Section 12 Part 2

Chi-square test

McNemar’s Test
Section 12 Part 2 Overview

• Section 12, Part 1 covered two inference methods for categorical data from 2 groups
  – Confidence Intervals for the difference of two proportions
  – Two sample z-test of equality of two proportions

• Section 12 Part 2 covers inference methods for categorical data
  – Chi-square test for comparisons between 2 categorical variables (Fisher’s exact test)
  – McNemar ‘s Chi-square (Binomial Test) test for paired categorical data
Chi-square test

• The Chi-square test can also be used to test for independence between two variables
  – The null hypothesis for this test is that the variables are independent (i.e. that there is no statistical association).
  – The alternative hypothesis is that there is a statistical relationship or association between the two variables.

• The Chi-square test can be used to test for equality of proportions between two or more groups.
  – The null hypothesis for this test is that the 2 proportions are equal.
  – The alternative hypothesis is that the proportions are not equal (test for a difference in either direction)
Contingency Tables

- Setting: Let $X_1$ and $X_2$ denote categorical variables, $X_1$ having $I$ levels and $X_2$ having $J$ levels. There are $IJ$ possible combinations of classifications.

<table>
<thead>
<tr>
<th></th>
<th>Level=1</th>
<th>Level=2</th>
<th>...........</th>
<th>Level=J</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level=1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level=2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level=I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- When the cells contain frequencies of outcomes, the table is called a contingency table.
Chi-square Test:  
Testing for Independence

Step 1: Hypothesis (always two-sided):
- $H_0$: Independent
- $H_A$: Not independent

Step 2: Calculate the test statistic:

$$ X^2 = \sum \frac{(x_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2 \text{ with df } = (I - 1)(J - 1) $$

Step 3: Calculate the p-value

$$ p\text{-value} = P(X^2 > X^2) \text{ <- value 2-sided} $$

Step 4: Draw a conclusion

- $p\text{-value} < \alpha$ reject independence
- $p\text{-value} > \alpha$ do not reject independence
Racial Differences and Cardiac Arrest

• In a large mid-western city, the association in the incidence of cardiac arrest and subsequent survival was studied in 6117 cases of non-traumatic, out of hospital cardiac arrest.

• During a 12 month period, fewer than 1% of African-Americans survived an arrest-to-hospital discharge, compared to 2.6% of Caucasians.
## Racial Differences and Cardiac Arrest

<table>
<thead>
<tr>
<th>Race</th>
<th>Survival to Discharge</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Caucasian</td>
<td>84</td>
<td>3123</td>
</tr>
<tr>
<td>African-American</td>
<td>24</td>
<td>2886</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>108</strong></td>
<td><strong>6009</strong></td>
</tr>
</tbody>
</table>
Racial Differences and Cardiac Arrest

Scientific Hypothesis:
An association exists between race (African-American/Caucasian) and survival to hospital discharge (Yes/No) in cases of non-traumatic out-of-hospital cardiac arrest.

Statistical Hypothesis:

\[ H_0: \text{Race and survival to hospital discharge are independent in cases of non-traumatic out-of-hospital cardiac arrest.} \]

\[ H_A: \text{Race and survival to hospital discharge are not independent in cases of non-traumatic out-of-hospital cardiac arrest.} \]
Chi-square Test:

Testing for Independence

1. Obtain a random sample of \( n \) independent observations \( (\text{the selection of one observation does not influence the selection of any other}) \).

2. Observations are classified subsequently according to cells formed by the intersection rows and columns in a contingency table.
   - Rows \( (r) \) consist of mutually exclusive categories of one variable.
   - Columns \( (c) \) consist of mutually exclusive categories of the other variable.

3. The frequency of observations in each cell is determined along with marginal totals.
Chi-square Test: Testing for Independence

4. Expected frequencies are calculated under the null hypothesis of independence (no association) and compared to observed frequencies.

