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

Calculator allowed · about 225 minutes · Write-in edition

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** A researcher counts the number of events per interval over 96 intervals: 0 events 27 times, 1 event 36 times, 2 events 23 times, and 3 or more events 10 times. It is suggested that the number of events per interval follows a Poisson distribution. Taking the mean of the data to be 1.167 (treating '3 or more' as 3), test the suggestion at the 5% level.

**(a)** Calculate the expected frequencies for a Poisson distribution with the given mean, using the categories 0, 1, 2 and 3 or more. **(3)**

**(b)** Calculate the test statistic and the degrees of freedom, and state the conclusion at the 5% level of significance, given that the critical value is 5.991. **(4)**

**(Total for Question 1 is 7 marks)**

**2** A researcher records the results of 190 people classified by group and by outcome, as shown. Rows: Group X: Grade A 44, Grade B 23, Grade C 43; Group Y: Grade A 26, Grade B 14, Grade C 40. A test is carried out at the 5% level to see whether the outcome is independent of the group.

| | Grade A | Grade B | Grade C |
|---|---|---|---|
| Group X | 44 | 23 | 43 |
| Group Y | 26 | 14 | 40 |

**(a)** State the hypotheses, and calculate the expected frequency for the first cell in each row. **(3)**

**(b)** Find the value of the test statistic, to 3 significant figures, and the number of degrees of freedom. **(3)**

**(c)** Given that the critical value is 5.991, state the conclusion of the test in context. **(3)**

**(Total for Question 2 is 9 marks)**

**3** A researcher counts the number of events per interval over 157 intervals: 0 events 46 times, 1 event 57 times, 2 events 35 times, and 3 or more events 19 times. It is suggested that the number of events per interval follows a Poisson distribution. Taking the mean of the data to be 1.172 (treating '3 or more' as 3), test the suggestion at the 5% level.

**(a)** Calculate the expected frequencies for a Poisson distribution with the given mean, using the categories 0, 1, 2 and 3 or more. **(3)**

**(b)** Calculate the test statistic and the degrees of freedom, and state the conclusion at the 5% level of significance, given that the critical value is 5.991. **(4)**

**(Total for Question 3 is 7 marks)**

**4** A researcher records the results of 157 people classified by group and by outcome, as shown. Rows: Group X: Grade A 44, Grade B 23, Grade C 13; Group Y: Grade A 13, Grade B 26, Grade C 38. A test is carried out at the 5% level to see whether the outcome is independent of the group.

| | Grade A | Grade B | Grade C |
|---|---|---|---|
| Group X | 44 | 23 | 13 |
| Group Y | 13 | 26 | 38 |

**(a)** State the hypotheses, and calculate the expected frequency for the first cell in each row. **(3)**

**(b)** Find the value of the test statistic, to 3 significant figures, and the number of degrees of freedom. **(3)**

**(c)** Given that the critical value is 5.991, state the conclusion of the test in context. **(3)**

**(Total for Question 4 is 9 marks)**

**5** A researcher counts the number of events per interval over 85 intervals: 0 events 15 times, 1 event 29 times, 2 events 23 times, and 3 or more events 18 times. It is suggested that the number of events per interval follows a Poisson distribution. Taking the mean of the data to be 1.518 (treating '3 or more' as 3), test the suggestion at the 5% level.

**(a)** Calculate the expected frequencies for a Poisson distribution with the given mean, using the categories 0, 1, 2 and 3 or more. **(3)**

**(b)** Calculate the test statistic and the degrees of freedom, and state the conclusion at the 5% level of significance, given that the critical value is 5.991. **(4)**

**(Total for Question 5 is 7 marks)**

**6** A researcher counts the number of events per interval over 146 intervals: 0 events 54 times, 1 event 56 times, 2 events 27 times, and 3 or more events 9 times. It is suggested that the number of events per interval follows a Poisson distribution. Taking the mean of the data to be 0.938 (treating '3 or more' as 3), test the suggestion at the 5% level.

**(a)** Calculate the expected frequencies for a Poisson distribution with the given mean, using the categories 0, 1, 2 and 3 or more. **(3)**

**(b)** Calculate the test statistic and the degrees of freedom, and state the conclusion at the 5% level of significance, given that the critical value is 5.991. **(4)**

**(Total for Question 6 is 7 marks)**

**7** A market researcher classifies the answers to a survey question into 4 categories. In a random sample of 123 observations the frequencies in the categories are 22, 39, 36, 26. A model predicts that the categories occur in the ratio 2:3:3:2. A test at the 5% significance level is to be carried out to see whether the model fits.

