2.3 Further Statistics: Chi-Squared Tests (Goodness of Fit and Contingency Tables): Test Pack

Calculator allowed · about 60 minutes

Total Marks

Name: _____ Date: ____ / ____ / _____

Answer ALL questions. Show all your working.

Key facts: chi-squared tests

Original content written for Revision Library.

A chi-squared test compares observed frequencies O with the expected frequencies E predicted by a proposed model, using the test statistic X^2 = sum of (O - E)^2 / E, calculated over all classes (or cells) after any pooling. For a goodness-of-fit test to a fully specified distribution (for example a uniform distribution, or given proportions), degrees of freedom = (number of classes) - 1. If one or more parameters of the model (for example p for a binomial, or the mean for a Poisson) are estimated from the same sample used in the test, one further degree of freedom is lost for each parameter estimated. For an r by c contingency table testing independence between two variables, expected frequencies are E = (row total * column total) / grand total, and degrees of freedom = (r - 1)(c - 1). Classes with an expected frequency below 5 should be pooled (combined with a neighbouring class) before the test statistic is calculated, reducing the number of classes, and hence the degrees of freedom, accordingly. For any test with 1 degree of freedom, including every 2 by 2 contingency table, Yates' continuity correction is applied: X^2 = sum of (|O - E| - 0.5)^2 / E. The calculated test statistic is compared with the critical value from chi-squared tables at the chosen significance level and the stated degrees of freedom; if the test statistic exceeds the critical value, the null hypothesis is rejected.

1 For each scenario below, state the number of degrees of freedom that should be used in the chi-squared test described, giving a brief reason for your answer. You are not required to state hypotheses or carry out any calculations.

(a) A six-sided spinner is spun 120 times to test whether each of the six outcomes is equally likely. **(2)**

(b) The number of accidents per week at a road junction is recorded for 80 weeks, to test whether it follows a Poisson distribution, with the mean estimated from the same sample. After pooling classes with small expected frequencies, 6 classes remain in total. **(2)**

(c) A 4 by 3 contingency table is used to test whether a person's favourite sport is associated with their year group. **(2)**

(d) The number of faulty items in independent samples of 8 items is tested against a binomial model B(8, p), where p is estimated from the data. After pooling classes with small expected frequencies, 5 classes remain in total. **(2)**

(Total for Question 1 is 8 marks)

© Revision Library | revisionlibrary.co.uk

2 At a summer fair in Bristol, a stallholder uses a spinner with six equal sectors, numbered 1 to 6, for a prize game. Some visitors suspect the spinner is biased. Over the course of the day the spinner is spun 180 times, with the results recorded below.

Score 1: 25 spins. Score 2: 34 spins. Score 3: 28 spins. Score 4: 32 spins. Score 5: 22 spins. Score 6: 39 spins.

(a) State suitable null and alternative hypotheses for a test of whether the spinner is fair. **(2)**

(b) State the expected frequency for each score under H0, and calculate the value of the chi-squared test statistic. **(3)**

(c) State the number of degrees of freedom and the critical value at the 5% significance level, and hence test, at the 5% level, whether the spinner is fair. State your conclusion clearly, in context. **(3)**

(Total for Question 2 is 8 marks)

© Revision Library | revisionlibrary.co.uk

3 National figures suggest that a randomly chosen child is equally likely to be a boy or a girl, independently of other children in the family. Amir, a statistics student, surveys 200 randomly chosen households in Leeds that each have exactly 4 children, and records X, the number of boys in each household. His results are shown below.

Number of boys (X): 0, 1, 2, 3, 4. Frequency: 10, 54, 80, 46, 10.

