

QQQ – Further Statistics 1 – Chapter 6 – Chi-Squared Tests (Version 1)

Total Marks: 25

(25 = Platinum, 23 = Gold, 20 = Silver, 17 = Bronze)

1. Sharma believes that each computer game he sells appeals equally to all age ranges.

To investigate this, he takes a random sample of 100 people who play these games and asks them which of the games A , B or C they prefer.

The results are summarised in the table below.

Computer game		A	B	C
Age range	< 20	8	15	6
	$20 - 30$	21	12	9
	> 30	6	10	13

(a) Write down hypotheses for a suitable test to assess Sharma's belief. (1)

(b) For the test, calculate the expected frequency for

(i) those players aged under 20 who prefer game C

(ii) those players aged between 20 and 30 who prefer game A (2)

(c) State the degrees of freedom of the test statistic for this test. (1)

Sharma correctly calculates the test statistic for this test to be 11.542 (to 3 decimal places).

(d) Using a 5% significance level, and stating your critical value, comment on Sharma's belief.

(2)

(Total 6 marks)

2. Kelly throws a tetrahedral die n times and records the number on which it lands for each throw.

She calculates the expected frequency for each number to be 43 if the die was unbiased.

The table below shows three of the frequencies Kelly records but the fourth one is missing.

Number	1	2	3	4
Frequency	47	34	36	x

(a) Show that $x = 55$ (1)

Kelly wishes to test, at the 5% level of significance, whether or not there is evidence that the tetrahedral die is unbiased.

(b) Explain why there are 3 degrees of freedom for this test. (1)

(c) Stating your hypotheses clearly and the critical value used, carry out the test. (5)

(Total 7 marks)

3. Table 1 below shows the number of car breakdowns in the *Snoreap* district in each of 60 months.

Number of car breakdowns	0	1	2	3	4	5
Frequency	12	11	19	14	3	1

Table 1

Anja believes that the number of car breakdowns per month in *Snoreap* can be modelled by a Poisson distribution. Table 2 below shows the results of some of her calculations.

Number of car breakdowns	0	1	2	3	4	≥ 5
Observed frequency (O_i)	12	11	19	14	3	1
Expected frequency (E_i)	9.92			9.64	4.34	

Table 2

- (a) State suitable hypotheses for a test to investigate Anja's belief. (1)
- (b) Explain why Anja has changed the label of the final column to ≥ 5 (1)
- (c) Showing your working clearly, complete Table 2 (4)
- (d) Find the value of $\frac{(O_i - E_i)^2}{E_i}$ when the number of car breakdowns is
- (i) 1
- (ii) 3 (2)
- (e) Explain why Anja used 3 degrees of freedom for her test. (2)

The test statistic for Anja's test is 6.54 to 2 decimal places.

- (f) Stating the critical value and using a 5% level of significance, complete Anja's test. (2)

(Total 12 marks)

Solutions

1.

Mark scheme (a)

Scheme	Marks	AO
H_0 : There is <u>no association</u> between <u>age</u> range and preferred <u>game</u> H_1 : There <u>is</u> an <u>association</u> between <u>age</u> range and preferred <u>game</u>	B1	1.2
	(1)	

Notes

B1 for both hypotheses in context

must mention independence/dependence or no association/association

must mention age and game at least once

Do not allow e.g. ' H_0 : games sold appeal equally to all age ranges'

Use of 'correlation', 'link', 'relationship', 'connection' is B0 but allow for 2nd B1 in (d)

ignore comments about sales

Mark scheme (b)

Scheme	Marks	AO
(i) $\frac{29 \times 28}{100} = \underline{\underline{8.12}}$ (ii) $\frac{42 \times 35}{100} = \underline{\underline{14.7}}$	M1 A1	1.1b 1.1b
	(2)	

Notes

M1 for one correct expression or correct value.

A1 for both correct values (correct to 1 dp). Accept 8.1

Mark scheme (c)

Scheme	Marks	AO
--------	-------	----

[Since (b)(i) > 5 there is no pooling so df = (3 - 1) × (3 - 1) =] 4	B1	1.1b
	(1)	

Notes

B1 for 4 cao

Mark scheme (d)

Scheme	Marks	AO
$\chi^2_4(0.05) = \underline{\underline{9.488}}$	B1ft	3.4/ 1.1b
[The test is significant so reject H_0 :] Sharma's <u>belief</u> is <u>not supported</u> / there is significant evidence of an association between <u>age</u> range and computer <u>game</u> preference.	B1	2.2b
	(2)	
	(6 marks)	

Notes

1st B1ft for awrt 9.488 or ft their df for a 5% critical value

may see $\chi^2_3(0.05) = 7.815$ $\chi^2_2(0.05) = 5.991$ $\chi^2_7(0.05) = 14.067$ $\chi^2_8(0.05) = 15.507$

2nd B1 Indep of hypotheses but dep on cv < 11.542

for correct statement in context about Sharma's belief

condone e.g 'games sold do **not** appeal equally to all age ranges'

for a contextual statement must mention age and game

B0 for contradictory statements e.g. "test is not significant so belief is not supported"

2.

