



Solutions

1.

<b>x</b>	0	1	2	3	≥4	<b>M1</b>
<b>O<sub>i</sub></b>	17	31	19	14	19	
<b>E<sub>i</sub></b>	12.2	27.0	28.5	19.0	13.3	
$\frac{(O - E)^2}{E}$	1.89	0.59	3.17	1.32	2.44	
$\sum \frac{(O-E)^2}{E} = 9.41$					AWRT 9.4	<b>M1A1</b>
$v = 5 - 2 = 3$						<b>B1</b>
$\chi^2_3 (5\%) = 7.815$						<b>B1</b>
<p>H<sub>0</sub>: Binomial distribution is a good/suitable model/fit                      [Condone: B(20, 0.1) is...]</p> <p>H<sub>1</sub>: Binomial distribution is not a suitable model</p> <p>(for hypotheses) allow just “X ~ B(20, 0.1)” for null etc.</p> <p>(Significant result) Binomial distribution is not a suitable model</p>						<b>both B1</b>
						<b>A1</b>





1. An area of grass was sampled by placing a  $1\text{ m} \times 1\text{ m}$  square randomly in 100 places. The numbers of daisies in each of the squares were counted. It was decided that the resulting data could be modelled by a Poisson distribution with mean 2. The expected frequencies were calculated using the model.

The following table shows the observed and expected frequencies.

Number of daisies	Observed frequency	Expected frequency
0	8	13.53
1	32	27.07
2	27	$r$
3	18	$s$
4	10	9.02
5	3	3.61
6	1	1.20
7	0	0.34
$\geq 8$	1	0.12

- a. Find values for  $r$  and  $s$ . (2)
- b. Using a 5% significance level, test whether or not this Poisson model is suitable. State your hypotheses clearly. (7)

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

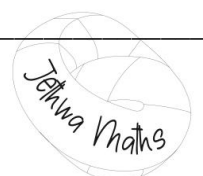
---

---

---

---

---



## Solutions

1a.

$r = 27.07$	<b>A1</b>
$s = 18.04$	<b>A1</b>

1b.

$H_0$ : A Poisson model $Po(2)$ is a suitable model. $H_1$ : A Poisson model $Po(2)$ is not a suitable model.	<b>Both</b> <b>B1</b>
Amalgamate data	<b>M1</b>
$\sum \frac{(O-E)^2}{E} = 3.28$ (awrt)	<b>M1A1</b>
$\nu = 6 - 1 = 5$	<b>B1</b>
$\chi^2_5$ (5 %) = 11.070 (follow through their degrees of freedom)	<b>B1</b>
$3.25 < 11.070$ There is insufficient evidence to reject $H_0$ , $Po(2)$ is a suitable model.	<b>A1</b>





1. Year 12 students at Coron Academy choose to participate in one of four sports during the Spring term, The students' choices are summarised in the table.

	Squash	Badminton	Archery	Hockey	Total
Male	5	16	30	19	70
Female	4	20	33	53	110
Total	9	36	63	72	180

Use a  $\chi^2$  test, at the 5% level of significance, to determine whether the choice of sport is independent of gender.

(10)

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

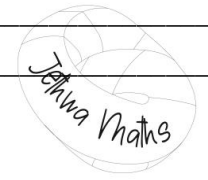
---

---

---

---

---



## Solutions

1.

$H_0$ : Choice independent of gender					<b>B1</b>
	<b>Squash</b>	<b>Badminton</b>	<b>Archery</b>	<b>Hockey</b>	<b>M1</b>
<b>Male</b>	5/3.5	16/14	30/24.5	19/28	
<b>Female</b>	4/3.5	20/22	33/38.5	53/44	
Combine Squash and Badminton					<b>M1</b>
	<b>Squash &amp; Badminton</b>	<b>Archery</b>	<b>Hockey</b>		<b>M1</b> <b>M1</b>
<b>Male</b>	21/17.5	30/24.5	19/28		
<b>Female</b>	24/27.5	33/38.5	53/44		
$\chi^2$ values					<b>M1</b>
	<b>Squash &amp; Badminton</b>	<b>Archery</b>	<b>Hockey</b>		
<b>Male</b>	0.7000	1.2347	2.8928		
<b>Female</b>	0.4455	0.7857	1.8409		
$\chi^2_{\text{calc}} = 7.90$					<b>A1</b>
$\nu = 2$					<b>B1</b>
$\chi^2_{5\%}(2) = 5.991$					<b>B1</b>
Reject $H_0$ Sufficient evidence, at the 1% level of significance, to support an association between the choice of sport and gender					<b>A1</b>





## Solutions

1.