Number of boys (X)	Frequency
0	10
1	54
2	80
3	46
4	10

(a) Using the national figure that a child is equally likely to be a boy or a girl, state suitable hypotheses to test whether X follows a binomial distribution. **(2)**

(b) Show that P(X=2) = 0.375, and hence state the expected frequency of each value of X in a sample of 200 households. **(3)**

(c) Calculate the value of the chi-squared test statistic. **(2)**

(d) Explain why this test has 4 degrees of freedom, state the critical value at the 5% significance level, and test whether the binomial model is a good fit. State your conclusion clearly, in context. **(3)**

(Total for Question 3 is 10 marks)

© Revision Library | revisionlibrary.co.uk

4 When carrying out a chi-squared test, classes with a small expected frequency are sometimes combined ('pooled') with a neighbouring class before the test statistic is calculated.

(a) State the usual condition used to decide whether an expected frequency is too small and pooling is required. **(1)**

(b) Explain the effect that pooling two classes into one has on the number of degrees of freedom used in the test. **(1)**

(c) A goodness-of-fit test for a Poisson distribution initially uses 7 classes. After pooling classes with expected frequency below 5, 5 classes remain in total. The mean of the Poisson distribution was estimated from the sample data. State the number of degrees of freedom used in the test. **(2)**

(d) State two conditions, other than the expected frequency condition, that should hold for a chi-squared test to be valid. **(2)**

(Total for Question 4 is 6 marks)

© Revision Library | revisionlibrary.co.uk

5 A college in Manchester surveys 240 A-level students about their preferred revision method (flashcards, past papers or online videos), grouped by subject area (STEM, Humanities or Arts). Some results are shown in the table below, where a and b represent two missing values.

	Flashcards	Past papers	Online videos	Row total
STEM	24	56	a	120
Humanities	18	30	12	60
Arts	b	14	28	60
Column totals	60	100	80	240

(a) Find the values of a and b. **(2)**

(b) State suitable null and alternative hypotheses for a test of association between subject area and preferred revision method. **(2)**

(c) Show that the expected frequency for Humanities students who prefer online videos is 20, and complete the expected frequency table. **(3)**

(d) Calculate the chi-squared test statistic. **(3)**

(e) State the degrees of freedom, and the critical values at the 5% and 1% significance levels. Hence state, with a reason, the strongest conclusion that can be drawn from this test. **(3)**

(f) By comparing observed and expected frequencies, identify the subject area and revision method combination that contributes most to the value of chi-squared, and briefly describe how its observed frequency differs from what would be expected under independence. **(2)**

(Total for Question 5 is 15 marks)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 2.3 Further Statistics: Chi-Squared Tests (Goodness of Fit and Contingency Tables): Test Pack

Question 1

(a) recognises 6 categories, fully specified with no parameters estimated from the data **B1**

(a) df = 6 - 1 = 5, cao **B1**

(a) Answer: df = 5

(b) recognises 6 classes remain after pooling, and 1 parameter (the mean) has been estimated from the data **B1**

(b) df = 6 - 1 - 1 = 4, cao **B1**

(b) Answer: df = 4

(c) recognises the table has r = 4 rows and c = 3 columns **B1**

(c) df = (4 - 1)(3 - 1) = 6, cao **B1**

(c) Answer: df = 6

(d) recognises 5 classes remain after pooling, and 1 parameter (p) has been estimated from the data **B1**

(d) df = 5 - 1 - 1 = 3, cao **B1**

(d) Answer: df = 3

Question 2

(a) H0: the spinner is fair, i.e. P(X=x) = 1/6 for x = 1, ..., 6 **B1**

(a) H1: the spinner is not fair, i.e. the probabilities are not all equal to 1/6 **B1**

(a) Answer: H0: the spinner is fair, P(X=x) = 1/6 for each score. H1: the spinner is not fair.

(b) E = 30 for each score **B1**

(b) attempts sum of $(O-E)^2/E$ across all 6 scores **M1**

(b) cao, awrt 6.47 **A1**

(b) Answer: Test statistic = 194/30 = awrt 6.47

(c) df = 5 **B1**

(c) critical value = 11.070 **B1**

(c) valid conclusion ft: since 6.47 < 11.070, do not reject H0; insufficient evidence that the spinner is biased **B1**

(c) Answer: df = 5, critical value = 11.070; 6.47 < 11.070 so do not reject H0 - no significant evidence the spinner is biased.

