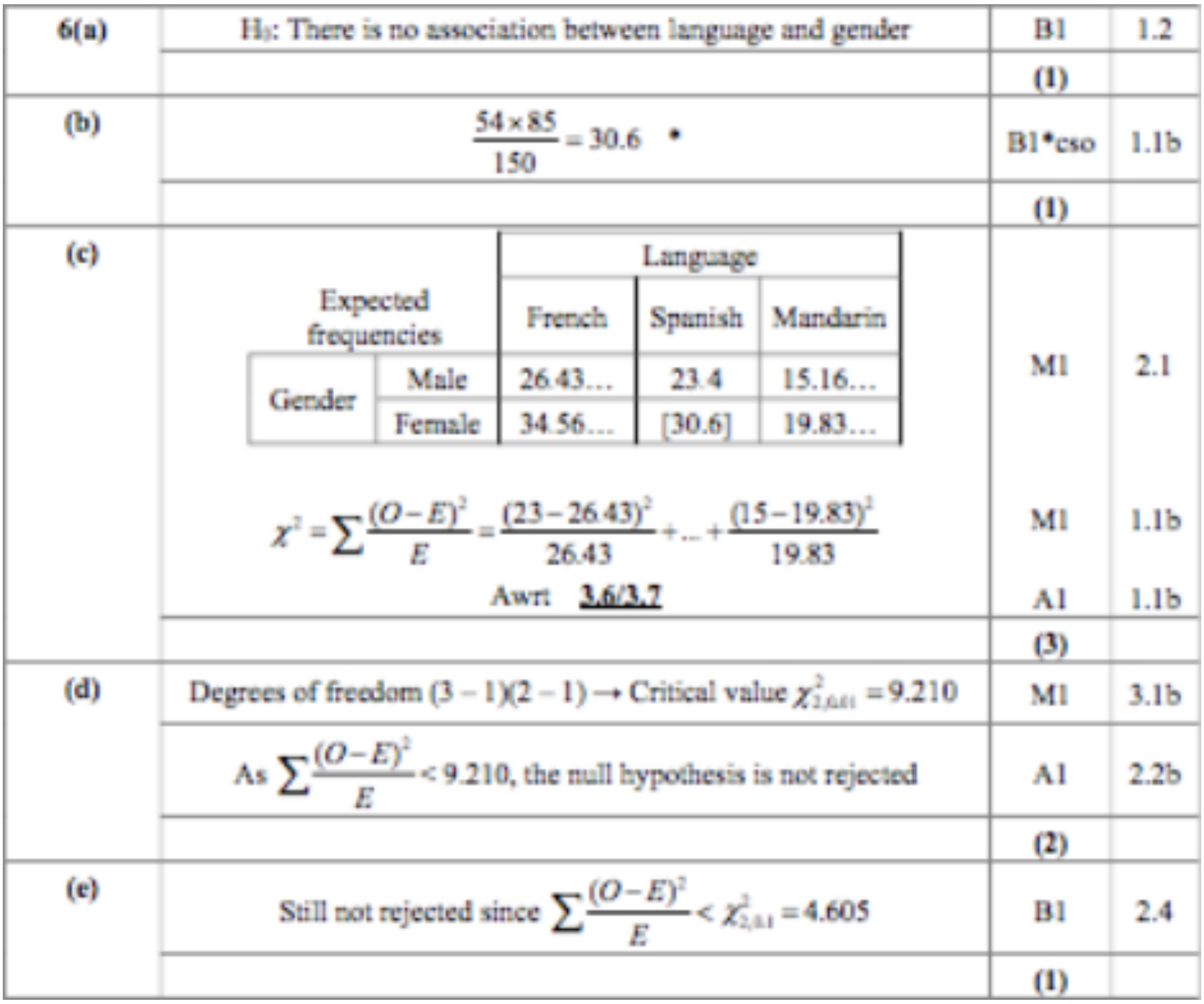


(a)	$S_{xx} = \sum(10s)^2 - \dfrac{\left(\sum 10s\right)^2}{10}$	M1
	$2658.9 = 100\sum(s)^2 - \dfrac{100\left(\sum s\right)^2}{10}$	M1
	$2658.9 = 100\,S_{ss}$	
	$S_{ss} = 26.589$ *	A1*cso
		(3)
b)	$64 = \sum_1^{10} 10(d_i - 9)$	M1
	$64 = 10\sum_1^{10} d_i - 900$	
	$\sum_1^{10} d_i = 96.4$	A1
	$S_{dd} = 1081.74 - \dfrac{(\text{"}96.4\text{"})^2}{10}$	M1
	$= 152.444$	
	$r = 0.935$	A1ft
		(4)
c)	Linear correlation is significant but scatter diagram suggests a non-linear relationship between the level of serum magnesium, and the level of the disease protein	B1
		(1)

2.

<table>
<tr><td>1</td><td>
<table>
<tr><td>Plate</td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td><td>F</td><td>G</td><td>H</td><td>I</td><td>J</td></tr>
<tr><td>Judge</td><td>7</td><td>3</td><td>2</td><td>1</td><td>8</td><td>10</td><td>4</td><td>6</td><td>5</td><td>9</td></tr>
<tr><td>Age</td><td>8</td><td>1</td><td>2</td><td>3</td><td>10</td><td>7</td><td>9</td><td>6</td><td>4</td><td>5</td></tr>
</table>
</td><td>M1
A1</td><td></td><td>1.1b
1.1b</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td>$\sum d^2 = 1+4+0+4+4+9+25+0+1+16 \quad [=64]$</td><td>M1</td><td></td><td>1.1b</td></tr>
<tr><td></td><td>$r_s = 1 - \dfrac{6\times\text{"}64\text{"}}{10\times 99}$</td><td>M1</td><td></td><td>1.1b</td></tr>
<tr><td></td><td>= 0.61212… = awrt <u>0.612</u></td><td>A1</td><td></td><td>1.1b</td></tr>
<tr><td></td><td colspan="4">Notes</td></tr>
<tr><td></td><td colspan="4">1st M1 an attempt to rank judge against actual ages</td></tr>
<tr><td></td><td colspan="4">1st A1 both correct ranks</td></tr>
<tr><td></td><td colspan="4">2nd M1 for an attempt to find $\sum d^2$ (some correct d values found and sum attempted)</td></tr>
