Table Analysis

1. Pearson's Chi-squared test
2. Exact tests (Fisher's exact test)
3. Larger Tables

Valeria Vitelli
Oslo Centre for Biostatistics and Epidemiology
Department of Biostatistics, UiO
valeria.vitelli@medisin.uio.no

MF9130E – Introductory Course in Statistics
11.04.2024

Chi-squared test via an example

Example 17.1 in Kirkwood & Sterne

We consider data from an **influenza vaccination trial**. In this case the exposure is vaccination (the row variable), so the table includes row percentages.

Observed numbers

	Influenza		
	Yes	**No**	**Total**
Vaccine	20 (8.3%)	220 (91.7%)	240
Placebo	80 (36.4%)	140 (63.6%)	220
Total	100 (21.7%)	360 (78.3%)	460

We want to assess the strength of the evidence that vaccination affected the probability of getting influenza.

Chi-squared test via an example

We start by calculating the **expected numbers** under the assumption of no association between vaccination and subsequent contraction of influenza.
Overall 100/460 people got influenza, so the expected numbers **getting influenza** are:

- $100/460 \times 240 = 52.2$ in the vaccine group, and
- $100/460 \times 220 = 47.8$ in the placebo group.

Further, overall 360/460 people escaped influenza, so the expected numbers **escaping influenza** are:

- $360/460 \times 240 = 187.8$ in the vaccine group, and
- $360/460 \times 220 = 172.2$ in the placebo group.

Chi-squared test via an example

The **test statistic** is

$$\chi^2 = \frac{(20 - 52.2)^2}{52.2} + \frac{(80 - 47.8)^2}{47.8} + \frac{(220 - 187.8)^2}{187.8} + \frac{(140 - 172.2)^2}{172.2}$$
$$= 19.86 + 21.69 + 5.52 + 6.02 = 53.09,$$

and the corresponding P-**value** is < 0.001. There is strong evidence against the null hypothesis of no effect of the vaccine on the probability of contracting influenza. It is therefore concluded that the vaccine is effective.

Chi-squared test via an example

Alternative formulation of the Chi-squared test for a 2×2 table

- A quicker formula for calculating the **test statistic** on a 2×2 table is

$$\chi^2 = \frac{n \times (d_1 \cdot h_0 - d_0 \cdot h_1)^2}{d \cdot h \cdot n_1 \cdot n_0}, \quad \text{d.f.} = 1,$$

  using the previous notation for a 2×2 table

Chi-squared test via an example

Example 17.1 in Kirkwood & Sterne

In the example of the **influenza vaccination trial**, the chi-squared is

$$\chi^2 = \frac{460 \cdot (20 \cdot 140 - 80 \cdot 220)^2}{100 \cdot 360 \cdot 240 \cdot 220} = 53.01,$$

which, apart from rounding error, is the same as the value obtained using the formula of observed and expected numbers

(Fisher's) Exact test for 2×2 tables

- When the numbers in the 2×2 table are very small, we need an **exact test** to compare two proportions
- This is based on calculating the **exact probabilities** of the observed table and of more extreme tables with the same row and column totals, using the following formula:

$$\text{Exact probability} = \frac{d! \cdot h! \cdot n_1! \cdot n_0!}{n! \cdot d_1! \cdot d_0! \cdot h_1! \cdot h_0!},$$

  with the standard notation for a 2×2 table
- Manually computing a p-value is quite time-consuming, but immediate with **R**!

Example: 17.2 in Kirkwood & Sterne

Consider the results from a study to compare two treatment regimes for **controlling bleeding** in haemophiliacs undergoing surgery

	Bleeding complications		
Treatment regime	**Yes**	**No**	**Total**
A (group 1)	1 (d_1)	12 (h_1)	13 (n_1)
B (group 0)	3 (d_0)	9 (h_0)	12 (n_0)
Total	4 (d)	21 (h)	25 (n)

Only one (8%) of the 13 haemophiliacs given treatment regime A suffered bleeding complications, compared to three (25%) of the 12 given regime B

These numbers are too small for the **chi-squared test** to be valid:

- the overall total, 25, is less than 40, and
- the smallest expected value, $4/25 \cdot 12 = 1.9$ (complications with regime B), is less than 5

