Chi-Square Test

Contingency Table Analysis:
A Chi-Square Test for Independence

Second Classification Category	First Classification Category					
	1	2	3	4	5	**Row Total**
1	O_{11}	O_{12}	O_{13}	O_{14}	O_{15}	$\mathbf{R_1}$
2	O_{21}	O_{22}	O_{23}	O_{24}	O_{25}	$\mathbf{R_2}$
3	O_{31}	O_{32}	O_{33}	O_{34}	O_{35}	$\mathbf{R_3}$
4	O_{41}	O_{42}	O_{43}	O_{44}	O_{45}	$\mathbf{R_4}$
5	O_{51}	O_{52}	O_{53}	O_{54}	O_{55}	$\mathbf{R_5}$
Column Total	$\mathbf{C_1}$	$\mathbf{C_2}$	$\mathbf{C_3}$	$\mathbf{C_4}$	$\mathbf{C_5}$	**n**

Contingency Table Analysis: A Chi-Square Test for Independence

A and B are independent if: $P(A \cap B) = P(A) \times P(B)$.

If the first and second classification categories are independent: $E_{ij} = (R_i)(C_j)/n$

Null and alternative hypotheses:

H0: The two classification variables are independent of each other

H1: The two classification variables are not independent

Chi-square test statistic for independence:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Degrees of freedom: df=(r-1)(c-1)

Expected cell count: $E_{ij} = \frac{R_i C_j}{n}$

Chi-Square Test for Equality of Proportions

Tests of equality of proportions across several populations are also called **tests of homogeneity.**

In general, when we compare c populations (or r populations if they are arranged as rows rather than columns in the table), then the Null and alternative hypotheses:

H0: p1 = p2 = p3 = … = pc

H1: Not all pi, I = 1, 2, …, c, are equal

Chi-square test statistic for equal proportions:

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$

Degrees of freedom: df = (r-1)(c-1)

Expected cell count: $E_{ij} = \frac{R_i C_j}{n}$