Recall: A and B are independent if:

\[ P(A \text{ and } B) = P(A) \times P(B) \]

5. Use the Chi-square \( (X^2) \) test statistic to observe the difference between the observed and expected frequencies.
\[ \chi^2 \] Distribution

- The probabilities associated with the chi-square distribution are in appendix D.
- The table is set up in the same way as the t-distribution.
- The chi-square distribution with 1 df is the same as the square of the Z distribution.
- Since the distribution only takes on positive values all the probability is in the right-tail.
Chi-square distributions and critical values for 1 df, 4 df and 20 df

For Chi-square with 20 df, the critical value $(\alpha = 0.05) = 31.4$

Critical value for $\alpha = 0.05$ and Chi-square with 1 df is 3.84

Critical value for $\alpha = 0.05$ and Chi-square with 4 df is 9.49

Since the Chi-square distribution is always positive, the rejection region is only in the right tail
How to Identify the critical value

• The rejection region of the Chi-square test is the upper tail so there is only one critical value
• First calculate the df to identify the correct Chi-square distribution
  – For a 2 X 2 table, there are (2-1)*(2-1) = 1 df

R commander:
> qchisq(0.95,1)
[1] 3.841459
State the conclusion

The p-value for $P( \chi^2 > X^2 )$

$> 1 - \text{pchisq}(X^2, 1)$

- Reject the null hypothesis by either the rejection region method or the p-value method

  $X^2 > \text{Critical Value}$

  or

  $P\text{value} < \alpha$
# Racial Differences and Cardiac Arrest

<table>
<thead>
<tr>
<th>Race</th>
<th>Survival to Discharge</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Caucasian</td>
<td>84</td>
<td>3123</td>
</tr>
<tr>
<td>African-American</td>
<td>24</td>
<td>2886</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
<td>6009</td>
</tr>
</tbody>
</table>

Under the assumption of independence:

\[ P(\text{YES and Caucasian}) = P(\text{YES}) \times P(\text{Caucasian}) \]

\[ = \frac{108}{6117} \times \frac{3207}{6117} = 0.009256 \]

Expected cell count \( e_{ij} \):

\[ e_{ij} = 0.009256 \times 6117 = 56.62 \]
Racial Differences and Cardiac Arrest

<table>
<thead>
<tr>
<th>Race</th>
<th>Survival to Discharge</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>YES</td>
<td>NO</td>
<td>Total</td>
</tr>
<tr>
<td>Caucasian</td>
<td>84</td>
<td>3123</td>
<td>3207</td>
</tr>
<tr>
<td></td>
<td>56.62</td>
<td>3151.43</td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>24</td>
<td>2886</td>
<td>2910</td>
</tr>
<tr>
<td></td>
<td>51.38</td>
<td>2854.82</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
<td>6009</td>
<td>6117</td>
</tr>
</tbody>
</table>

Expected Cell Counts = (Marginal Row total * Marginal Column Total)/ n

Rule of Thumb: Check to see if expected frequencies are > 2

No more than 20% of cells with expected frequencies < 5
Racial Differences and Cardiac Arrest

Step 1: Hypothesis (always two-sided):

\( H_0: \) Independent (Race/Survival)

\( H_A: \) Not independent

Step 2: Calculate the test statistic:

\[
X^2 = \sum \frac{(x_{ij} - e_{ij})^2}{e_{ij}}
\]

\[
= \frac{(84 - 56.62)^2}{56.62} + \frac{(3123 - 3151.43)^2}{3151.43} + \frac{(24 - 51.38)^2}{51.38} + \frac{(2886 - 2854.82)^2}{2854.82}
\]

\[
= 13.24 + 0.26 + 14.59 + .34 = 28.42
\]
Racial Differences and Cardiac Arrest

Step 3: Calculate the p-value

\[ p-value = P(X^2 > 28.42) < 0.00000001 \]

\[ > 1-pchisq(28.42,1) \]

[1] 9.765127e-08

Step 4: Draw a conclusion

A significant association exists between race and survival to hospital discharge in cases of non-traumatic out-of-hospital cardiac arrest.
Chi-square Test:
Testing for Equality or Homogeneity of Proportions

*Testing for equality or homogeneity of proportions* — examines differences between two or more independent proportions.