$H_0: X \sim \text{Geo}(0.5)$ is a suitable model $H_1: X \sim \text{Geo}(0.5)$ is not a suitable model					<b>B1</b>										
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center;"><math>x</math></td> <td style="text-align: center;">1</td> <td style="text-align: center;">2</td> <td style="text-align: center;"><math>\geq 3</math></td> <td style="text-align: center;"><b>Total</b></td> </tr> <tr> <td style="text-align: center;"><b>Expected frequency <math>E_x</math></b></td> <td style="text-align: center;">50</td> <td style="text-align: center;">25</td> <td style="text-align: center;">25</td> <td style="text-align: center;">100</td> </tr> </table>					$x$	1	2	$\geq 3$	<b>Total</b>	<b>Expected frequency <math>E_x</math></b>	50	25	25	100	<b>M1</b>
$x$	1	2	$\geq 3$	<b>Total</b>											
<b>Expected frequency <math>E_x</math></b>	50	25	25	100											
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center;"><math>x</math></td> <td style="text-align: center;">1</td> <td style="text-align: center;">2</td> <td style="text-align: center;"><math>\geq 3</math></td> <td style="text-align: center;"><b>Total</b></td> </tr> <tr> <td style="text-align: center;"><math>\frac{(O_x - E_x)^2}{E_x}</math></td> <td style="text-align: center;">0.72</td> <td style="text-align: center;">0.16</td> <td style="text-align: center;">2.56</td> <td style="text-align: center;">3.44</td> </tr> </table>					$x$	1	2	$\geq 3$	<b>Total</b>	$\frac{(O_x - E_x)^2}{E_x}$	0.72	0.16	2.56	3.44	<b>M1</b>
$x$	1	2	$\geq 3$	<b>Total</b>											
$\frac{(O_x - E_x)^2}{E_x}$	0.72	0.16	2.56	3.44											
$X^2 = 3.44$															
$v = 3 - 1 = 2$					<b>B1</b>										
$\chi^2_2(5\%) = 5.991$					<b>B1</b>										
Since $5.991 > 3.44$ there is insufficient evidence to reject $H_0$ at the 5% level $X \sim \text{Geo}(0.5)$ is a suitable model for the data.					<b>A1</b>										







## Solutions

1a.

$p = \frac{1}{\bar{x}}$ $\bar{x} = \frac{\sum fx}{\sum f} = \frac{(34 \times 1) + (22 \times 2) + (14 \times 3) + (18 \times 4) + (12 \times 5)}{100} = 2.52$	<b>M1</b>
$p = \frac{1}{2.52} \approx 0.397$	<b>A1</b>

1b.

$H_0$ : A geometric distribution is a suitable model $H_1$ : A geometric distribution is not a suitable model							<b>B1</b>														
<table border="1" style="width: 100%; border-collapse: collapse; margin-bottom: 5px;"> <thead> <tr> <th style="padding: 2px 5px;"><math>x</math></th> <th style="padding: 2px 5px;">1</th> <th style="padding: 2px 5px;">2</th> <th style="padding: 2px 5px;">3</th> <th style="padding: 2px 5px;">4</th> <th style="padding: 2px 5px;"><math>\geq 5</math></th> <th style="padding: 2px 5px;">Total</th> </tr> </thead> <tbody> <tr> <td style="padding: 2px 5px;"><b>Expected frequency <math>E_i</math></b></td> <td style="padding: 2px 5px;">39.7</td> <td style="padding: 2px 5px;">23.9</td> <td style="padding: 2px 5px;">14.4</td> <td style="padding: 2px 5px;">8.7</td> <td style="padding: 2px 5px;">13.3</td> <td style="padding: 2px 5px;">100</td> </tr> </tbody> </table>							$x$	1	2	3	4	$\geq 5$	Total	<b>Expected frequency <math>E_i</math></b>	39.7	23.9	14.4	8.7	13.3	100	<b>M1</b>
$x$	1	2	3	4	$\geq 5$	Total															
<b>Expected frequency <math>E_i</math></b>	39.7	23.9	14.4	8.7	13.3	100															
<table border="1" style="width: 100%; border-collapse: collapse; margin-bottom: 5px;"> <thead> <tr> <th style="padding: 2px 5px;"><math>x</math></th> <th style="padding: 2px 5px;">1</th> <th style="padding: 2px 5px;">2</th> <th style="padding: 2px 5px;">3</th> <th style="padding: 2px 5px;">4</th> <th style="padding: 2px 5px;"><math>\geq 5</math></th> <th style="padding: 2px 5px;">Total</th> </tr> </thead> <tbody> <tr> <td style="padding: 2px 5px;"><math>\frac{(O_i - E_i)^2}{E_i}</math></td> <td style="padding: 2px 5px;">0.818</td> <td style="padding: 2px 5px;">0.151</td> <td style="padding: 2px 5px;">0.011</td> <td style="padding: 2px 5px;">9.941</td> <td style="padding: 2px 5px;">0.127</td> <td style="padding: 2px 5px;">11.048</td> </tr> </tbody> </table>							$x$	1	2	3	4	$\geq 5$	Total	$\frac{(O_i - E_i)^2}{E_i}$	0.818	0.151	0.011	9.941	0.127	11.048	<b>M1</b>
$x$	1	2	3	4	$\geq 5$	Total															
$\frac{(O_i - E_i)^2}{E_i}$	0.818	0.151	0.011	9.941	0.127	11.048															
$X^2 = 11.048$																					
$v = 5 - 1 = 4$							<b>B1</b>														
$\chi_4^2(2.5\%) = 11.143$							<b>B1</b>														
Since $11.143 > 11.048$ there is insufficient evidence to reject $H_0$ at the 2.5% level. Geometric distribution is a suitable model for the data.							<b>A1</b>														

