

Chi Square Tests

Chi-Square Test (χ^2)

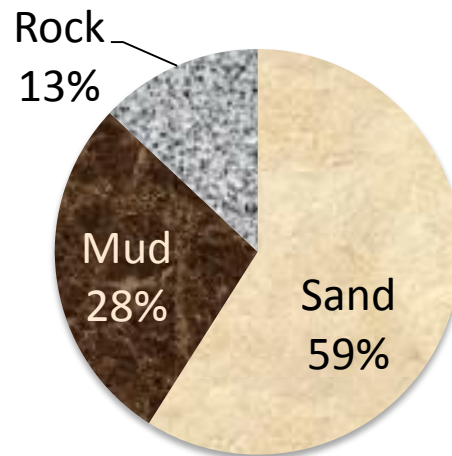
- Nonparametric test for nominal independent variables
 - These variables, also called "attribute variables" or "categorical variables," classify observations into a small number of categories.
 - Examples?
 - The dependent variable: the count of observations in each category of the nominal variable
 - e.g. the number of men versus women; or the number of accidents in dry weather versus in wet weather

Two uses of χ^2

- **Chi Square Goodness of Fit:** when we have one independent variable (e.g. weather)
 - we compare the numbers of observations in the categories of this variable (e.g. wet and dry) to what we would expect if the variable did not make any difference (e.g. out of 100 accidents 50 in dry weather and 50 in wet weather)
- **Chi Square Test of Independence:** when we have two or more independent variables
 - we compare the numbers of observations in each category of each variable to the numbers we would expect if the variables were independent of each other

Chi-Square Goodness of Fit

- The observed counts of numbers of observations in each category are compared with the expected counts, which are calculated using some kind of theoretical expectation, such as a 1:1 sex ratio.
- An example:
 - An area of shore that has 59% of the area covered in sand, 28% mud and 13% rocks;
 - if seagulls were standing in random places, your null hypothesis would be that 59% of the seagulls were standing on sand, 28% on mud and 13% on rocks.
 - the independent variable is type of shore, the dependent variable is the observed number of seagulls



Tabulating Chi Square Goodness of Fit

| | Sand | Mud | Rocks |
|--------------------------------------|------|-----|-------|
| Observed seagulls (Total: 100) | 35 | 8 | 57 |
| Expected seagulls (Total: 100) | 59 | 28 | 13 |

Calculating the test value

➤ The test statistic is calculated by taking an observed number (O), subtracting the expected number (E), then squaring this difference. The larger the deviation from the null hypothesis, the larger the difference between observed and expected.

$$X^2 = \sum \left[\frac{(O - E)^2}{E} \right]$$

➤ Each squared difference is divided by the expected number, and these standardized ratios are summed: the more differences between what you would expect and what you get the bigger the number.

Calculate Chi square for the seagull data

$(O - E)^2$ divided by E:

Now check your answer on Graphpad:

<http://www.graphpad.com>

Or on Social Science Statistics:

<http://www.socscistatistics.com/tests/Default.aspx>

(this will also give you the steps of the calculation)

Goodness of fit in SPSS

- Create a variable column (Surface)
- Create frequency column
- Type the observed frequencies for each category of the independent variable
- From the data menu, weight cases by frequency
- Go to Analyse – Nonparametric – One Sample - Chi square
- Select Surface as test field
- In Options, you can set the expected values to be equal percentages of the categories (33% here) or you can assign expected values
- Run the analysis

How the test works

1. Identify Pop. Distribution & Assumptions

- a) Two populations, one distribution that matches expected outcomes and another where distribution matches observed outcomes.
- b) Null hypothesis: the two distributions do not differ
- c) Comparison distribution is chi-square