**(a)** State the hypotheses for the test and find the expected frequency for each category. **(3)**

**(b)** Calculate the test statistic, to 3 significant figures, and state the number of degrees of freedom. **(3)**

**(c)** The critical value of chi$^2$ for 3 degrees of freedom at the 5% level is 7.815. State your conclusion in context. **(3)**

**(Total for Question 7 is 9 marks)**

**8** A market researcher classifies the answers to a survey question into 4 categories. In a random sample of 58 observations the frequencies in the categories are 19, 18, 16, 5. A model predicts that the categories occur in the ratio 4:3:2:1. A test at the 1% significance level is to be carried out to see whether the model fits.

**(a)** State the hypotheses for the test and find the expected frequency for each category. **(3)**

**(b)** Calculate the test statistic, to 3 significant figures, and state the number of degrees of freedom. **(3)**

**(c)** The critical value of $chi^2$ for 3 degrees of freedom at the 1% level is 11.345. State your conclusion in context. **(3)**

**(Total for Question 8 is 9 marks)**

**9** A researcher records the results of 206 people classified by group and by outcome, as shown. Rows: Morning class: Pass 41, Merit 15, Distinction 37; Afternoon class: Pass 48, Merit 40, Distinction 25. A test is carried out at the 1% level to see whether the outcome is independent of the group.

| | Pass | Merit | Distinction |
|---|---|---|---|
| Morning class | 41 | 15 | 37 |
| Afternoon class | 48 | 40 | 25 |

**(a)** State the hypotheses, and calculate the expected frequency for the first cell in each row. **(3)**

**(b)** Find the value of the test statistic, to 3 significant figures, and the number of degrees of freedom. **(3)**

**(c)** Given that the critical value is 9.21, state the conclusion of the test in context. **(3)**

**(Total for Question 9 is 9 marks)**

**10** A teacher classifies the results of a class test into 4 categories. In a random sample of 104 observations the frequencies in the categories are 20, 33, 27, 24. A model predicts that the categories occur in the ratio 2:3:3:2. A test at the 5% significance level is to be carried out to see whether the model fits.

**(a)** State the hypotheses for the test and find the expected frequency for each category. **(3)**

**(b)** Calculate the test statistic, to 3 significant figures, and state the number of degrees of freedom. **(3)**

**(c)** The critical value of $chi^2$ for 3 degrees of freedom at the 5% level is 7.815. State your conclusion in context. **(3)**

**(Total for Question 10 is 9 marks)**

**11** A bakery in York claims that its daily sales split across four product types, cakes, pastries, bread and biscuits, in the ratio 4:3:2:1. On a randomly chosen day, 200 items are sold, with the following breakdown.

Cakes: 70. Pastries: 64. Bread: 44. Biscuits: 22.

**(a)** State suitable null and alternative hypotheses to test the bakery's claim. **(2)**

**(b)** Find the expected number of each product type sold, based on the claimed ratio. **(2)**

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

**(d)** State the number of degrees of freedom, explaining your answer, and the critical value at the 5% significance level. Hence test the bakery's claim, stating your conclusion clearly, in context. **(3)**

**(Total for Question 11 is 9 marks)**

**12** A cafe in Norwich records the number of coffees sold in each of four two-hour periods during a working day, over a sample of 180 coffees sold in total, to test whether sales are uniformly distributed across the four periods.

Period 1 (8am-10am): 38. Period 2 (10am-12pm): 52. Period 3 (12pm-2pm): 61. Period 4 (2pm-4pm): 29.

Critical values for chi-squared with 3 degrees of freedom: 10% level 6.251; 5% level 7.815; 2.5% level 9.348; 1% level 11.345; 0.1% level 16.266.

