

QQQ – Further Statistics 1 – Chapter 6 – Chi-Squared Tests (Version 2)

Total Marks: 22

(22 = Platinum, 20 = Gold, 18 = Silver, 15 = Bronze)

1. A leisure club offers a choice of one of three activities to its 150 members on a Tuesday evening. The manager believes that there may be an association between the choice of activity and the age of the member and collected the following data.

Activity Age a years	Badminton	Bowls	Snooker
$a < 20$	9	3	3
$20 \leq a < 40$	10	10	14
$40 \leq a < 50$	16	15	5
$50 \leq a < 60$	15	13	11
$a \geq 60$	4	19	3

(a) Write down suitable hypotheses for a test of the manager's belief. (1)

The manager calculated expected frequencies to use in the test.

(b) Calculate the expected frequency of members aged 60 or over who choose snooker, used by the manager. (1)

(c) Explain why there are 6 degrees of freedom used in this test. (2)

The test statistic used to test the manager's belief is 19.583

(d) Using a 5% level of significance, complete the test of the manager's belief. (2)

(Total 6 marks)

2. A spinner used for a game is designed to give scores with the following probabilities

Score	1	2	3	4	6
Probability	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{5}$	$\frac{1}{10}$

The spinner is spun 80 times and the results are as follows

Score	1	2	3	4	6
Frequency	15	4	12	41	8

Test, at the 10% level of significance, whether or not the spinner is giving scores as it is designed to do. Show your working and state your hypotheses clearly. (7)

(Total 7 marks)

Edexcel AS June 2019 Q2

3. A researcher is investigating the number of female cubs present in litters of size 4

He believes that the number of female cubs in a litter can be modelled by $B(4, 0.5)$

He randomly selects 100 litters each of size 4 and records the number of female cubs.

The results are recorded in the table below.

Number of female cubs	0	1	2	3	4
Observed number of litters	10	33	33	15	9

He calculated the expected frequencies as follows

Number of female cubs	0	1	2	3	4
Expected number of litters	6.25	r	s	r	6.25

(a) Find the value of r and the value of s (3)

(b) Carry out a suitable test, at the 5% level of significance, to determine whether or not the number of female cubs in a litter can be modelled by $B(4, 0.5)$

You should clearly state your hypotheses and the critical value used. (6)

(Total 9 marks)

Edexcel A2 June 2022 Q1

Solutions

1.

Mark scheme (a)

Scheme	Marks	AO
H_0 : There is no association between age and activity H_1 : There is an association between age and activity	B1	1.2
	(1)	

Notes

B1 for both hypotheses in terms of “association” or “independence”. Must mention age and activity (or sport). [Use of “relationship” or “link” here is B0 but allow for last B1ft]

Mark scheme (b)

Scheme	Marks	AO
$\frac{26 \times 36}{150} = 6.24$	B1	1.1b
	(1)	

Notes

B1 for 6.24

Mark scheme (c)

Scheme	Marks	AO
Since expected value in $a < 20$ and snooker = $3.6 < 5$ we amalgamate two rows	B1	2.4
Table is now 4×3 so degrees of freedom is $(4 - 1) \times (3 - 1) = 6 (*)$	B1*	1.1b
	(2)	

Notes

1st B1 for a reason to get a 4 x 3 table based on amalgamation of rows

Must mention $a < 20$ and snooker and see 3.6 and be combining rows (not columns)

2nd B1* for 6 degrees of freedom clearly coming from a 4 x 3 table formed from amalgamation of rows. [8
– 2 = 6 is B0]

Mark scheme (d)

Scheme	Marks	AO
Critical value $\chi^2_6(5\%) = 12.592$	B1	1.1b
[Significant result]: so there is evidence to support manager's belief	B1ft	2.2b
	(2)	
	(6 marks)	

Notes

1st B1 for correct critical value (allow 12.6 or 12.59 or awrt 12.592)

NB p -value = 0.0032839 . . . so allow awrt 0.00328

2nd B1ft for a correct comparison and conclusion (ft their cv) [Independent of hypotheses]

e.g. there is an “association” or “relationship” or “link” between age and activity is OK

BUT there is a “correlation” between age and activity is B0

Do not accept contradictory contextual statements e.g. “manager's belief supported , there is no association between age and activity”

2.

Scheme	Marks	AO																								
H_0 : Spinner is working as designed (o.e.) H_1 : Spinner is not working as designed (o.e.)	B1	1.2																								
<table><tr><td>E_i</td><td>24</td><td>8</td><td>8</td><td>32</td><td>8</td></tr><tr><td>O_i</td><td>15</td><td>4</td><td>12</td><td>41</td><td>8</td></tr><tr><td>$\frac{(O_i - E_i)^2}{E_i}$</td><td>$\frac{81}{24}$</td><td>$\frac{16}{8}$</td><td>$\frac{16}{8}$</td><td>$\frac{81}{32}$</td><td>0</td></tr><tr><td>$\frac{O_i^2}{E_i}$</td><td>$\frac{225}{24}$</td><td>$\frac{16}{8}$</td><td>$\frac{144}{8}$</td><td>$\frac{1681}{32}$</td><td>$\frac{64}{8}$</td></tr></table>	E_i	24	8	8	32	8	O_i	15	4	12	41	8	$\frac{(O_i - E_i)^2}{E_i}$	$\frac{81}{24}$	$\frac{16}{8}$	$\frac{16}{8}$	$\frac{81}{32}$	0	$\frac{O_i^2}{E_i}$	$\frac{225}{24}$	$\frac{16}{8}$	$\frac{144}{8}$	$\frac{1681}{32}$	$\frac{64}{8}$	M1 A1 M1	3.4 1.1b 1.1b
E_i	24	8	8	32	8																					
O_i	15	4	12	41	8																					
$\frac{(O_i - E_i)^2}{E_i}$	$\frac{81}{24}$	$\frac{16}{8}$	$\frac{16}{8}$	$\frac{81}{32}$	0																					
$\frac{O_i^2}{E_i}$	$\frac{225}{24}$	$\frac{16}{8}$	$\frac{144}{8}$	$\frac{1681}{32}$	$\frac{64}{8}$																					

