2.3 Further Statistics: Chi-Squared Tests (Goodness of Fit and Contingency Tables): 10-minute Pack

Calculator allowed · about 10 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 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)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 2.3 Further Statistics: Chi-Squared Tests (Goodness of Fit and Contingency Tables): 10-minute Pack

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	

© Revision Library | revisionlibrary.co.uk