# Exercise 2.b)

The expected values remain the same as in exercise 2.a):

| **Expected** | *Woman* | not *Woman* | total |
|---|---|---|---|
| $I$ | 333 | 667 | $1'000$ |
| not $I$ | $19'667$ | $39'333$ | $59'000$ |
| total | $20'000$ | $40'000$ | $60'000$ |

The given observed data is

| **Observed** | *Woman* | not *Woman* | total |
|---|---|---|---|
| $I$ | 450 | 550 | $1'000$ |
| not $I$ | $19'550$ | $39'450$ | $59'000$ |
| total | $20'000$ | $40'000$ | $60'000$ |

and therefore $\chi^2$ is given by

$$\begin{aligned}\chi^2 &= \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} \\ &= \frac{(450-333)^2}{333} + \frac{(550-667)^2}{667} + \frac{(19'550-19'667)^2}{19'667} + \frac{(39'450-39'333)^2}{39'333} \\ &\approx 62.67541.\end{aligned}$$

For a significance level of 1% and dof[1] = 1, the critical value for $\chi^2$ is 6.635. Because $\chi^2 \approx 62.67541 > 6.635$, it's clear that the Null-Hypothesis is **not** accepted and that there **is** a correlation between the term $I$ and female authors.

[1] dof = degrees of freedom