<tr><td></td><td colspan="4">3rd M1 for using their $\sum d^2$ in formula for r_s with $n = 10$</td></tr>
<tr><td></td><td colspan="4">2nd A1 for awrt 0.612 or exact fraction e.g. $\frac{101}{165}$</td></tr>
</table>

3.



<table>
<tr><td>6(a)</td><td>H₀: There is no association between language and gender</td><td>B1</td><td>1.2</td></tr>
<tr><td></td><td></td><td>(1)</td><td></td></tr>
<tr><td>(b)</td><td>$\frac{54 \times 85}{150} = 30.6$ *</td><td>B1*cso</td><td>1.1b</td></tr>
<tr><td></td><td></td><td>(1)</td><td></td></tr>
<tr><td>(c)</td><td>
<table>
<tr><td colspan="2">Expected frequencies</td><td colspan="3">Language</td></tr>
<tr><td></td><td></td><td>French</td><td>Spanish</td><td>Mandarin</td></tr>
<tr><td>Gender</td><td>Male</td><td>26.43…</td><td>23.4</td><td>15.16…</td></tr>
<tr><td></td><td>Female</td><td>34.56…</td><td>[30.6]</td><td>19.83…</td></tr>
</table>
</td><td>M1</td><td>2.1</td></tr>
<tr><td></td><td>$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(23-26.43)^2}{26.43} + \ldots + \frac{(15-19.83)^2}{19.83}$</td><td>M1</td><td>1.1b</td></tr>
<tr><td></td><td>Awrt <u>3.6/3.7</u></td><td>A1</td><td>1.1b</td></tr>
<tr><td></td><td></td><td>(3)</td><td></td></tr>
<tr><td>(d)</td><td>Degrees of freedom (3 – 1)(2 – 1) → Critical value $\chi^2_{2,0.01} = 9.210$</td><td>M1</td><td>3.1b</td></tr>
<tr><td></td><td>As $\sum \frac{(O-E)^2}{E} < 9.210$, the null hypothesis is not rejected</td><td>A1</td><td>2.2b</td></tr>
<tr><td></td><td></td><td>(2)</td><td></td></tr>
<tr><td>(e)</td><td>Still not rejected since $\sum \frac{(O-E)^2}{E} < \chi^2_{2,0.1} = 4.605$</td><td>B1</td><td>2.4</td></tr>
<tr><td></td><td></td><td>(1)</td><td></td></tr>
</table>

