Paper 2 Option G

Further Statistics 1 Mark Scheme (Section A)

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AOs</th></tr>
<tr><td>1(a)</td><td>H_0: 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" rowspan="2">Expected frequencies</td><td colspan="3">Language</td></tr>
<tr><td>French</td><td>Spanish</td><td>Mandarin</td></tr>
<tr><td rowspan="2">Gender</td><td>Male</td><td>26.43…</td><td>23.4</td><td>15.16…</td></tr>
<tr><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>$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(23-26.43)^2}{26.43} + ... + \frac{(15-19.83)^2}{19.83}$</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>
<tr><td colspan="4">(8 marks)</td></tr>
</table>

Notes:

(a)
B1: For correct hypothesis in context

(b)
B1*: For a correct calculation leading to the given answer and no errors seen

(c)
M1: For attempt at $\frac{\text{(Row Total)(Column Total)}}{\text{(Grand Total)}}$ to find expected frequencies
M1: For applying $\sum \frac{(O-E)^2}{E}$
A1: awrt 3.6 or 3.7

(d)
M1: For using degrees of freedom to set up a χ^2 model critical value
A1: For correct comparison and conclusion

(e)
A1ft: For correct conclusion with supporting reason

Question	Scheme	Marks	AOs
2(a)	$-4 = 2 - 5\text{E}(X)$	M1	3.1a
	$\text{E}(X) = 1.2$		
	$-1\times c + 0\times a + 1\times a + 2\times b + 3\times c = 1.2$	M1	1.1b
	$a + 2b + 2c = 1.2 \quad \boxed{1}$		
	$\text{P}(Y \geqslant -3) = 0.45$ gives $\text{P}(2 - 5X \geqslant -3) = 0.45$ i.e. $\text{P}(X \leqslant 1) = 0.45$	M1	2.1
	$2a + c = 0.45 \quad \boxed{2}$		
	$2a + b + 2c = 1 \quad \boxed{3}$	M1	1.1b
	$\begin{pmatrix} 1 & 2 & 2 \\ 2 & 0 & 1 \\ 2 & 1 & 2 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1.2 \\ 0.45 \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 2 & -3 \\ -2 & -3 & 4 \end{pmatrix}\begin{pmatrix} 1.2 \\ 0.45 \\ 1 \end{pmatrix}$ <u>or</u>	M1	1.1b
	e.g. $\boxed{3} - \boxed{2} \Rightarrow b + c = 0.55$ sub. $2(b+c)$ into $\boxed{1} \Rightarrow a = 0.1$ etc		
	$a = 0.1 \quad b = 0.3 \quad c = 0.25$	A1 A1	1.1b 1.1b
		(7)	
(b)	$\text{Var}(Y) = 75 - (-4)^2$ <u>or</u> 59	M1	1.1a
	$[\text{Var}(Y) = 5^2\text{Var}(X)$ implies$]$ $\text{Var}(X) = 2.36$	A1	1.2
		(2)	
(c)	$\text{P}(Y > X) = \text{P}(2 - 5X > X) \rightarrow \text{P}(X < \tfrac{1}{3})$	M1	3.1a
	$\text{P}(X < \tfrac{1}{3}) = a + c = 0.35$	A1ft	1.1b
		(2)	
		(11 marks)	

Notes:

(a)

M1: For using given information to find an expression for E(X) i.e. use of $\text{E}(Y) = 2 - 5\text{E}(X)$

M1: For use of $\sum x\text{P}(X = x) = \text{‘}1.2\text{’}$

M1: For use of $\text{P}(Y \geqslant -3) = 0.45$ to set up the argument for solving by forming an equation in a and c

M1: For use of $\sum \text{P}(X = x) = 1$

M1: For solving their 3 linear equations (matrix or elimination)

A1: For any 2 of a, b or c correct

A1: For all 3 correct values

Question 2 notes continued:

Another method for part (a) is:

M1: For using given information to find the probability distribution for Y leading to an expression for E(Y)

M1: For use of $\sum y\text{P}(Y = y) = -4$

M1: For use of P($Y \geqslant -3$) = 0.45 to set up the argument for solving by forming an equation in a and c

M1: For use of $\sum \text{P}(Y = y) = 1$

M1: For solving their 3 linear equations (matrix or elimination)

A1: For any 2 of a, b or c correct

A1: For all 3 correct values

(b)

M1: For use of Var(Y) = E(Y^2) – [E(Y)]^2 (may be implied by a correct answer)

A1: For use of Var(aX) = a^2 Var(X) to reach 2.36 or exact equivalent

(c)

M1: For rearranging to the form P($X < k$)

A1ft: 0.1’ + ‘025’ (provided their a and c and their $a + c$ are all probabilities)

Another method for part (c) is:

