

# THE CHI-SQUARED STATISTIC

Comparing observed and expected results:

Meaningful v. meaningless data

#### 

### EXAMPLE PROBLEM: LIKERT-SCALE QUESTIONS

- Did the instructor dress appropriately for class?
  - (a) always
  - (b) most of the time
  - (c) seldom
  - (d) never



• 100 students responds:

- (a) always 40
- (b) most of the time 30
- (c) seldom 20
- (d) never 10

#### 

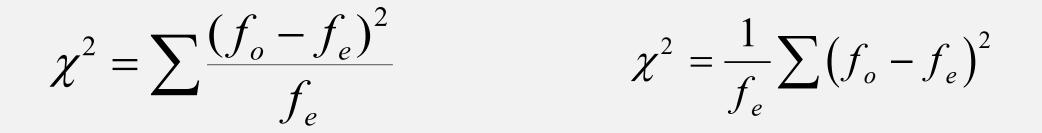
## THE CHI-SQUARED FORMULAS

- N = the number of data value(students responding) = 100
- k = the number of categories (choices) = 4
- $f_e =$  the expected frequency if the data are random = 100 / 4 = 25
- $f_o =$  the observed frequency = 40, 30, 20, and 10

$$f_e = \frac{N}{k} \qquad \qquad \chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$



# **OBSERVED FREQUENCIES**



$$\chi^{2} = \frac{1}{25} (40 - 25)^{2} + (30 - 25)^{2} + (20 - 25)^{2} + (10 - 25)^{2}$$

### CHI-SQUARED SOLUTION: TWO IMPLEMENTATIONS

• Single function solution

Ē

- Focus on data input and output
- Converting formula, especially  $\sum$ , to C++
- Client-supplier solution
  - Separating input and output from computation is good program design
  - A supplier defines and supplies reusable services
  - The client is an application program that uses the supplier's services