

# Contingency Tables

## Hypothesis Testing for Independence/An Association with small E Values

Important Note:

It should be noted that  $\sum \frac{(O-E)^2}{E}$  approximately follows a  $\chi^2$  distribution but this is only the case if  $E > 5$ .

If  $E < 5$  then  $\sum \frac{(O-E)^2}{E}$  does not reasonably follow a  $\chi^2$  distribution

Assumptions of the  $\chi^2$  test:

1. The data in the contingency table must be frequencies
2. The categories in the contingency table must be mutually exclusive
3. Each subject may contribute data to one and only one cell in the contingency table
4. Expected values should be at least 5 when calculating a  $\chi^2$  test.

If an expected value is found to be less than 5 we must combine classes to increase this number.

- The combinations must be sensible.
- The number of combinations should be minimal
- Combining classes in turn reduces the degrees of freedom.
- Reducing the degrees of freedom reduced the critical value
- A lower critical value mean a Type I error is more likely to occur