4.

4(i)	2608*p* $p = e^{-3.87}3.87^6/6!$ (×2608=253.82)	M1 A1	*p* from Poisson From 253.8 or 254 seen
	$(273-253.82)^2/253.82$ =1.449	M1 A1 4	 Answer between 1.445 and 1.460
(ii)	Number of cells – 1 (estimated mean) – 1(same totals)	B1 1	Not 11-1
(iii)	H$_0$: A Poisson distribution fits the data H$_1$: A Poisson distribution does not fit the data CV = 15. 99 13.0 < CV and do not reject Ho accept that there is insufficient evidence that a Poisson distribution does not fit data	 B1 B1 M1 A1 4 (9)	For both hypotheses Their CV Sufficient evidence that Poisson distribution fits data , OK

5.

(i) Correct subst in $\geq$ two S formulae $$\frac{14464.1-\frac{265\times 274.6}{5}}{\sqrt{\left(14176.54-\frac{265^2}{5}\right)\left(15162.22-\frac{274.6^2}{5}\right)}}$$ $= -0.868$ (3 sfs)	M1 M1 A1 3	Any correct version or $$\frac{14464.1-5\times 53\times 54.92}{\sqrt{(14176.54-5\times 53^2)(15162.22-5\times 54.92^2)}}$$ or fully correct method with $(x - \overline{x})^2$ etc
(ii) No difference oe	B1 1	Or slightly diff or more acc because of rounding errors when mult by 2.54 oe Not just "more accurate"
(iii)Choose y on x stated	B1ind	or implied, eg by S_{xy}/S_{xx} or $y = ax + b$
$$\frac{14464.1-\frac{265\times 274.6}{5}}{14176.54-\frac{265^2}{5}}$$ or -0.682 $y - {}^{274.6}/_{5} = (\text{their} - 0.682)(x - {}^{265}/_{5})$ $y = 91(.1) - 0.68(2)\,x$ 49.9 (3sfs) or 50	M1 M1ind A1 5 A1	If state x on y, but wking is y on x: B1 or their $\frac{-89.7}{131.54}$ seen or $\frac{14464.1-5\times 53\times 54.92}{14176.54-5\times 53^2}$ or correct subst into a correct formula $\frac{S_{xy}}{S_{xx}}$ or $a = {}^{274.6}/_{5}$ - (their $-$ 0.682) x ${}^{265}/_{5}$ Simplif to 3 terms. Coeffs to $\geq$ 2 sfs cao Use of x on y: equiv M mks as above
	9	

6.

(i)		Proper					
		P	F				
	P	**31**	**11**	42	B1		Two correct
Trial							
	F	**5**	13	**18**	B1		Others correct
		36	**24**	60		**2**	

(ii) (H_0: Trial results and Proper results are independent.)			
E-values: 25.2 16.8	M1		One correct. Ft marginals in (i)
10.8 7.2	A1		All correct
$\chi^2 = 5.3^2(25.2^{-1}+10.8^{-1}+16.8^{-1}+7.2^{-1})$	M1		Allow two errors
	A1		With Yates' correction
$= 9.289$	A1		art 9.29
Compare correctly with 7.8794	M1		Or 7.88
There is evidence that results are not independent.	A1 √	7	Ft χ^2_{calc}.