**(a)** State suitable null and alternative hypotheses for this test. **(2)**

**(b)** State the expected number of coffees sold in each period under H0, and calculate the chi-squared test statistic. **(3)**

**(c)** Using the critical values given, state the range within which the p-value of this test lies, justifying your answer. **(2)**

**(d)** Hence state your conclusion at the 1% significance level, in the context of this question. **(2)**

**(Total for Question 12 is 9 marks)**

**13** A GP surgery in Cardiff investigates whether there is an association between having a flu vaccine and catching flu during the winter, among 160 randomly chosen patients. The results are shown in the table below.

| | Caught flu | Did not catch flu | Row total |
|---|---|---|---|
| Vaccinated | 18 | 72 | 90 |
| Not vaccinated | 22 | 48 | 70 |
| Column totals | 40 | 120 | 160 |

**(a)** State suitable null and alternative hypotheses for this test. **(2)**

**(b)** Show that the expected frequency for vaccinated patients who caught flu is 22.5, and complete the expected frequency table. **(3)**

**(c)** Explain why Yates' continuity correction should be used in this test, and calculate the value of chi-squared with the correction applied. **(5)**

**(d)** State the number of degrees of freedom and the critical value at the 5% significance level, and test whether there is an association between vaccination and catching flu. State your conclusion clearly, in context. **(3)**

**(Total for Question 13 is 13 marks)**

**14** A students' union at a university in Sheffield investigates whether attending an exam workshop is associated with passing a resit exam, among 120 students. The results are shown below.

Attended workshop: passed resit 42, failed resit 18, row total 60.
Did not attend: passed resit 24, failed resit 36, row total 60.
Column totals: passed resit 66, failed resit 54, grand total 120.

| | Passed resit | Failed resit | Total |
|---|---|---|---|
| Attended workshop | 42 | 18 | 60 |
| Did not attend | 24 | 36 | 60 |
| Total | 66 | 54 | 120 |

**(a)** State suitable null and alternative hypotheses for this test. **(2)**

**(b)** Show that the expected frequency for students who attended the workshop and passed the resit is 33, and complete the expected frequency table. **(3)**

**(c)** Calculate the chi-squared test statistic, applying Yates' continuity correction. **(3)**

**(d)** State the number of degrees of freedom, and determine the lowest of the standard significance levels (10%, 5%, 1%) at which the null hypothesis would be rejected. **(3)**

**(e)** State your conclusion in context, and give one reason why this result does not necessarily mean that attending the workshop causes students to pass the resit. **(2)**

**(Total for Question 14 is 13 marks)**

**15** A factory inspects lightbulbs in boxes of 5 for defects. Over a period of production, 180 boxes are sampled, and X, the number of defective bulbs in each box, is recorded. The total number of defective bulbs found across all 180 boxes is 150. The factory wants to test whether X follows a binomial distribution B(5, p), with p estimated from the data.

Number of defective bulbs (X): 0, 1, 2, 3, 4, 5. Frequency: 77, 67, 27, 7, 2, 0.

**(a)** State suitable null and alternative hypotheses for this test. **(2)**

**(b)** Show that an estimate for p, the probability that a bulb is defective, is 1/6, and find P(X=2), giving your answer to 4 decimal places. **(4)**

**(c)** Using p = 1/6, find the expected frequency for each value of X, giving each to 3 decimal places. **(2)**

**(d)** Explain why pooling is required, and give the pooled observed and expected frequency tables. **(3)**

**(e)** Calculate the chi-squared test statistic for the pooled data. **(2)**

**(f)** State the number of degrees of freedom, explaining your reasoning, 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 15 is 16 marks)**

**16** A quality inspector at a car windscreen manufacturer records X, the number of scratches found on each of a sample of 100 windscreens, to test whether X can be modelled by a Poisson distribution, with the mean estimated from the data.

Number of scratches (X): 0, 1, 2, 3, 4, 5. Frequency: 22, 34, 25, 12, 5, 2.

No windscreen in the sample had more than 5 scratches, but since a Poisson distribution allows any non-negative integer value, the final class should be treated as '$X \geq 5$' when finding expected frequencies from the model.