The **exact test** should therefore be used

The **exact probability** of the observed table is

$$\begin{aligned}\text{Exact probability} &= \frac{4! \cdot 21! \cdot 13! \cdot 12!}{25! \cdot 1! \cdot 3! \cdot 12! \cdot 9!} \\ &= 0.2261\end{aligned}$$

In addition, we need to calculate the probability that a **more extreme table** (with the same row and column totals as the observed table) could occur by chance under the null hypothesis that there is no difference between the two treatment regimes

$\rightarrow$ with **R**: $p > 0.05$, no evidence to reject

Larger contingency tables

- The **chi-squared test** can also be applied to larger tables, generally called $r \times c$ **tables**, where r denotes the number of rows in the table and c the number of columns
- The **test statistic** is:

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}, \quad \text{d.f.} = (r-1)\cdot(c-1), \qquad (1)$$

  which is chi-squared distributed with $(r-1)\cdot(c-1)$ degrees of freedom under the null hypothesis
- The general rule for calculating an **expected number** is:

$$E = \frac{\text{column total} \cdot \text{row total}}{\text{overall total}} \qquad (2)$$

Test validity for the case of larger contingency tables

- **Rule of thumb:**
  The approximation of **the chi-squared test is valid when**:
  - less than 20% of the expected numbers are under 5, and
  - none of the expected numbers is less than 1
- Sometimes this restriction **can be overcome by combining rows (or columns)** with low expected numbers, providing that these combinations make biological sense

Example: 17.3 in Kirkwood & Sterne

Consider the results from a survey to compare the **principal water sources** in three villages in West Africa.
The numbers of households using a river, a pond, or a spring are given. We will treat the **water source** as outcome and **village** as exposure, so column precentages are displayed.

Observed numbers

	Water source			
Village	**River**	**Pond**	**Spring**	**Total**
A	20 (40.0%)	18 (36.0%)	12 (24.0%)	50 (100.0%)
B	32 (53.3%)	20 (33.3%)	8 (13.3%)	60 (100.0%)
C	18 (45.0%)	12 (30.0%)	10 (25.0%)	40 (100.0%)
Total	70 (46.7%)	50 (33.3%)	30 (20.0%)	150 (100.0%)

Overall, 70 of the 150 households use a river. If there were no difference between villages, one would expect this same proportion of river usage in each village. Thus the **expected numbers** of households using a river in villages A, B and C, respectively, are:

$$\frac{70}{150} \cdot 50 = 23.3, \quad \frac{70}{150} \cdot 60 = 28.0 \quad \text{and} \quad \frac{70}{150} \cdot 40 = 18.7.$$

We use the same procedure to calculate the expected numbers of households using a pond and a spring in the villages.

Expected numbers

	Water source			
Village	**River**	**Pond**	**Spring**	**Total**
A	23.3	16.7	10.0	50
B	28.0	20.0	12.0	60
C	18.7	13.3	8.0	40
Total	70	50	30	150

The observed value of the **test statistic** is:

$$\begin{aligned}
\chi^2 &= \frac{(20-23.3)^2}{23.3}+\frac{(18-16.7)^2}{16.7}+\frac{(12-10.0)^2}{10.0} \\
&\quad +\frac{(32-28.0)^2}{28.0}+\frac{(18-18.7)^2}{18.7}+\frac{(20-20.0)^2}{20.0} \\
&\quad +\frac{(8-12.0)^2}{12.0}+\frac{(12-13.3)^2}{13.3}+\frac{(10-8.0)^2}{8.0} \\
&= 3.53,
\end{aligned}$$

with $\mathrm{d.f.} = (r-1)\cdot(c-1) = 2\cdot 2 = 4$ degrees of freedom.

The corresponding ***P*-value** is 0.47. This means that there is no evidence of a difference between the villages in the proportion of households using different water sources.

Final Remarks

McNemars test

- A special case of the 2x2 table for **paired categorical data**
- For example when measuring the presence/absence of something at two time points for each individual

Chi squared versus Fisher's exact test

- If the exact test works for both small and large samples, it is **natural to ask if we shall always use the exact test**
- Answer: **NO!** Why?
  - The two tests **have different assumptions**
  - The **exact test can be (very) conservative**; giving to high p-values and low power

Read more in e.g. Lydersen, Fagerland, Laake, *Recommended tests for association in 2x2 tables*, Statistics in Medicine, 2009: Fisher's exact should never be used without the mid-P correction