

Yates' Correction for 2 by 2 Tables

Starter

1. **(Review of last lesson)** In addition to being classed into grades 9 – 1, 200 students are classified as female, male and prefer not to say. Assuming all expected values are 5 or more, the statistic $\sum \frac{(O_i - E_i)^2}{E_i}$ was 33.85. Stating your hypotheses and using a 1 % significance level investigate whether or not sex and grade are associated.

Working: H_0 : there is no relationship between sex and grades
 H_1 : there is a relationship between sex and grades
 $\chi^2_{calc} = 33.85$
 Degrees of freedom, $\nu = 8 \times 2 = 16$
 The critical value at the 1 % level is $\chi^2_{16}(1\%) = 32.00$
 Since $\chi^2_{calc} = 33.85 < 32.00 = \chi^2_{16}(1\%)$, we reject H_0 .
 There is evidence of an association between sex and grades

- E.g. 1** A research worker studying the age of adults and the number of credit cards they possess obtained these results.

		Number of cards	
		≤ 3	> 3
Age	< 30	74	20
	≥ 30	50	35

Carry out a test at the 5 % level to decide whether or not there is an association between age and credit cards owned.

Working: First calculate expected frequencies

Expected frequencies		Number of cards	
		≤ 3	> 3
Age	< 30	65.12	28.88
	≥ 30	58.88	26.12

$$\chi^2_{calc} = \sum \frac{(|O_i - E_i| - 0.5)^2}{E_i}$$

$$= \frac{(|74 - 65.12| - 0.5)^2}{65.12} + \frac{(|20 - 28.88| - 0.5)^2}{28.88} + \frac{(|50 - 58.88| - 0.5)^2}{58.88} + \frac{(|35 - 26.12| - 0.5)^2}{26.12}$$

$$\chi^2_{calc} = 7.396$$

H_0 : there is not association between age and credit cards owned

H_1 : there is an association between age and credit cards owned

Degrees of freedom, $\nu = 1 \times 1 = 1$

The critical value at the 5 % level is $\chi^2_1(5\%) = 3.841$

Since $\chi^2_{calc} = 7.396 > 3.841 = \chi^2_1(5\%)$, we reject H_0 .

There is evidence of an association between age and credit cards owned

E.g. 2 During the trial of a new drug, 60 volunteers out of 200 were treated with the drug. Those experiencing relief of their symptoms and those who did not were recorded as follows:

Observed frequencies	Relief	No relief	Totals
Treated	10	50	60
Not treated	40	100	140
Totals	50	150	200

Test at the 10 % level to see if there is an association between treatment with the drug and relief of symptoms.

Working: First calculate expected frequencies:

Expected frequencies	Relief	No relief	Totals
Treated	15	45	60
Not treated	35	105	140
Totals	50	150	200

$$\chi_{calc}^2 = \sum \frac{(|O_i - E_i| - 0.5)^2}{E_i}$$

$$= \frac{(|10 - 15| - 0.5)^2}{15} + \frac{(|50 - 45| - 0.5)^2}{45} + \frac{(|40 - 35| - 0.5)^2}{35} + \frac{(|100 - 105| - 0.5)^2}{105}$$

$$\chi_{calc}^2 = 2.571$$

H_0 : there is not an association between treatment and relief of symptoms

H_1 : there an association between treatment and relief of symptoms

Degrees of freedom, $\nu = 1 \times 1 = 1$

The critical value at the 10 % level is $\chi_1^2(10\%) = 2.706$

Since $\chi_{calc}^2 = 2.571 < 2.706 = \chi_1^2(10\%)$, we do not reject H_0 .

There is no evidence of an association between treatment with the drug and relief of symptoms.

Video: [Yates' correction](#)

Video: [Yates' correction \(from 14:25\)](#)

[Solutions to Starter and E.g.s](#)

Exercise

p104 6B Qu 1i, 2-5, (6 red)