**(a)** Calculate the mean number of scratches per windscreen for this sample. **(2)**

**(b)** Using this mean, find P(X=3) for a Poisson distribution, giving your answer to 4 decimal places, and state the expected frequency for X=3 in the sample of 100 windscreens. **(3)**

**(c)** The expected frequencies for X = 0, 1, 2 and 4 are 22.313, 33.470, 25.102 and 4.707 respectively (to 3 dp), and P(X≥5) gives an expected frequency of 1.858. Explain why pooling is required, and give the pooled observed and expected frequency tables. **(3)**

**(d)** Calculate the chi-squared test statistic for the pooled data. **(2)**

**(e)** State the number of degrees of freedom, explaining your reasoning, and the critical value at the 5% significance level. Hence test whether a Poisson distribution is a suitable model, stating your conclusion clearly, in context. **(3)**

**(f)** The value of the test statistic found in part (d) is very small. Comment briefly on what this suggests about the fit of the Poisson model, and suggest one reason, other than chance, why an unusually small chi-squared value might sometimes arise in practice. **(2)**

**(Total for Question 16 is 15 marks)**

**17** A factory in Coventry making phone cases records the type of defect found on faulty items over two shifts, Day and Night, across 200 faulty items in total. Defects are classified as Scratch, Crack, Discolour or Misalign.

Day: Scratch 34, Crack 28, Discolour 16, Misalign 22, row total 100.
Night: Scratch 25, Crack 31, Discolour 25, Misalign 19, row total 100.
Column totals: Scratch 59, Crack 59, Discolour 41, Misalign 41, grand total 200.

**(a)** State suitable null and alternative hypotheses to test whether defect type is associated with shift. **(2)**

**(b)** Show that the expected frequency for Day-shift Scratch defects is 29.5, and complete the expected frequency table. **(3)**

**(c)** Calculate the chi-squared test statistic for this test of association. **(2)**

**(d)** State the number of degrees of freedom and the critical value at the 5% significance level, and test whether defect type is associated with shift. State your conclusion clearly, in context. **(3)**

**(e)** The QA manager separately claims that, overall (combining both shifts), the four defect types should occur in the ratio 3:3:2:2 (Scratch:Crack:Discolour:Misalign). Using the column totals from the table, test this claim at the 5% significance level. State your hypotheses, expected frequencies, test statistic, degrees of freedom and conclusion. **(6)**

**(f)** Explain why the result of the test in part (e) cannot, by itself, tell us whether defect type is associated with shift. **(2)**

**(Total for Question 17 is 18 marks)**

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

## Question 1

(a) $P(X = j) = e^{-1.167}\ 1.167^j/j!$ for j = 0, 1, 2 **M1**

(a) $P(X \geq 3) = 1 - P(X \leq 2)$ **M1**

(a) expected: 29.88, 34.88, 20.35, 10.89 **A1**

**(a) Answer: 29.88, 34.88, 20.35, 10.89**

(b) $chi^2$ = sum of $(O - E)^2/E$ **M1**

(b) 0.733 **A1**

(b) degrees of freedom = 4 - 1 - 1 = 2 (the mean was estimated from the data) **B1**

(b) do not reject H0: the Poisson model is a suitable fit **A1**

**(b) Answer: $chi^2$ = 0.733, nu = 2, do not reject H0**

## Question 2

(a) H0: outcome and group are independent; H1: they are associated **B1**

(a) expected = (row total x column total)/grand total **M1**

(a) $E_{11}$ = 770/19 = 40.53, $E_{21}$ = 560/19 = 29.47 **A1**

**(a) Answer: $E_{11}$ = 40.53, $E_{21}$ = 29.47**

(b) sum of $(O - E)^2/E$ over all six cells **M1**

(b) $chi^2$ = 2.25 **A1**

(b) degrees of freedom = (2 - 1)(3 - 1) = 2 **B1**

**(b) Answer: $chi^2$ = 2.25, nu = 2**

(c) compares 2.25 with 5.991 **M1**

(c) do not reject H0 **A1**

(c) there is no evidence of an association between group and outcome **B1**

**(c) Answer: do not reject H0**

## Question 3

(a) $P(X = j) = e^{-1.172}\ 1.172^j/j!$ for j = 0, 1, 2 **M1**

(a) $P(X \geq 3) = 1 - P(X \leq 2)$ **M1**

(a) expected: 48.63, 56.99, 33.4, 17.98 **A1**

**(a) Answer: 48.63, 56.99, 33.4, 17.98**

(b) $chi^2$ = sum of $(O - E)^2/E$ **M1**

(b) 0.277 **A1**

(b) degrees of freedom = 4 - 1 - 1 = 2 (the mean was estimated from the data) **B1**