Mark scheme (a)

Scheme	Marks	AO
$x = 4 \times 43 - 47 - 34 - 36 = 55^*$	B1*	3.4
	(1)	

Notes

B1*: Using the uniform model to show the missing observed value eg

$$x = \frac{43 - 0.25 \times (47 + 34 + 36)}{0.25} = 55$$

Mark scheme (b)

Scheme	Marks	AO
$\nu = 4 - 1 = 3$ since the only constraint is that the totals agree	B1	2.4
	(1)	

Notes

B1: $4 - 1 = 3$ (may be in words) and explanation of what the constraint is

Mark scheme (c)

Scheme	Marks	AO
H_0 : The die is unbiased H_1 : The die is biased	B1	2.1
Test Statistic $= \frac{(47 - 43)^2}{43} + \frac{(34 - 43)^2}{43} + \frac{(36 - 43)^2}{43} + \frac{(55 - 43)^2}{43}$	M1	1.1b
$= 6.744 \dots$	A1	1.1b
$\chi^2_{(3,0.05)} = 7.815$	B1	1.1b
Not in the critical region since $7.815 > "6.74 \dots"$ therefore insufficient evidence to reject H_0 Inconclusive test - consistent with the die being unbiased.	A1	3.5a
	(5)	
	(7 marks)	

Notes

B1: Both hypotheses correct. eg The data fits a discrete uniform distribution

M1: Attempting to find $\sum \frac{(O - E)^2}{E}$ or $\sum \frac{O^2}{E} - N$ May be implied by awrt 6.74 or p value of 0.0805...

A1: awrt 6.74 or $\frac{290}{43}$ oe May be implied by p value of 0.0805...

B1: awrt 7.82 (Calc 7.8147...)

A1: Drawing correct inference in context. Need the word die or tetrahedral

3.

Mark scheme (a)

Scheme	Marks	AO
H_0 : No. of car <u>breakdowns</u> per month follows a <u>Poisson</u> distribution H_1 : No. of car <u>breakdowns</u> per month does not follow a <u>Poisson</u> distribution	B1	2.5
	(1)	

Notes

B1 correct hypotheses mentioning “breakdowns” and “Poisson”. Po(1.8) o.e. is B0

Mark scheme (b)

Scheme	Marks	AO
A <u>Poisson</u> distribution will assign some probability to (all) values greater than 5 (o.e.)	B1	2.4
	(1)	

Notes

B1 for a suitable reason mentioning the Poisson distribution taking values beyond 5
e.g. sample space for Poisson being $[0, \infty)$

Mark scheme (c)

Scheme	Marks	AO
Need λ : $\left[\hat{\lambda} = \right] \frac{0 \times 12 + 1 \times 11 + 2 \times 19 + 3 \times 14 + 4 \times 3 + 5 \times 1}{(12 + 11 + 19 + 14 + 3 + 1)}$ <u>or</u> 1.8	M1	1.1b
[Under H_0 $X \sim \text{Po}(1.8)$] $E_1 = 60 \times P(X = 1) = 17.85(227 \dots)$	M1	3.4 1.1b
$E_2 = 60 \times P(X = 2) = 16.06(705 \dots)$	A1	1.1b

$E_{\geq 5} = 60 - \sum_0^4 E_i = 2.18(43996 \dots)$	Blft	
	(4)	

Notes

1st M1 for an expression for the mean with least 3 correct products and correct denominator
M0 if they have found λ by working backwards from $P(X = 0) \Rightarrow 60e^{-\lambda} = 9.92$

2nd M1 for a correct method for finding E_1 or E_2 (implied by one correct value)

1st A1 for awrt 17.85 and awrt 16.07

Blft for awrt 2.18, or using $60 - (\text{sum of their } E_i)$ or $60 \times (1 - P(X \leq 4))$

Mark scheme (d)

Scheme	Marks	AO
$\frac{(11 - "17.85 \dots")^2}{"17.85 \dots"} = 2.6287 \dots \text{ or}$ $\frac{(14 - 9.64)^2}{9.64} = 1.97195 \dots$	M1	1.1b
awrt 2.63 and awrt 1.97	A1	1.1b
	(2)	

Notes

M1 for either correct expression (ft their E_1)

A1 for awrt 2.63 and awrt 1.97

Mark scheme (e)

Scheme	Marks	AO
Need to combine last two columns since E_i are < 5	B1	1.1a
Degrees of freedom therefore $5 - 2$ since mean for Poisson estimated from O_i	B1	1.1a
	(2)	

Notes

1st B1 for explaining need to pool columns since $E_i < 5$

allow candidates describing combining last three columns as long as they say $E_i < 5$

2nd B1 for mentioning mean/rate/parameter/ λ estimated from O_i and 2 constraints

Mark scheme (f)

Scheme	Marks	AO
$\chi^2_3(5\%) = 7.815$	B1	1.1b
(Not significant) insufficient evidence to <u>reject Anja's belief</u>	B1	2.2b
	(2)	
	(12)	

Notes

1st B1 for correct cv of 7.815 (or better)

2nd B1 for correct conclusion in context mentioning “breakdowns” and “Poisson”

Allow equivalent words for belief, such as “theory” etc.

Must be consistent with their cv

Allow $Po(1.8)$ in conclusion instead of Poisson