M1: For comparing distribution of X with distribution of Y to identify X = –1 and X = 0

A1ft: ‘0.1’ + ‘025’ (provided their a and c and their $a + c$ are all probabilities)

Question	Scheme	Marks	AOs
3(a)	$X \sim \text{Po}(2.6)$ $Y \sim \text{Po}(1.2)$		
	P(each hire 2 in 1 hour) $= \text{P}(X=2) \times \text{P}(Y=2) = 0.25104\ldots \times 0.21685\ldots$	M1	3.3
	$= 0.05444\ldots$ awrt **<u>0.0544</u>**	A1	1.1b
		(2)	
(b)	$W = X + Y \rightarrow \ W \sim \text{Po}(3.8)$	M1	3.4
	$\text{P}(W = 3) = 0.20458\ldots.$ awrt **<u>0.205</u>**	A1	1.1b
		(2)	
(c)	$T \sim \text{Po}((2.6+1.2)\times 2)$	M1	3.3
	$\text{P}(T < 9) = 0.64819\ldots$ awrt **<u>0.648</u>**	A1	1.1b
		(2)	
(d)	**(i)** Mean = np = **<u>2.4</u>**	B1	1.1b
	(ii) Variance = $np(1-p) = 2.3904$ awrt **<u>2.39</u>**	B1	1.1b
		(2)	
(e)	**(i)** [$D \sim \text{Po}(2.4)$ $\text{P}(D \leqslant 4)$] $= 0.9041\ldots$ awrt **<u>0.904</u>**	B1	1.1b
	(ii) Since n is large and p is small/mean is approximately equal to variance	B1	2.4
		(2)	
			(10 marks)

Notes:

(a)

M1: For $\text{P}(X=2) \times \text{P}(Y=2)$ from $X \sim \text{Po}(2.6)$ and $Y \sim \text{Po}(1.2)$ i.e. correct models (may be implied by correct answer)

A1: **awrt 0.0544**

(b)

M1: For combining Poisson distributions and use of Po('3.8') (may be implied by correct answer)

A1: **awrt 0.205**

(c)

M1: For setting up a new model and attempting mean of Poisson distribution (may be implied by correct answer)

A1: **awrt 0.648**

(d)(i)

B1: For **2.4**

(d)(ii)

B1: For **awrt 2.39**

(e)(i)

B1: For **awrt 0.904**

(e)(ii)

B1: For a correct explanation to support use of Poisson approximation in this case

Question	Scheme	Marks	AOs
4(a)	**(i)** $P(X = 1) = 0.34523\ldots$ awrt **0.345**	B1	1.1b
	(ii) $P(X \leqslant 4) = 0.98575\ldots$ awrt **0.986**	B1	1.1b
		(2)	
(b)	$\frac{(0\times10)+1\times16+2\times7+3\times4+4\times2+(5\times0)+6\times1}{40}=1.4*$	B1*cso	1.1b
		(1)	
(c)	$r = 40 \times$ '0.34523…' $s = 40 \times$ '1 – 0.986…'	M1	3.4
	$r =$ **13.81** $s =$ **0.57**	A1ft	1.1b
		(2)	
(d)	H_0: The Poisson distribution is a suitable model H_1: The Poisson distribution is not a suitable model	B1	3.4
	[Cells are combined when expected frequencies < 5] So combine the last 3 cells	M1	2.1
	$\chi^2 = \sum\frac{(O-E)^2}{E} = \frac{(10-9.86)^2}{9.86} + \ldots + \frac{(7-(4.51+1.58+0.57))^2}{(4.51+1.58+0.57)}$	M1	1.1b
	awrt **1.1**	A1	1.1b
	Degrees of freedom = 4 – 1 – 1 = 2	B1	3.1b
	(Do not reject H_0 since $1.10 < \chi^2_{2,(0.05)} = 5.991$). The number of mortgages approved each week follows a Poisson distribution	A1	3.5a
		(6)	
			(11 marks)

Notes:

(a)(i)
B1: awrt 0.345

(a)(ii)
B1: awrt 0.986

(b)
B1*: For a fully correct calculation leading to given answer with no errors seen

(c)
M1: For attempt at r or s (may be implied by correct answers)
A1ft: For both values correct (follow through their answers to part (a))

(d)
B1: For both hypotheses correct (lambda should not be defined so correct use of the model)
M1: For understanding the need to combine cells before calculating the test statistic (may be implied)
M1: For attempt to find the test statistic using $\chi^2 = \sum\frac{(O-E)^2}{E}$
A1: awrt 1.1
B1: For realising that there are 2 degrees of freedom leading to a critical value of $\chi_2^{\,2}(0.05) = 5.991$
A1: Concluding that a Poisson model is suitable for the number of mortgages approved each week

Further Statistics 2 Mark Scheme (Section B)