Question 3

(a) H0: X ~ B(4, 0.5) **B1**

(a) H1: X does not follow a B(4, 0.5) distribution **B1**

(a) Answer: H0: X ~ B(4, 0.5). H1: X does not follow a B(4, 0.5) distribution.

(b) correct use of $P(X=2) = C(4,2) * 0.5^2 * 0.5^2$ **M1**

(b) 0.375 cso **A1**

(b) complete expected frequency table: 12.5, 50, 75, 50, 12.5 **B1**

(b) Answer: P(X=2) = 0.375; expected frequencies 12.5, 50, 75, 50, 12.5.

(c) attempts sum of $(O-E)^2/E$ across all 5 classes **M1**

(c) cao, awrt 1.97 **A1**

(c) Answer: Test statistic = awrt 1.97

(d) p = 0.5 was given, not estimated from the sample, so df = (number of classes) - 1 = 5 - 1 = 4 **B1**

(d) critical value = 9.488 **B1**

(d) valid conclusion ft: since 1.97 < 9.488, do not reject H0; the B(4, 0.5) model is a good fit for the data **B1**

(d) Answer: df = 4 (p not estimated); critical value = 9.488; 1.97 < 9.488 so do not reject H0.

Question 4

(a) expected frequency less than 5, oe **B1**

(a) Answer: E < 5

(b) each time two classes are combined into one, the number of classes (and hence the degrees of freedom) is reduced by 1, oe **B1**

(b) Answer: Each pooling of two classes into one reduces the degrees of freedom by 1.

(c) df = (number of classes after pooling) - 1 - (number of estimated parameters) = 5 - 1 - 1 **M1**

(c) df = 3, cao **A1**

(c) Answer: df = 3

(d) any one valid condition, e.g. observations must be independent of each other **B1**

(d) a second valid condition, e.g. the data must come from a random sample and be recorded as frequencies/counts (not measurements), oe **B1**

(d) Answer: Any two of: observations are independent; the sample is random; data are frequencies/counts, not measurements; the total number of observations is fixed.

© Revision Library | revisionlibrary.co.uk

Question 5

(a) a = 40 **B1**

(a) b = 18 **B1**

(a) Answer: a = 40, b = 18

(b) H0: subject area and preferred revision method are independent **B1**

(b) H1: subject area and preferred revision method are associated (not independent) **B1**

(b) Answer: H0: subject area and revision method are independent. H1: they are associated.

(c) uses E = (row total * column total) / grand total = 60*80/240 **M1**

(c) 20 cso **A1**

(c) remaining 8 values correct: STEM 30, 50, 40; Humanities 15, 25; Arts 15, 25, 20 **A1**

(c) Answer: Expected frequencies - STEM: 30, 50, 40. Humanities: 15, 25, 20. Arts: 15, 25, 20.

(d) attempts sum of $(O-E)^2/E$ across all 9 cells **M1**

(d) at least 6 correct individual terms (unsimplified) **A1**

(d) cao, 15.36 **A1**

(d) Answer: Test statistic = 15.36

(e) df = (3-1)(3-1) = 4 **B1**

(e) critical values 9.488 (5%) and 13.277 (1%) **B1**

(e) valid conclusion: 15.36 exceeds both critical values, so reject H0 even at the 1% level; strong evidence of association between subject area and preferred revision method **B1**

(e) Answer: df = 4; critical values 9.488 (5%), 13.277 (1%); reject H0 at the 1% level.

(f) identifies Arts students and past papers (the largest single contribution, 4.84) **B1**

(f) correct description: Arts students use past papers far less than expected (14, compared with an expected frequency of 25) **B1**

(f) Answer: Arts and past papers contributes most; Arts students use past papers much less than expected (14 vs 25).

© Revision Library | revisionlibrary.co.uk