

# Hypothesis Testing

## Overview

In Comparative Studies, often the best way to analyse and report the results is to use confidence intervals. However, statistical hypothesis tests are still widely used in scientific work and thus, we need to consider how these tests work. Most statistical analyses involve comparisons (usually between treatments and procedures or between groups of subjects).

## Types of Hypotheses

The medical hypothesis is the basis of the statistical hypotheses (i.e. null hypothesis  $H_0$  and alternative hypothesis  $H_1$ ).

### Medical Hypothesis

The medical hypothesis is the starting point. It is statement that presents an idea of association between variables. For example, progressive polyarthritis (PAP) is associated with the HLA-DR4 antigen.

### Statistical Hypotheses

#### Null Hypothesis $H_0$

A null hypothesis states that there is no association between the variables of interest. For example, there is no association of PAP with HLA-DR4.

#### Alternative Hypothesis $H_1$

An alternative hypothesis states that there is a real association between the variables of interest. For example, PAP is associated with HLA- DR4.

## Fourfold & Contingency Tables

Fourfold and Contingency Tables are used to calculate **observed frequencies, expected frequencies, chi-squared statistic  $\chi^2$  and degree of freedom (df)**.

A chi-squared test is used to determine whether there is an association between two variables, which may be:

- qualitative
- discrete quantitative
- continuous quantitative (where values have been grouped)

Data from two such variables may be arranged in a contingency table. The categories for one variable define the rows, and the categories for the other variable define the columns. Individuals are assigned to the appropriate cell. When the table only has two rows or two columns, the table is called a four-fold table.

### Observed and Expected Frequencies

Observed frequencies are simply the observed number of subjects that apply to each cell. The expected frequencies for each cell are the numbers that we would expect to find if the null hypothesis was true (i.e. no association). To calculate the expected frequency for each cell, we use the observed frequencies. For a given cell, the sum of the cell's row multiplied by the sum of the cell's column, divided by the total number of subjects  $n$ . See the example below:

***Observed frequencies in the sample of 308 patients divided according to presence of PAP and HLA-DR4***

PAP	HLA-DR4		Total
	Present	Absent	
Present	46	28	74
Absent	50	184	234
Total	96	212	308

### Expected frequencies in the sample of 308 patients divided according to presence of PAP and HLA-DR4

PAP	HLA-DR4		Total
	Present	Absent	
Present	23	51	74
Absent	73	161	234
Total	96	212	308

You can see that for the first cell, the expected value was calculated as follows:

$$(46 + 50) \times (46 + 28) \div 308 = 23$$

### Chi-Squared Statistic $\chi^2$

For the alternative hypothesis to be true (i.e. presence of association), the observed and expected frequencies, should be close together, any discrepancy being due to random variation. The best way to this is a chi-squared test, where the observed and expected values are compared by this formula:

$$\chi^2 = \sum (\text{Observed} - \text{Expected})^2 \div \text{Expected}$$

For the example above, we would like to perform the statistical test with 5% level of significance. The test result is  $\chi^2=43.61$ . In order to interpret this chi-squared statistic, the number of degrees of freedom (df) is needed. For a contingency table, df is calculated by the formula:

$$df = (\text{No. of Rows} - 1) \times (\text{No. of Columns} - 1)$$

For the example table, df=1. Using a reference table which shows the percentage points of the  $\chi^2$  distribution, we can see that 43.62 is greater than 3.84 (the critical value for 5% level of significance). Please see this chi-squared distribution reference table:

Degrees of Freedom	Probability										
	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.01	0.001
1	0.004	0.02	0.06	0.15	0.46	1.07	1.64	2.71	3.84	6.64	10.83
2	0.10	0.21	0.45	0.71	1.39	2.41	3.22	4.60	5.99	9.21	13.82
3	0.35	0.58	1.01	1.42	2.37	3.66	4.64	6.25	7.82	11.34	16.27
4	0.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	13.28	18.47
5	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	15.09	20.52
6	1.63	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	16.81	22.46
7	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	18.48	24.32
8	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	20.09	26.12
9	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	21.67	27.88
10	3.94	4.86	6.18	7.27	9.34	11.78	13.44	15.99	18.31	23.21	29.59
	Nonsignificant								Significant		

Thus the probability is less than 5% that such a large observed difference could have risen by chance. Thus we can reject the null hypothesis. The use of chi-squared test is not confined to nominal and ordinal data but also can be used for continuous variables hat have been categorised.

## Links

## Bibliography

BENCKO, Vladimír. *Hygiene & Epidemiology : Selected Chapters*. 2nd edition. Karolinum, 2011. Chapter 6.2: Statistical methodology in epidemiology. ISBN 978-80-246-0793-1.