

Goodness of Fit

Tests whether the categorical data matches a hypothesized distribution.

Null and alternative hypotheses

- H_0 : The stated distribution is correct
- H_a : The stated distribution is not correct

Calculating χ^2

$$\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

$$\text{Degrees of freedom} = \# \text{categories} - 1$$

Validity requirements

χ^2 tests are valid if:

- Random data
- 10% rule
 - $n \leq 0.1N$
- Large counts
 - Expected counts > 5
 - $n \cdot \hat{p} \geq 10$
 - $n(1 - \hat{p}) \geq 10$

Two-Way Tables

Test for Homogeneity

Tests whether the distribution of a categorical variable is identical in two populations.

Null and alternative hypotheses

- H_0 : There is no difference in the distribution of the categories
- H_a : There is a difference in the distribution of the categories

Calculating Expected Counts & Degrees of Freedom

$$\text{Expected Count} = \frac{(\text{Row total})(\text{Column total})}{\text{Table total}}$$

$$\text{Degrees of freedom} = (\# \text{rows} - 1)(\# \text{columns} - 1)$$

Calculator Note

- χ^2 *cdf* P-value from χ^2
- χ^2 *GOF-wTest* G'ness Fit
- χ^2 *Test* Ind. & Homog.

Test for Independence

Tests whether there is an association between two categorical variables. Calculation same as above.

Null and alternative hypotheses

- H_0 : There is no association among the categories (they are independent)
- H_a : There is an association among the categories (they are not independent)