(b) do not reject H0: the Poisson model is a suitable fit **A1**

**(b) Answer: $chi^2$ = 0.277, nu = 2, do not reject H0**

## Question 4

(a) H0: outcome and group are independent; H1: they are associated **B1**

(a) expected = (row total x column total)/grand total **M1**

(a) $E_{11}$ = 4560/157 = 29.04, $E_{21}$ = 4389/157 = 27.96 **A1**

**(a) Answer: $E_{11}$ = 29.04, $E_{21}$ = 27.96**

(b) sum of $(O - E)^2/E$ over all six cells **M1**

(b) $chi^2$ = 29.3 **A1**

(b) degrees of freedom = (2 - 1)(3 - 1) = 2 **B1**

**(b) Answer: $chi^2$ = 29.3, nu = 2**

(c) compares 29.3 with 5.991 **M1**

(c) reject H0 **A1**

(c) there is evidence of an association between group and outcome **B1**

**(c) Answer: reject H0**

## Question 5

(a) $P(X = j) = e^{-1.518}\ 1.518^j/j!$ for j = 0, 1, 2 **M1**

(a) $P(X \geq 3) = 1 - P(X \leq 2)$ **M1**

(a) expected: 18.63, 28.28, 21.46, 16.63 **A1**

**(a) Answer: 18.63, 28.28, 21.46, 16.63**

(b) $chi^2$ = sum of $(O - E)^2/E$ **M1**

(b) 0.947 **A1**

(b) degrees of freedom = 4 - 1 - 1 = 2 (the mean was estimated from the data) **B1**

(b) do not reject H0: the Poisson model is a suitable fit **A1**

**(b) Answer: $chi^2$ = 0.947, nu = 2, do not reject H0**

## Question 6

(a) $P(X = j) = e^{-0.938}\ 0.938^j/j!$ for j = 0, 1, 2 **M1**

(a) $P(X \geq 3) = 1 - P(X \leq 2)$ **M1**

(a) expected: 57.15, 53.6, 25.14, 10.11 **A1**

**(a) Answer: 57.15, 53.6, 25.14, 10.11**

(b) $chi^2$ = sum of $(O - E)^2/E$ **M1**

(b) 0.540 **A1**

(b) degrees of freedom = 4 - 1 - 1 = 2 (the mean was estimated from the data) **B1**

(b) do not reject H0: the Poisson model is a suitable fit **A1**

**(b) Answer: $chi^2$ = 0.540, nu = 2, do not reject H0**

## Question 7

(a) H0: the model (ratio 2:3:3:2) is a suitable fit; H1: it is not **B1**

(a) expected = N x (ratio part)/10 **M1**

(a) expected: 24.6, 36.9, 36.9, 24.6 **A1**

**(a) Answer: 24.6, 36.9, 36.9, 24.6**

(b) sum of $(O - E)^2/E$ **M1**

(b) $chi^2 = 0.496$ **A1**

(b) degrees of freedom = 4 - 1 = 3 **B1**

**(b) Answer: $chi^2 = 0.496$, nu = 3**

(c) compares 0.496 with 7.815 **M1**

(c) do not reject H0 **A1**

(c) there is no evidence against the model: the data are consistent with the ratio **B1**

**(c) Answer: do not reject H0**

## Question 8

(a) H0: the model (ratio 4:3:2:1) is a suitable fit; H1: it is not **B1**

(a) expected = N x (ratio part)/10 **M1**

(a) expected: 23.2, 17.4, 11.6, 5.8 **A1**

**(a) Answer: 23.2, 17.4, 11.6, 5.8**

(b) sum of $(O - E)^2/E$ **M1**

(b) $chi^2 = 2.56$ **A1**

(b) degrees of freedom = 4 - 1 = 3 **B1**

**(b) Answer: $chi^2 = 2.56$, nu = 3**

(c) compares 2.56 with 11.345 **M1**

(c) do not reject H0 **A1**

(c) there is no evidence against the model: the data are consistent with the ratio **B1**