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AOs</th></tr>
<tr><td>5(a)</td><td>
<table>
<tr><th>Competitor</th><th>A</th><th>B</th><th>C</th><th>D</th><th>E</th><th>F</th><th>G</th><th>H</th></tr>
<tr><td>Judge 1’s ranks</td><td>8</td><td>4</td><td>7</td><td>6</td><td>5</td><td>1</td><td>3</td><td>2</td></tr>
<tr><td>Judge 2’s ranks</td><td>8</td><td>5</td><td>6</td><td>7</td><td>3</td><td>1</td><td>4</td><td>2</td></tr>
<tr><td>d^2</td><td>0</td><td>1</td><td>1</td><td>1</td><td>4</td><td>0</td><td>1</td><td>0</td></tr>
</table>
$\sum d^2 = 8$
$r_s = 1 - \frac{6 \times 8}{8(64 - 1)}$</td><td>M1
M1
dM1</td><td>1.1b
1.1b
1.1b</td></tr>
<tr><td></td><td>$r_s = 0.90476 \ldots$ awrt <u>0.905</u></td><td>A1</td><td>1.1b</td></tr>
<tr><td></td><td></td><td>(4)</td><td></td></tr>
<tr><td>(b)</td><td>H0: $\rho_s = 0$ H1: $\rho_s > 0$</td><td>B1</td><td>2.5</td></tr>
<tr><td></td><td>Critical value $\rho_s = 0.8333$</td><td>B1</td><td>1.1b</td></tr>
<tr><td></td><td>$r_s = 0.905$ lies in the critical region/reject H_0</td><td>M1</td><td>2.1</td></tr>
<tr><td></td><td>The two judges are in agreement.</td><td>A1</td><td>2.2b</td></tr>
<tr><td></td><td></td><td>(4)</td><td></td></tr>
<tr><td>(c)</td><td>E.g. The data is unlikely to be from a bivariate normal distribution (competitor A)/The emphasis here is on the ranks and not the individual scores.</td><td>B1</td><td>2.4</td></tr>
<tr><td></td><td></td><td>(1)</td><td></td></tr>
<tr><td>(d)</td><td>Both show positive correlation, but the judges agree more on the beam (since 0.952 is closer to 1)</td><td>B1</td><td>2.2b</td></tr>
<tr><td></td><td></td><td>(1)</td><td></td></tr>
<tr><td colspan="4" align="right">(10 marks)</td></tr>
</table>

Notes:

(a)

M1: For an attempt to rank at least one row (at least four correct)

M1: For an attempt at d^2 row for their ranks

M1: Dependent on 1st M1 for use of $r_s = 1 - \frac{6 \times 8}{8(64 - 1)}$ with their $\sum d^2$

A1: For awrt 0.905

(b)

B1: Both hypotheses stated in terms of ρ_s

B1: For correct critical value

M1: For comparing their ‘0.905’ with their ‘0.8333’

A1: For a correct contextual conclusion with no contradictions seen

(c)

B1: For a correct explanation to support the use of Spearman

(d)

B1: For a correct comparison of the correlation coefficients

Question	Scheme	Marks	AOs
6(a)	P(X < 3) = $\int_1^3 \frac{1}{18}(11-2x)\,\mathrm{d}x$ <u>or</u> area of trapezium	M1	1.1a
	$= \left[\frac{1}{18}(11x - x^2)\right]_1^3$		
	$= \dfrac{7}{9}$	A1	1.1b
		(2)	
(b)	Since P(X < 3) > 0.75, the upper quartile is less than 3	B1ft	2.2a
		(1)	
(c)	$\mathrm{E}(X^2) = \int_1^4 \frac{1}{18}x^2(11-2x)\,\mathrm{d}x \left[= \dfrac{23}{4}\right]$	M1	1.1b
	$\mathrm{Var}(X) = \dfrac{23}{4} - \left(\dfrac{9}{4}\right)^2$	M1	1.1b
	$= \dfrac{11}{16}$	A1	1.1b
		(3)	
(d)	F(4) = 1 → $\frac{1}{18}(11(4) - 4^2 + c) = 1$ <u>or</u> F(1) = 0 → $\frac{1}{18}(11(1) - 1^2 + c) = 0$	M1	2.1
	c = –10 *	A1*cso	1.1b
		(2)	
(e)	F(m) = 0.5	M1	1.2
	$\frac{1}{18}(11m - m^2 - 10) = 0.5 \rightarrow m^2 - 11m + 19 = 0$ and attempt to solve	M1	1.1b
	$m = \dfrac{11 \pm \sqrt{11^2 - 4(19)}}{2}$ [= 2.1458 or 8.8541…]		
	m = 2.1458… <u>**2.15**</u> (only)	A1	2.2a
		(3)	
		(11 marks)	

Notes:

(a)