Chi-Square Test for Goodness-of-Fit distribution

Characteristics of the comparison distribution

- Degrees of Freedom: N of categories $- 1$

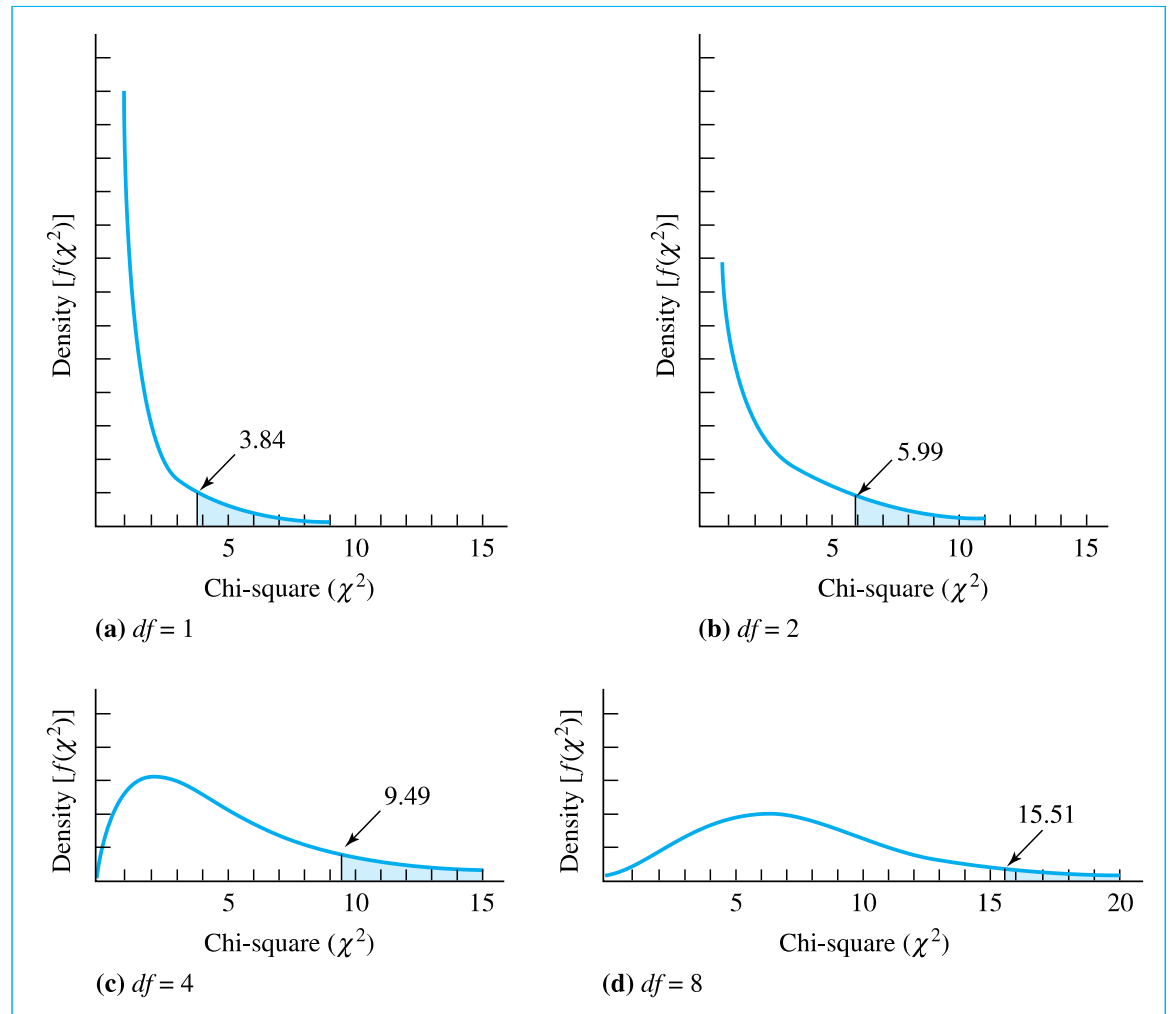


Figure 19.1

Chi-Square distribution for $df = 1, 2, 4,$ and 8 with critical values for $\alpha = .05$.

Chi-Square Goodness of Fit:

Test Assumptions

1. Random and independent sampling.
2. Sample size must be sufficiently large (no more than 20% of cells should have an expected value of less than 5)
3. Values of the variable are mutually exclusive and exhaustive. Every subject must fall in only one category.

Note: If these values are not met, the critical values in the chi-square table are not necessarily correct.

Exercise 1: Popularity

- A Psychology course is offered by three different professors.
- The table shows the number of students enrolled in the course of each.
- Is one professor more popular than another or are the different enrollment numbers due to chance?
- Complete the table and run the analysis.

| | Prof A | Prof B | Prof C | Prof D |
|----------|--------|--------|--------|--------|
| Observed | 25 | 29 | 22 | 17 |
| ????? | | | | |

Chi Square Test of Independence

- Two or more nominal variables
- We test the independence of the variables (whether they affect each other)

Chi-Square Test of Independence

Example

A researcher wants to know if there is a significant difference in the frequencies with which males come from small, medium, or large cities as contrasted with females. The two variables are hometown size (small, medium, or large) and sex (male or female). Another way of putting our research question is: Is gender independent of size of hometown?

