Yates' Correction (A Level Only)

In a chi-squared test for association, for two independent variables, the calculated χ^2 value is assumed to follow the distribution of χ^2_v, where v is the degrees of freedom and is given by:

$$v = (\text{number of rows} - 1)(\text{number of columns} - 1)$$

For a 2×2 contingency table, when $v = 1$, the approximation is not appropriate. A better approximation is to use Yates' correction, which is given by:

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

Example 1: The contingency table below shows the gender of pupils in a class and whether they have passed or failed an exam. Test, at the 10% significance level, whether gender and exam grade are independent.

	Male	Female
Failed	8	7
Passed	19	16

Step	Working
State the null and alternative hypotheses.	H_0: Gender and exam grade are independent. H_1: Gender and exam grade are not independent.
Find the row total, column total and overall total from the contingency table.	(see table below)
Find the expected value for each cell using $E_i = \frac{\text{row total} \times \text{column total}}{\text{overall total}}$.	(see table below)
Calculate the degrees of freedom.	$v = (2-1)(2-1)$ $= 1$
Since $v = 1$, use Yates' correction to calculate the value of χ^2.	(see working below)
Compare the chi-squared value with the critical value and state your conclusion.	$0.0613 < 2.706$ ∴ Do not reject H_0. There is insufficient evidence to show that gender and exam grades are dependent.

	Male	Female	Total
Failed	8	7	15
Passed	19	16	35
Total	27	23	50

	Male	Female
Failed	$\frac{15 \times 27}{50} = 8.1$	$\frac{15 \times 23}{50} = 6.9$
Passed	$\frac{35 \times 27}{50} = 18.9$	$\frac{35 \times 23}{50} = 16.1$

$$\chi^2_{\text{Yates}} = \sum \frac{(|O_i - E_i| - 0.5)^2}{E_i}$$
$$= \frac{(|8-8.1|-0.5)^2}{8.1} + \frac{(|7-6.9|-0.5)^2}{6.9} + \frac{(|19-18.9|-0.5)^2}{18.9} + \frac{(|16-16.1|-0.5)^2}{16.1}$$
$$= \frac{(-0.4)^2}{8.1} + \frac{(-0.4)^2}{6.9} + \frac{(-0.4)^2}{18.9} + \frac{(-0.4)^2}{16.1}$$
$$= 0.0613 \text{ (3s.f.)}$$

Sometimes it is not apparent that Yates' correction needs to be used from the beginning. If a contingency table becomes a 2×2 table after 2 or more rows or columns are merged, Yates' correction needs to be used as well.

Example 2: The contingency table below shows the age of applicants to a company and the outcome of their application. Test, at 10% significance level, whether applicants' age and their application outcome are independent.

	≤ 25	$26-35$	> 35
Accepted	14	5	1
Rejected	36	38	6

Step	Working
State the null and alternative hypotheses.	H_0: Applicants' age and their application outcomes are independent. H_1: Applicants' age and their application outcomes are dependent.
Find the row total, column total and overall total for the contingency table.	(see table below)
Construct a contingency table showing the expected values using $E_i = \frac{\text{row total} \times \text{column total}}{\text{overall total}}$.	(see table below)
Notice that the last column has expected values of ≤ 5, so the last two columns need to be merged.	(see table below)
Calculate the degrees of freedom from the merged table.	$v = (2-1)(2-1)$ $= 1$
Since $v = 1$, calculate χ^2 using Yates' correction.	(see working below)
Compare χ^2 with the critical value and state your conclusion.	$3.0625 > 2.706$ ∴ Reject H_0. There is sufficient evidence to show that applicants' age and their application outcome are dependent.

	≤ 25	$26-35$	> 35	Total
Accepted	14	5	1	20
Rejected	36	38	6	80
Total	50	43	7	100

	≤ 25	$26-35$	> 35
Accepted	$\frac{20 \times 50}{100} = 10$	$\frac{20 \times 43}{100} = 8.6$	$\frac{20 \times 7}{100} = 1.4$
Rejected	$\frac{80 \times 50}{100} = 40$	$\frac{80 \times 43}{100} = 34.4$	$\frac{80 \times 7}{100} = 5.6$

	≤ 25	≥ 26
Accepted	10	$8.6 + 1.4 = 10$
Rejected	40	$34.4 + 5.6 = 40$

$$\chi^2_{\text{Yates}} = \sum \frac{(|O_i - E_i| - 0.5)^2}{E_i}$$
$$= \frac{(|14-10|-0.5)^2}{10} + \frac{(|6-10|-0.5)^2}{10} + \frac{(|36-40|-0.5)^2}{40} + \frac{(|44-40|-0.5)^2}{40}$$
$$= \frac{(3.5)^2}{10} + \frac{(3.5)^2}{10} + \frac{(3.5)^2}{40} + \frac{(3.5)^2}{40}$$
$$= 3.0625$$

CC BY NC ND