

CIMT Statistics p210 Example

Subject choice and gender:

f_o	French	Poetry	Russian	Sculpture
Male	2	8	15	10
Female	10	17	21	37

H_0 : no association between gender and subject

H_1 : there is an association between gender and subject

Assume H_0 to be true

$\alpha = 5\%$, 1 tail test

generating expected frequencies:

f_e	French	Poetry	Russian	Sculpture
Male	3.5	7.3	10.5	13.7
Female	8.5	17.7	25.5	33.3

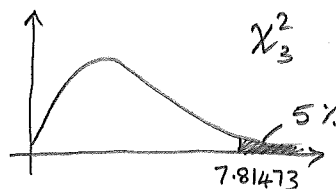
we check that none of f_e are < 1 , and at least 80% are ≥ 5 . \checkmark

we do have one entry out of eight which is < 5 , which is 12.5% of the f_e , which is acceptable

$$\text{so } X^2 = \sum \frac{(f_o - f_e)^2}{f_e} = \frac{(2 - 3.5)^2}{3.5} + \dots + \frac{(37 - 33.3)^2}{33.3} = 5.14363$$

we have $(4-1) \times (2-1) = 3 \times 1 = 3$ degrees of freedom

so X^2 has a χ^2_3 distribution:



as $5.14363 < 7.81473$, we are not in critical region

\therefore we have no evidence to reject H_0

\therefore We conclude that from this sample, there is no association between gender and subjects.