

§11.2: Comparing Counts (Chi-Square)

Skills

- Conduct a hypothesis test using the Chi Square procedures (Independence of Variables, Homogeneity of Proportions)
-
-
-
-
-
-
-
-
-
-
-

The Chi Square Independence of Variables Hypothesis Test

H_0 : the variables are independent

H_a : the variables are not independent

Conditions (Independence)

- Random sample
 - Sample should be smaller than 10% of the population (if sampling without replacement)
 - All expected counts are at least 5
-
-
-
-
-
-
-
-
-
-

A Twist

What if you measure *one* qualitative variable across *several populations*?

Chi Square can be used to determine if that one variable has the same distribution across all populations

The Chi Square Homogeneity of Proportions Hypothesis Test

H_0 : the variable has the same distribution across all populations

H_a : the variable does not have the same distribution across all populations

Conditions (Homogeneity)

- Stratified random sample* or randomized experiment
 - *10% condition if sampling without replacement
 - All expected counts are at least 5
-
-
-
-
-
-
-
-
-
-

Mechanics (both tests)

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

$$df = (\text{rows} - 1)(\text{columns} - 1)$$

How to tell the difference

- One variable, one population:

Goodness of Fit

- One variable, multiple populations:

Homogeneity of Proportions

- Two variables, one population:

Independence of Variables

Consider how the data were collected!

Example

A study was conducted to determine if there is a relationship between jogging and blood pressure. A random sample of 210 people was obtained, and each person was classified according to these two criteria.

The data obtained are as follows...

Example

| Jog? | Low BP | Moderate BP | High BP |
|------|--------|-------------|---------|
| Yes | 34 | 57 | 21 |
| No | 15 | 63 | 20 |

Do the data provide evidence that jogging and blood pressure are not related?