$\sum \frac{(O_i - E_i)^2}{E_i} = 3.375 + 2 + 2 + 2.53125 + 0 = 9.90625$ or $\sum \frac{O_i^2}{E_i} - N = 9.375 + 2 + 18 + 52.53125 + 8 - 80 = 9.90625$	A1	1.1b
$\nu = 5 - 1 = 4$ so $\chi_4^2(10\%)$ cv = 7.779 or better	B1	3.4
Result is significant so there is evidence that the spinner is not operating as designed	A1cso	3.5a
	(7 marks)	

Notes

1st B1 for both hypotheses given in suitable context

1st M1 for using the model to find at least 2 correct expected frequencies

1st A1 for all correct E_i

2nd M1 for attempt to find test statistic (at least two correct expressions, fractions or decimals)

2nd A1 for a correct test statistic (awrt 9.91) [accept $\frac{317}{32}$]

2nd B1 for correct critical value (allow 7.78)

NB p -value = 0.042036... so allow awrt 0.042

3rd A1cso dep on all previous marks for a correct conclusion in context

(can be in terms of model or spinner's design) Must mention spinner **and** scores or design

Accept "spinner is not accurate"

3.

Mark scheme (a)

Scheme	Marks	A0
$r = P(X = 3) \times 100$ or $r = P(X = 1) \times 100$ or $s = P(X = 2) \times 100$	M1	3.4
$r = \underline{25}$ (value may be in table)	A1	1.1b
$s = \underline{37.5}$ (value may be in table)	A1	1.1b
	(3)	

Notes

M1 Using the Binomial model to expected value. Allow if both probs 0.25 and 0.375 seen
May be implied by a correct value of r or s . Alternatives $r = 6.25 \times 4$ or $s = 6.25 \times 6$

1st A1 for $r = 25$

2nd A1 for $s = 37.5$

SC "B1" If M0 scored but their values of r and s satisfy $2r + s = 87.5$ score as **M0A0A1**

Mark scheme (b)

Scheme	Marks	AO												
H_0 : B(4, 0.5) is a suitable model (o.e.) Condone B(0.5, 4) H_1 : B(4, 0.5) is not a suitable model (o.e.)	B1	2.5												
<table><tr><td>$\frac{(O_i - E_i)^2}{E_i}$</td><td>2.25</td><td>2.56</td><td>0.54</td><td>4</td><td>1.21</td></tr><tr><td>$\frac{O_i^2}{E_i}$</td><td>16</td><td>43.56</td><td>29.04</td><td>9</td><td>12.96</td></tr></table>	$\frac{(O_i - E_i)^2}{E_i}$	2.25	2.56	0.54	4	1.21	$\frac{O_i^2}{E_i}$	16	43.56	29.04	9	12.96	M1	1.1b
$\frac{(O_i - E_i)^2}{E_i}$	2.25	2.56	0.54	4	1.21									
$\frac{O_i^2}{E_i}$	16	43.56	29.04	9	12.96									
$\sum \frac{(O_i - E_i)^2}{E_i} = 10.56$ or $\sum \frac{O_i^2}{E_i} - N = 110.56 - 100 = 10.56 \left(= \frac{264}{25} \right)$	A1	1.1b												
$\nu = 5 - 1 = 4$	B1	1.1b												
CV = 9.488 (Calc 9.487729035...)	B1ft	1.1b												
Significant so there is evidence that the researcher's model is not suitable	A1	2.2b												
	(6)													
	Total 9													

Notes

1st B1 Both hypotheses correct using the correct notation in at least one or written in full e.g. binomial with $n = 4$ and $p = 0.5$

M1 Calculating either $\frac{(O_i - E_i)^2}{E_i}$ or $\frac{O_i^2}{E_i}$ at least 4 correct. Implied by sight of awrt 10.6

1st A1 Allow 10.6 (from correct working)

2nd B1 Correct dof May be implied by CV of 9.48 or 9.49 or better

3rd B1ft For 9.488 or better. Can fit their dof NB $\chi^2_3(5\%) = 7.815$ (allow awrt 7.815)

2nd A1 Indep' of hypotheses but dep on 1st A1

Evaluating the outcome by drawing a correct inference.

Compatible with comparison of 10.56 or 10.6 with their CV (which must be > 1)

They must say **model not suitable** (o.e.)

They do not need to state the comparison or say reject H_0 etc

No need to explicitly see $B(4, 0.5)$ mentioned here