In chi-square test for independence, we examine the cross-classification of a *single sample* of observations on two qualitative variables.

The chi-square test can also be used for problems involving *two or more* independent populations.
Patients with evolving myocardial infarction were assigned independently and randomly to one of four thrombolytic treatments, and then followed to determine 30 day mortality.

<table>
<thead>
<tr>
<th>30 day outcome</th>
<th>Streptokinase and SC Heparin</th>
<th>Streptokinase and IV Heparin</th>
<th>Accelerated t-PA and IV Heparin</th>
<th>Accelerated t-PA Streptokinase with IV Heparin</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survived</td>
<td>9091</td>
<td>9609</td>
<td>9692</td>
<td>9605</td>
<td>37997</td>
</tr>
<tr>
<td>Died</td>
<td>705</td>
<td>768</td>
<td>652</td>
<td>723</td>
<td>2848</td>
</tr>
<tr>
<td>Total</td>
<td>9796</td>
<td>10377</td>
<td>10344</td>
<td>10328</td>
<td>40845</td>
</tr>
</tbody>
</table>

Are these four treatment populations equal with respect to 30-day mortality?
Chi-square Test: Testing for Equality or Homogeneity of Proportions

### Example

<table>
<thead>
<tr>
<th>30 day outcome</th>
<th>Streptokinase and SC Heparin</th>
<th>Streptokinase and IV Heparin</th>
<th>Accelerated t-PA and IV Heparin</th>
<th>Accelerated t-PA Streptokinase with IV Heparin</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survived</td>
<td>9091</td>
<td>9609</td>
<td>9692</td>
<td>9605</td>
<td>37997</td>
</tr>
<tr>
<td></td>
<td>9112.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Died</td>
<td>705</td>
<td>768</td>
<td>652</td>
<td>723</td>
<td>2848</td>
</tr>
<tr>
<td>Total</td>
<td>9796</td>
<td>10377</td>
<td>10344</td>
<td>10328</td>
<td>40845</td>
</tr>
</tbody>
</table>

Under the assumption of independence:

\[
P(\text{Streptokinase and SC Heparin and Survival}) = P(\text{Streptokinase and SC Heparin}) \times P(\text{Survival})
\]

\[
= \frac{9796}{40845} \times \frac{37997}{40845} = 0.223
\]

Expected cell count \(e_{ij} = 0.223 \times 40845 = 9112.95\)
Chi-square Test: Testing for Equality or Homogeneity of Proportions

Example

<table>
<thead>
<tr>
<th>30 day outcome</th>
<th>Streptokinase and SC Heparin</th>
<th>Streptokinase and IV Heparin</th>
<th>Accelerated t-PA and IV Heparin</th>
<th>Accelerated t-PA Streptokinase with IV Heparin</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survived</td>
<td>9091</td>
<td>9609</td>
<td>9692</td>
<td>9605</td>
<td>37997</td>
</tr>
<tr>
<td></td>
<td>9112.95</td>
<td>9653.44</td>
<td>9622.74</td>
<td>9607.86</td>
<td></td>
</tr>
<tr>
<td>Died</td>
<td>705</td>
<td>768</td>
<td>652</td>
<td>723</td>
<td>2848</td>
</tr>
<tr>
<td></td>
<td>683.05</td>
<td>723.56</td>
<td>721.26</td>
<td>720.14</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>9796</td>
<td>10377</td>
<td>10344</td>
<td>10328</td>
<td>40845</td>
</tr>
</tbody>
</table>

Under the assumption of independence:

Expected Cell Counts = (Marginal Row total * Marginal Column Total)/ n
Chi-square Test: Testing for Equality or Homogeneity of Proportions

Step 1: Hypothesis (always two-sided):

H₀: The four treatment options are homogeneous with respect to 30 day survival.