**(c) Answer: do not reject H0**

## Question 9

(a) H0: outcome and group are independent; H1: they are associated **B1**

(a) expected = (row total x column total)/grand total **M1**

(a) $E_{11} = 8277/206 = 40.18$, $E_{21} = 10057/206 = 48.82$ **A1**

**(a) Answer: $E_{11} = 40.18$, $E_{21} = 48.82$**

(b) sum of $(O - E)^2/E$ over all six cells **M1**

(b) $chi^2 = 12.4$ **A1**

(b) degrees of freedom = (2 - 1)(3 - 1) = 2 **B1**

**(b) Answer: $chi^2 = 12.4$, nu = 2**

(c) compares 12.4 with 9.21 **M1**

(c) reject H0 **A1**

(c) there is evidence of an association between group and outcome **B1**

**(c) Answer: reject H0**

## Question 10

(a) H0: the model (ratio 2:3:3:2) is a suitable fit; H1: it is not **B1**

(a) expected = N x (ratio part)/10 **M1**

(a) expected: 20.8, 31.2, 31.2, 20.8 **A1**

**(a) Answer: 20.8, 31.2, 31.2, 20.8**

(b) sum of $(O - E)^2/E$ **M1**

(b) $chi^2 = 1.19$ **A1**

(b) degrees of freedom = 4 - 1 = 3 **B1**

**(b) Answer: $chi^2 = 1.19$, nu = 3**

(c) compares 1.19 with 7.815 **M1**

(c) do not reject H0 **A1**

(c) there is no evidence against the model: the data are consistent with the ratio **B1**

**(c) Answer: do not reject H0**

## Question 11

(a) H0: sales occur in the ratio 4:3:2:1, as claimed **B1**

(a) H1: sales do not occur in the ratio 4:3:2:1 **B1**

**(a) Answer: H0: sales occur in the ratio 4:3:2:1. H1: they do not.**

(b) converts the ratio 4:3:2:1 to proportions 0.4, 0.3, 0.2, 0.1 and multiplies by 200 **M1**

(b) all four correct: 80, 60, 40, 20 **A1**

**(b) Answer: Expected frequencies: 80, 60, 40, 20.**

(c) attempts sum of $(O-E)^2/E$ across all 4 classes **M1**

(c) cao, awrt 2.12 **A1**

**(c) Answer: Test statistic = awrt 2.12**

(d) the ratio was fully specified (not estimated from the sample), so df = 4 - 1 = 3 **B1**

(d) critical value = 7.815 **B1**

(d) valid conclusion ft: since 2.12 < 7.815, do not reject H0; the data are consistent with the claimed ratio **B1**

**(d) Answer: df = 3; critical value = 7.815; 2.12 < 7.815 so do not reject H0.**

## Question 12

(a) H0: coffee sales are uniformly distributed across the four periods **B1**

(a) H1: sales are not uniformly distributed across the four periods **B1**

**(a) Answer: H0: sales are uniform across the four periods. H1: they are not.**

(b) E = 45 for each period **B1**

(b) attempts sum of $(O-E)^2/E$ across all 4 periods **M1**

(b) cao, awrt 13.6 **A1**

**(b) Answer: Test statistic = 610/45 = awrt 13.6**

(c) compares 13.56 with the 1% (11.345) and 0.1% (16.266) critical values, showing 13.56 exceeds 11.345 but is less than 16.266 **M1**

(c) correct bracket: $0.001 < p < 0.01$ **A1**

**(c) Answer: $0.001 < p < 0.01$**

(d) 13.56 > 11.345, so reject H0 at the 1% level **B1**

(d) conclusion in context: there is significant evidence that coffee sales are not uniformly distributed across the four periods **B1**

**(d) Answer: Reject H0 at the 1% level: sales are not uniform across the four periods.**

## Question 13

(a) H0: whether a patient catches flu is independent of whether they were vaccinated **B1**

(a) H1: catching flu is associated with (not independent of) vaccination status **B1**

**(a) Answer: H0: catching flu is independent of vaccination status. H1: they are associated.**

(b) uses E = (row total * column total) / grand total = 90*40/160 **M1**

(b) 22.5 cso **A1**

(b) remaining three values correct: 67.5, 17.5, 52.5 **A1**

**(b) Answer: Expected frequencies: 22.5, 67.5, 17.5, 52.5.**

(c) the table is 2 by 2, giving 1 degree of freedom, so Yates' correction is applied to improve the approximation of the discrete data by the continuous chi-squared distribution **B1**