Contingency table for the data for 30 females and 6 males:

Frequency with which males and females come from small, medium, and large cities

| | Small | Medium | Large | Totals |
|--------|-------|--------|-------|--------|
| Female | 10 | 14 | 6 | 30 |
| Male | 4 | 1 | 1 | 6 |
| Totals | 14 | 15 | 7 | 36 |

The formula for chi-square is:

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

Where:

O is the observed frequency, and

E is the expected frequency.

The degrees of freedom for the 2-D chi-square statistic is:

$$\mathbf{df = (Columns - 1) \times (Rows - 1)}$$

Computing Expected Frequencies

Frequency with which males and females come from small, medium, and large cities

| | Small | Medium | Large | Totals |
|--------|-------|--------|-------|--------|
| Female | 10 | 14 | 6 | 30 |
| Male | 4 | 1 | 1 | 6 |
| Totals | 14 | 15 | 7 | 36 |

Expected Frequency for each Cell:

The cell's Column Total x the cell's Row Total / Grand Total

In our example:

Column Totals are 14 (small), 15 (medium), and 7 (large).

Row Totals are 30 (female) and 6 (male).

Grand total is 36.

Computing Expected Frequencies

Frequency with which males and females come from small, medium, and large cities

| | Small | Medium | Large | Totals |
|--------|-------|--------|-------|--------|
| Female | 10 | 14 | 6 | 30 |
| Male | 4 | 1 | 1 | 6 |
| Totals | 14 | 15 | 7 | 36 |

The expected frequency:

1. Small female cell: $14 \times 30 / 36 = 11.667$
2. Medium female cell: $15 \times 30 / 36 = 12.500$
3. Large female cell: $7 \times 30 / 36 = 5.833$
4. Small male cell: $14 \times 6 / 36 = 2.333$
5. Medium male cell: $15 \times 6 / 36 = 2.500$
6. Large male cell: $7 \times 6 / 36 = 1.167$

Observed frequencies, expected frequencies, and $(O - E)^2/E$ for males and females from small, medium, and large cities

| | Small | | | Medium | | | Large | | | Totals |
|---------------|----------|----------|-------------|----------|----------|-------------|----------|----------|-------------|--------|
| | Observed | Expected | $(O-E)^2/E$ | Observed | Expected | $(O-E)^2/E$ | Observed | Expected | $(O-E)^2/E$ | |
| Female | 10 | 11.667 | 0.238 | 14 | 12.500 | 0.180 | 6 | 5.833 | 0.005 | 30 |
| Male | 4 | 2.333 | 1.191 | 1 | 2.500 | 0.900 | 1 | 1.167 | 0.024 | 6 |
| Totals | 14 | | | 15 | | | 7 | | | 36 |

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

$$= 0.238 + 0.180 + 0.005 + 1.191 + 0.900 + 0.024 = 2.538$$

- For practice:
 - Check the accuracy of the hand calculations on Social Science Statistics. Are the two variables independent of each other?

Fischer's Exact Test

- Chi square test is not accurate when we have a small number of observations (expected frequency of less than 5 in more than 20% of cells)
- We can substitute Fischer's exact in a 2 x 2 design

Exercise 2: Cycling (handlebar.sav)

- Are the Dutch reckless cyclists?
 - Keep only one hand on the handlebar
- Variables:
 - Nationality (English, Dutch)
 - Hands on handlebar (One, Two)

Observed frequencies

| | Dutch | English |
|------------|-------|---------|
| One Handed | 120 | 17 |
| Two Handed | 578 | 154 |

In SPSS:

- Weight cases by frequency
- Go to Analyze -> Descriptive Statistics -> Crosstabs
- Choose Chi square from Statistics
- You can also choose Phi and Cramer's V – an effect size
- From Exact, you can choose Fischer's exact
- From Cells, choose the data you need (expected, possibly percentages)
- Run

Other way of running Chi-Square using SPSS: ResponseIncentive.sav

A researcher is interested in whether people are more likely to return survey questionnaires if the questionnaire offers an incentive. He sends out 100 questionnaires:

- 20 promises that the respondent will get the survey results
- 30 says the respondent will be entered in a prize draw
- 50 has no incentive

We have two categorical variables: **Incentive** (results vs. prize draw vs. none) and **Response** (questionnaire returned or not).