Hₐ: The four treatment options are not homogeneous with respect to 30 day survival.

Step 2: Calculate the test statistic:

\[ X^2 = \sum \frac{(x_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2 \text{ with df } = (I - 1)(J - 1) \]

Step 3: Calculate the p-value = P(\( X^2 > X^2 \))

Step 4: Draw a conclusion

p-value < \( \alpha \) reject independence

p-value > \( \alpha \) do not reject independence
Chi-square Test: Testing for Equality or Homogeneity of Proportions

Step 1: Hypothesis (always two-sided):

\( H_0: \) The four treatment options are homogeneous with respect to 30 day survival.

\( H_A: \) The four treatment options are not homogeneous with respect to 30 day survival.

Step 2: Calculate the test statistic:

\[
X^2 = \sum \frac{(x_{ij} - e_{ij})^2}{e_{ij}} = 10.85 \ \text{with df} = (2 - 1)(4 - 1) = 3
\]
Chi-square Test: Testing for Equality or Homogeneity of Proportions

Step 3: Calculate the p-value
\[ p\text{-value} = P(X^2 > 10.85) = 0.0126 \]

\[ > 1 - \text{pchisq}(10.85, 3) \]
[1] 0.01256526

Step 4: Draw a conclusion
\[ p\text{-value} < \alpha = 0.05 \text{ reject null} \]

The four treatment groups are not equal with respect to 30 day mortality. The largest relative departure from expected was noted in patients receiving accelerated t-PA and IV heparin, with fewer patients than expected dying.
Chi-square online calculator

This website will calculate the Chi-square statistic and p-value for data in a 2 X 2 table. Enter the cell counts in the table. Choose the Chi-square test without Yate’s correction to obtain the same results as in the example

www.graphpad.com/quickcalcs/contingency1.cfm
Chi-Square Testing: Rules of Thumb

• All expected frequencies should be equal to or greater than 2 (observed frequencies can be less than 2).

• No more than 20% of the cells should have expected frequencies of less than 5.

• What if these rules of thumb are violated?
Small Expected Frequencies

• Chi-square test is an approximate method.

• The chi-square distribution is an *idealized* mathematical model.

• In reality, the statistics used in the chi-square test are qualitative (have discrete values and not continuous).

• For 2 X 2 tables, use *Fisher’s Exact Test* (i.e. $P(x=k) \sim B(n,p)$) if your expected frequencies are less than 2. (Section 6.6)
Tests for Categorical Data

• To compare proportions between two groups or to test for independence between two categorical variables, use the Chi-square test.

• If more than 20% of the expected cell frequencies < 5, use the Fisher’s exact test.

• When categorical data are paired, the McNemar test is the appropriate test.
Comparing Proportions with Paired data

- When data are paired and the outcome of interest is a proportion, the **McNemar Test** is used to evaluate hypotheses about the data.
  - Developed by Quinn McNemar in 1947
  - Sometimes called the McNemar Chi-square test because the test statistic has a Chi-square distribution
Examples of Paired Data for Proportions

• Pair-Matched data can come from
  – Case-control studies where *each* case has a matching control (matched on age, gender, race, etc.)
  – Twins studies – the matched pairs are twins.

• Before - After data
  – the outcome is presence (+) or absence (-) of some characteristic measured on the same individual at two time points.
Summarizing the Data

• Like the Chi-square test, data need to be arranged in a contingency table before calculating the McNemar statistic
• The table will always be 2 X 2 but the cell frequencies are numbers of ‘pairs’ not numbers of individuals
• Examples for setting up the tables are in the following slides for
  – Case – Control paired data
  – Twins paired data: one exposed and one unexposed
  – Before – After paired data
Pair-Matched Data for Case-Control Study: outcome is exposure to some risk factor

<table>
<thead>
<tr>
<th>Case</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exposed</td>
</tr>
<tr>
<td>Exposed</td>
<td>a</td>
</tr>
<tr>
<td>Unexposed</td>
<td>c</td>
</tr>
</tbody>
</table>

The counts in the table for a case-control study are numbers of *pairs* not numbers of individuals.
Paired Data for Before-After Counts

The counts in the table for a before-after study are numbers of *pairs* and number of individuals.