(c) uses $X^2$ = sum of $(|O-E| - 0.5)^2 / E$ **M1**

(c) first two terms correct (unsimplified): 16/22.5, 16/67.5 **A1**

(c) remaining two terms correct (unsimplified): 16/17.5, 16/52.5 **A1**

(c) cao, awrt 2.17 **A1**

**(c) Answer: Test statistic (with Yates' correction) = awrt 2.17**

(d) df = 1 **B1**

(d) critical value = 3.841 **B1**

(d) valid conclusion ft: since 2.17 < 3.841, do not reject H0; insufficient evidence of an association between vaccination and catching flu **B1**

**(d) Answer: df = 1, critical value = 3.841; 2.17 < 3.841 so do not reject H0.**

## Question 14

(a) H0: workshop attendance is independent of passing the resit exam **B1**

(a) H1: workshop attendance and passing the resit are associated (not independent) **B1**

**(a) Answer: H0: workshop attendance and passing the resit are independent. H1: they are associated.**

(b) uses E = (row total * column total) / grand total = 60*66/120 **M1**

(b) 33 cso **A1**

(b) remaining three values correct: 27, 33, 27 **A1**

**(b) Answer: Expected frequencies: 33, 27, 33, 27.**

(c) uses $X^2$ = sum of $(|O-E| - 0.5)^2 / E$ **M1**

(c) all four terms correct (unsimplified), e.g. 72.25/33, 72.25/27, 72.25/33, 72.25/27 **A1**

(c) cao, awrt 9.73 **A1**

**(c) Answer: Test statistic (with Yates' correction) = awrt 9.73**

(d) df = 1 **B1**

(d) compares 9.73 with the critical values 2.706 (10%), 3.841 (5%) and 6.635 (1%) **M1**

(d) 1%, since 9.73 > 6.635 **A1**

**(d) Answer: df = 1; H0 is rejected even at the 1% level.**

(e) strong evidence of an association between attending the workshop and passing the resit **B1**

(e) valid caveat, e.g. association does not imply causation; a confounding variable such as general motivation or prior attainment could affect both workshop attendance and passing, oe **B1**

**(e) Answer: Strong evidence of association; however, a confounding variable (e.g. motivation) could explain the link without attendance causing the outcome.**

## Question 15

(a) H0: X ~ B(5, p), with p estimated as in part (b) **B1**

(a) H1: X does not follow this binomial distribution **B1**

**(a) Answer: H0: X ~ B(5, $p_{hat}$). H1: X does not follow this binomial distribution.**

(b) $p_{hat}$ = (total defectives) / (total bulbs inspected) = 150 / (180*5) **M1**

(b) 1/6 cso **A1**

(b) correct use of P(X=2) = C(5,2) * $(1/6)^2$ * $(5/6)^3$ **M1**

(b) cao, awrt 0.1608 **A1**

**(b) Answer: $p_{hat}$ = 1/6; P(X=2) = awrt 0.1608**

(c) attempts E = 180 * P(X=x) for each value of X **M1**

(c) all six values correct (ft): 72.338, 72.338, 28.935, 5.787, 0.579, 0.023 **A1**

**(c) Answer: Expected frequencies: 72.338, 72.338, 28.935, 5.787, 0.579, 0.023**

(d) E(X=4)=0.579 and E(X=5)=0.023 are both less than 5, and even combined (0.602) they remain below 5, so X=3, 4 and 5 must all be combined into a single class 'X≥3' **B1**

(d) pooled observed table: 77, 67, 27, 9 **B1**

(d) pooled expected table: 72.338, 72.338, 28.935, 6.389 **B1**

**(d) Answer: Pooled classes 0, 1, 2, ≥3: O = 77, 67, 27, 9; E = 72.338, 72.338, 28.935, 6.389.**

(e) attempts sum of $(O-E)^2/E$ across the 4 pooled classes **M1**

(e) cao, awrt 1.89 **A1**

**(e) Answer: Test statistic = awrt 1.89**

(f) df = (4 pooled classes) - 1 - 1 (p estimated) = 2 **B1**

(f) critical value = 5.991 **B1**

(f) valid conclusion ft: since 1.89 < 5.991, do not reject H0; the binomial model B(5, 1/6) is a good fit for the number of defective bulbs per box **B1**

