

Inference Guide – Categorical Data Distributions (χ^2)

One Variable

One Sample [df = # of cells/categories – 1]

↳ compared with population model

H_0 : distribution = specified model

H_A : distribution \neq specified model (right sided)

A0 Data are counts.

C0 (Are they?)

A1 Individuals/data independent.

C1 SRS and $n < 10\%$ population.

A2 Sample large enough

C2 All expected counts ≥ 5 .

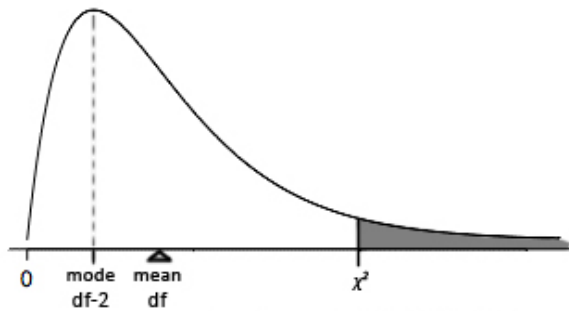
χ^2 test for Goodness-of-Fit [df = # of cells – 1]

$$\chi^2 = \sum_{\text{all cells}} \frac{(\text{Obs} - \text{Exp})^2}{\text{Exp}}$$

One Variable ↓	Obs Counts	Exp Value (Counts)	Residuals (Obs-Exp)	(Resid) ²	Component $\frac{(\text{Obs}-\text{Exp})^2}{\text{Exp}}$
Cat. 1		$\sum \text{Obs} * \text{hyp}$			
Cat. 2		$\sum \text{Obs} * \text{hyp}$			
Cat. 3		$\sum \text{Obs} * \text{hyp}$			
Cat. 4		$\sum \text{Obs} * \text{hyp}$			

df = #cat-1 $\sum \text{Obs}$

$\chi^2 = \sum$



P-value = $\chi^2 \text{cdf}(\chi^2, 999, \text{df})$

Or use: $\chi^2 \text{GOF-Test}(L_{\text{Obs}}, L_{\text{Exp}}, \text{df})$

If reject H_0 , then ☆