

CHI SQUARE

$$\chi^2$$

PEARSON CHI SQUARE TEST OF ASSOCIATION

Tom Malloy

EXPERIMENTAL SITUATION

Criteria for Using this test statistic

Essentially the same as the Goodness of fit test, except that...

We don't need prior probabilities

Each person (or thing) can be classified two ways instead of one

1. two partitions (or classification schemes)...

Example:

- ▶ Gender (Male vs Female)
- ▶ Job description (Management, Clerical)

A two-way classificatin table:

2. N independent observations (each classified two ways)

3 Frequency data...

What is the question being asked? In the example

Is the first classification scheme independent of the second classification scheme?

Suppose you suspect that a certain corporation has sexist hiring practices, that is, you think it hires men to be managers and women to be clerical help.

Sci Hyp:

PCH of chance:

To decide between the hypotheses, you do research:

Go to the corporation and...

The essential question is....

Example if there is COMPLETE dependence: (complete association)

Example if there is strong dependence (strong association)

If the two classification schemes are INDEPENDENT then you cannot predict one from the other. (No association)

Suppose the data turn out like this:

Pearson's Chi-Square test of Association evaluates PCH of chance

How to calculate expected frequencies

Review example

Data (f_{jk}):

$fe_{jk} =$

Example of calculating expected frequencies (fe_{jk})

Statistical Hypotheses

H_0

H_1

FORMULA FOR CHI-SQUARE

df =

EXAMPLE CALCULATIONS

Statistical Conclusion Validity

CAVEATS in the use of **CHI-SQUARE**

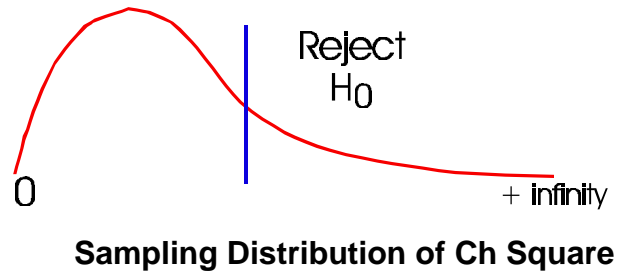
1. EACH f_{ej} must be 10 or greater if $df = 1$

EACH f_{ej} must be 5 or greater if $df = 2$ or more.

2. EVERY observation must be independent; no person can contribute two frequencies to the data.

Sampling Distribution of Chi Square

Assumptions about
where data come from



Sampling Distribution of Ch Square

Frequency Data:

Sample Statistic =