**(f) Answer: df = 2; critical value = 5.991; 1.89 < 5.991 so do not reject H0.**

## Question 16

(a) attempts sum of (fx) / (sum of f) **M1**

(a) cao, 1.5 **A1**

**(a) Answer: Mean = 1.5 scratches per windscreen**

(b) correct use of P(X=3) = $e^{-1.5}$ * $1.5^3$ / 3! **M1**

(b) cao, awrt 0.1255 **A1**

(b) expected frequency for X=3 (ft) = 12.551 **B1**

**(b) Answer: P(X=3) = awrt 0.1255; expected frequency = 12.551**

(c) E(X=4)=4.707 and E(X≥5)=1.858 are both less than 5, so these two classes must be combined into a single class 'X≥4' **B1**

(c) pooled observed table: 22, 34, 25, 12, 7 **B1**

(c) pooled expected table: 22.313, 33.470, 25.102, 12.551, 6.564 **B1**

**(c) Answer: Pooled classes 0, 1, 2, 3, ≥4: O = 22, 34, 25, 12, 7; E = 22.313, 33.470, 25.102, 12.551, 6.564.**

(d) attempts sum of $(O-E)^2/E$ across the 5 pooled classes **M1**

(d) cao, awrt 0.0663 **A1**

**(d) Answer: Test statistic = awrt 0.0663**

(e) df = (5 pooled classes) - 1 - 1 (mean estimated) = 3 **B1**

(e) critical value = 7.815 **B1**

(e) valid conclusion ft: since 0.0663 < 7.815, do not reject H0; a Poisson distribution is a suitable model for the number of scratches per windscreen **B1**

**(e) Answer: df = 3; critical value = 7.815; 0.0663 < 7.815 so do not reject H0.**

(f) suggests an unusually close (excellent) agreement between observed and expected frequencies **B1**

(f) any sensible reason, e.g. the data may have been rounded, adjusted, or not recorded entirely independently, producing values suspiciously close to the model, oe **B1**

**(f) Answer: The very small value suggests an unusually close fit; possible reasons include the data being rounded/adjusted or not collected fully independently.**

## Question 17

(a) H0: defect type is independent of shift **B1**

(a) H1: defect type is associated with shift **B1**

**(a) Answer: H0: defect type and shift are independent. H1: they are associated.**

(b) uses E = (row total * column total) / grand total = 100*59/200 **M1**

(b) 29.5 cso **A1**

(b) remaining values correct: 29.5, 20.5, 20.5 for each shift (both rows identical since row totals are equal) **A1**

**(b) Answer: Expected frequencies (each shift): Scratch 29.5, Crack 29.5, Discolour 20.5, Misalign 20.5.**

(c) attempts sum of $(O-E)^2/E$ across all 8 cells **M1**

(c) cao, awrt 3.72 **A1**

**(c) Answer: Test statistic = awrt 3.72**

(d) df = (2-1)(4-1) = 3 **B1**

(d) critical value = 7.815 **B1**

(d) valid conclusion ft: since 3.72 < 7.815, do not reject H0; no evidence that defect type is associated with shift **B1**

**(d) Answer: df = 3; critical value = 7.815; 3.72 < 7.815 so do not reject H0.**

(e) H0: defect types occur in the ratio 3:3:2:2; H1: they do not **B1**

(e) converts the ratio to proportions 0.3, 0.3, 0.2, 0.2 and multiplies by 200 **M1**

(e) expected frequencies all correct: 60, 60, 40, 40 **A1**

(e) attempts sum of $(O-E)^2/E$ using the pooled column totals 59, 59, 41, 41 **M1**

(e) cao, awrt 0.0833 **A1**

(e) df = 3, critical value = 7.815; since 0.0833 < 7.815, do not reject H0 - the combined data are consistent with the claimed ratio **B1**

**(e) Answer: Test statistic = awrt 0.0833; df = 3, critical value = 7.815; do not reject H0.**

(f) part (e) uses the combined totals across both shifts, so any difference between Day and Night has been lost once the data are pooled **B1**

(f) the question of whether shift affects defect type is answered only by the test in parts (a) to (d), which uses the full two-way breakdown **B1**

**(f) Answer: Part (e) pools both shifts together, so it cannot detect any difference between Day and Night; only the two-way test in (a) to (d) can address association with shift.**