Steps

- Analyse – Descriptive Statistics – Crosstabs
- Choose rows and columns
- Click display bar charts
- Choose statistics (Chi² and Cramer's V for effect size)
- Choose which cells you want displayed (observed and expected)
- Run the analysis

SPSS Output

| Incentive * Response Crosstabulation | | | | | |
|--------------------------------------|----------------|----------------|----------|------|-------|
| | | | Response | | Total |
| | | | 0 | 1 | |
| Incentive | SurveyResults | Count | 11 | 9 | 20 |
| | | Expected Count | 12.6 | 7.4 | 20.0 |
| | PrizeDraw | Count | 14 | 16 | 30 |
| | | Expected Count | 18.9 | 11.1 | 30.0 |
| | NoReward | Count | 38 | 12 | 50 |
| | | Expected Count | 31.5 | 18.5 | 50.0 |
| Total | Count | | 63 | 37 | 100 |
| | Expected Count | | 63.0 | 37.0 | 100.0 |

| Chi-Square Tests | | | |
|------------------------------|--------------------|----|-----------------------|
| | Value | df | Asymp. Sig. (2-sided) |
| Pearson Chi-Square | 7.608 ^a | 2 | .022 |
| Likelihood Ratio | 7.702 | 2 | .021 |
| Linear-by-Linear Association | 4.568 | 1 | .033 |
| N of Valid Cases | 100 | | |

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 7.40.

How big is the effect?: Cramer's V

| Symmetric Measures | | |
|------------------------|-------|--------------|
| | Value | Approx. Sig. |
| Nominal by Nominal Phi | .276 | .022 |
| Cramer's V | .276 | .022 |
| N of Valid Cases | 100 | |

a. Not assuming the null hypothesis.
b. Using the asymptotic standard error assuming the null hypothesis.

.27 out of 1 = a medium association between type of incentive and whether people return a questionnaire.

Can be viewed like a correlation coefficient. The significance level indicates it is unlikely the observed pattern of data is due to chance.

A more useful effect size: Odds Ratio

1. Odds that a Q was returned given a promise of results.

$$\text{Odds}_{(\text{responding to results})} = \frac{\text{number that responded to results} = 9}{\text{number that didn't respond} = 11} = .82$$

2. Odds that a Q was returned given a promise of prize draw.

$$\text{Odds}_{(\text{responding to draw})} = \frac{\text{number that responded to draw} = 16}{\text{number that didn't respond} = 14} = 1.14$$

3. Odds ratio.

$$\frac{\text{Odds}_{(\text{responding to draw})}}{\text{Odds}_{(\text{responding to results})}} = 1.39$$

Odds ratios can be calculated for any pairs of categories.

Writing up the results

- There was a significant association between the type of incentive and whether people returned the questionnaire, $X^2(2) = 7.61$, $p = .022$. People offered either type of incentive were more likely to respond than those not offered any incentive but, based on the odds ratio, the odds of returning the questionnaire were 1.39 times higher if people were promised a prize draw than if they were promised the results of the survey.

Exercise 4

The relationship between drug companies and medical researchers is under scrutiny because of possible conflict of interest. The issue that started the controversy was a 1995 case control study that suggested that the use of calcium-channel blockers to treat hypertension led to an increase risk of heart disease. This led to an intense debate. Researchers writing in the *New England Journal of Medicine* (“Conflict of Interest in the Debate over Calcium Channel Antagonists,” January 8, 1998, p. 101) looked at the 70 research reports that appeared during 1996–1997, classifying them as favorable, neutral, or critical toward the drugs. The researchers then contacted the authors of the reports and questioned them about financial ties to drug companies. Results in ResearchBribes.sav

Homework

- `Sonnentag(2012).sav`
 - Is there an association between the amount of time pressure at work and whether we can relax when not working (`SwitchOff`)?
- `Births.sav`
 - Are births equally distributed over the year or are more babies born in some months than in others?

Make-up Homework

Does a negative example on TV make us more negative in our relationships? Eastenders.sav

- Couples watched three types of TV programmes:
 - EastEnders (British soap opera with very miserable and mean people)
 - Friends (British soap opera with exaggeratedly nice and helpful people)
 - A neutral nature programme
- After watching each programme, the couples were left alone for an hour and the number of sharp/nasty/unfriendly comments they make to each other was counted.

Choose the appropriate test, run it and report the results.