<table>
<thead>
<tr>
<th>Before treatment</th>
<th>After treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>a</td>
</tr>
<tr>
<td>-</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>d</td>
</tr>
</tbody>
</table>
Null hypotheses for Paired Data

• The null hypothesis for case-control pair matched data is that the proportion of subjects exposed to the risk factor is equal for cases and controls.

• The null hypothesis for twin paired data is that the proportions with the event are equal for exposed and unexposed twins.

• The null hypothesis for before-after data is that the proportion of subjects with the characteristic (or event) is the same before and after treatment.
McNemar’s test

• For any of the paired data Null Hypotheses the following are true if the null hypothesis is true:

\[
\begin{align*}
\text{Ho: } & b = c \\
\text{Ho: } & b/(b+c) = 0.5
\end{align*}
\]

• Cells ‘b’ and ‘c’ are called the *discordant* cells because they represent pairs with a difference.

• Cells ‘a’ and ‘d’ are the *concordant* cells. These cells do not contribute any information about a difference between pairs or over time so they aren’t used to calculate the test statistic.
McNemar Statistic

• The McNemar’s Chi-square statistic is calculated using the counts in the ‘b’ and ‘c’ cells of the table:

\[ \chi^2 = \frac{(b - c)^2}{b + c} \]

• Rule of thumb: \( b + c \geq 20 \)
• If the null hypothesis is true the McNemar Chi-square statistic = 0.
McNemar statistic distribution

- The sampling distribution of the McNemar statistic is a Chi-square distribution.

- For a test with alpha = 0.05, the critical value for the McNemar statistic = 3.84.
  - The null hypothesis is not rejected if the McNemar statistic < 3.84.
  - The null hypothesis is rejected if the McNemar statistic > 3.84.
You can find the p-value for the McNemar statistic using R

1 - pchisq(X^2, 1)

If the test statistic is > 3.84, the p-value will be < 0.05 and the null hypothesis of equal proportions between pairs or over time will be rejected.
McNemar test Example

- Breast cancer patients receiving mastectomy followed by chemotherapy were matched to each other on age and cancer stage.

- By random assignment, one patient in each matched pair received chemo perioperatively and for an additional 6 months, while the other patient in each matched pair received chemo perioperatively only.
## Chemo Study

<table>
<thead>
<tr>
<th></th>
<th>Periop. only</th>
<th>Survived 5 years</th>
<th>Died within 5 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periop. + 6 Months</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survived 5 years</td>
<td>510</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>Died within 5 years</td>
<td>5</td>
<td></td>
<td>90</td>
</tr>
</tbody>
</table>
McNemar test hypotheses

Scientific Question: Does survival to 5 years differ by treatment group?

\[
H_0: b = c
\]

\[
H_A: b \text{ not equal } c
\]
McNemar test

• Check: \( b + c \geq 20 \)
• Critical value for Chi-square distribution with 1 df = 3.84
• Calculate the test statistic

\[
\chi^2 = \frac{(b - c)^2}{b + c} = \frac{(17 - 5)^2}{17 + 5} = 6.54
\]

• P-value = \( P(\chi^2 > 6.54) = 0.01 \)

> 1-pchisq(6.54,1)

[1] 0.01054753
Decision and Conclusion

• Decision: Reject Ho
  - By the rejection region method: 6.54 > 3.84
  - By the p-value method: 0.01 < 0.05

• Conclusion: The data provide evidence that an extra 6 months of chemotherapy results in a different survival rate compared to treatment with perioperative chemo alone. (p = 0.01).
McNemar test online calculator

This website will calculate the McNemar test statistic and p-value

http://www.graphpad.com/quickcalcs/McNemar1.cfm