

Validity requirements

 $X^2$  tests are valid if:

• Random data

Large counts

 $n \le 0.1N$ 

Expected counts > 5

• 10% rule

## **Goodness of Fit**

Tests whether the categorical data matches a hypothesized distribution.

#### Null and alternative hypotheses

- $H_0$ : The stated distribution is correct
- $H_{\alpha}$ : The stated distribution is not correct

## Calculating $\chi^{\scriptscriptstyle 2}$

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

Degrees of freedom = #categories – 1

# **Two-Way Tables**

#### **Test for Homogeneity**

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

#### Null and alternative hypotheses

- H<sub>0</sub>: There is no difference in the distribution of the categories between the populations.
- *H*<sub>a</sub>: There is a difference in the distribution of the categories between the populations.

# Calculator Note χ<sup>2</sup> cdf P-value from χ<sup>2</sup>

- $\chi^2 GOF-wTest$  G'ness Fit
- $\chi^2$ -Test Ind. & Homog.

## Calculating Expected Counts & Degrees of Freedom

Expected Count = (Row total)(Column total) Table total

Degrees of freedom = (#rows - 1)(#columns - 1)

#### **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)
- $H_{\alpha}$ : There is an association among the categories (they are not independent)