M1: For integrating f(x) with correct limits **or** for finding area of trapezium

A1: For $\dfrac{7}{9}$ (allow awrt 0.778)

(b)

B1ft: For comparison of their (a) with 0.75 and concluding that the upper quartile is less than 3

(c)

M1: For an attempt to find $\mathrm{E}(X^2)$

M1: For use of $\mathrm{Var}(X) = \mathrm{E}(X^2) - \left(\dfrac{9}{4}\right)^2$

A1: For $\dfrac{11}{16}$ (allow awrt 0.688)(M1 marks may be implied by a correct answer)

Question 6 notes continued:
(d) **M1:** For use of F(4) = 1 or F(1) = 0 **A1*cso:** For a fully correct solution leading to given answer with no errors seen
(e) **M1:** For use of F(m) = 0.5 **M1:** For setting up quadratic and attempt to solve **A1:** For 2.15 and rejecting the other solution

Question	Scheme	Marks	AOs
7(a)	$r = \frac{284.4 - \frac{251(12)}{10}}{\sqrt{10.36 \times 40.9}}$	M1	1.1b
	$r = -0.79671...$ awrt **-0.797**	A1	1.1b
		(2)	
(b)	$b = \frac{'-16.4'}{10.36}$	M1	3.3
	$a = \frac{251}{10} - 'b'\frac{12}{10}$	M1	1.1b
	$y = 27.0 - 1.58x$	A1	1.1b
		(3)	
(c)	$y = [27.0 - 1.58(2)]$ = 23.84 awrt **23.8**	B1ft	3.4
		(1)	
(d)	$\text{RSS} = 40.9 - \frac{(-16.4)^2}{10.36}$	M1	1.1b
	$\text{RSS} = 14.938...$ awrt **14.9**	A1	1.1b
		(2)	
(e)	$\sum$residuals = 0 → –0.63 + (–0.32) + … + f + (–1.88) = 0	M1	3.1a
	f = **–1.04**	A1	1.1b
		(2)	
(f)	The residuals should be randomly scattered above and below zero so linear model may not be appropriate	B1	3.5b
		(1)	
			(11 marks)

Notes:

(a)
M1: For a complete correct method for finding r
A1: For awrt –0.797

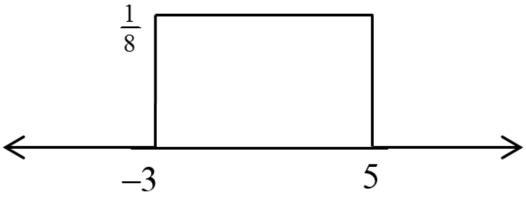
(b)
M1: For use of a correct model i.e. a correct expression for b (ft their S_{xy})
M1: For use of a correct model i.e. a correct (ft) expression for a
A1: For $y = 27.0 - 1.58x$ [a correct answer here can imply both method marks]

(c)
B1: For awrt 23.8 (evaluating their model found in part (b) with $x = 2$)

(d)
M1: For a correct expression for RSS
A1: For awrt 14.9

(e)
M1: For use of $\sum$residuals = 0 [Use of regression equation needs correct sign]
A1: For –1.04

(f)
B1: For identifying that the residuals are not randomly scattered above and below zero and concluding the linear regression model may not be appropriate

Question	Scheme	Marks	AOs
8(a)		B1 (shape) B1 (labels)	1.1b 1.1b
		(2)	
(b)	$\mathrm{P}(X < 2(k - X)) = \mathrm{P}(X < \tfrac{2}{3}k)$	M1	3.1a
	$\dfrac{\frac{2}{3}k-(-3)}{5-(-3)} = 0.25$	M1	1.1b
	$k = -\dfrac{3}{2}$	A1	1.1b
		(3)	
(c)	$\mathrm{E}(X^3) = \displaystyle\int_{-3}^{5} \frac{1}{5-(-3)} x^3 \mathrm{d}x$	M1	2.1
	$= \left[\dfrac{1}{32}x^4\right]_{-3}^{5} = \tfrac{1}{32}(5^4 - (-3)^4)$	dM1	1.1b
	$= 17$*	A1*cso	1.1b
		(3)	
		(8 marks)	

Notes:

(a)
B1: For correct shape
B1: For correct labels

(b)
M1: For simplifying to $\mathrm{P}(X < \tfrac{2}{3}k)$
M1: For equating probability expression to 0.25
A1: For $-\dfrac{3}{2}$

Another method for part (b) is:
M1: For understanding $2[k - x] = -1$ and $x = -1$
M1: For substitution and attempt to solve
A1: For $-\dfrac{3}{2}$

(c)
B1: For integrating $x^3\,\mathrm{f}(x)$
M1: For use of correct limits (dependent on previous M1)
A1*: For fully correct solution leading to the given answer with no errors seen