Comparing means and proportions

Johan A. Elkink
School of Politics & International Relations
University College Dublin

5 November 2018

Outline

One-sample t-test

In some cases, we compare means within the same sample—e.g. the average scores on two different questions, answered by the same individuals. This is a **paired sample t-test**:

$$t = \frac{\bar{d}}{\hat{\sigma}_d / \sqrt{n}},$$

with degrees of freedom $n - 1$, whereby

$$d_i = x_{i1} - x_{i2}$$

and x_1 and x_2 are two different variables. Note that d is the difference and σ_d therefore refers to the standard deviation of the difference, not of either x.

Outline

Two-sample t-test

To compare the means of two independent samples (e.g. men vs women):

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}},$$

with degrees of freedom $min(n_1 - 1, n_2 - 1)$.

A somewhat more precise estimate of the degrees of freedom can be calculated and is used by SPSS.

Two-sample t-test: example

Imagine an experiment of extra teaching for half a class, with x representing the score on a test at the end.

Group	n	$\bar{x}$	$\hat{\sigma}_x$
Treatment	21	51.48	11.01
Control	23	41.52	17.15

Perform a t-test that the extra teaching significantly helped.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} = \frac{51.48 - 41.52}{\sqrt{\frac{(11.01)^2}{21} + \frac{(17.15)^2}{23}}} = 2.31$$

The degrees of freedom are the lowest of $n_1 - 1$ and $n_2 - 1$, i.e. 20. We can look this up in a table to find $0.01 < p < 0.02$.

(Moore, McCabe and Craig, 2012, 432–433)

Confidence interval of the difference

The confidence interval follows then as:

$$CI = (\bar{x}_1 - \bar{x}_2) \pm t^c \cdot \sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}$$

Comparing proportions

Recall, for a dichotomous variable,

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = p(1-p).$$

Thus for two proportions we get:

$$t = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

Only for large samples: see Moore, McCabe and Craig (2012, 588–589) for an adjusted version.

Outline

Chi-squared test

To check for dependence between two categorical variables, we can make use of the chi-squared test (χ^2-test). Based on the **margins** of the table, we can determine the expected values of the cells under independence and then calculate whether the cells differ significantly:

$$\chi^2 = \sum \frac{(n_{observed} - n_{expected})^2}{n_{expected}}$$

The degrees of freedom is calculated as $df = (r-1)(c-1)$, with r the number of rows and c the number of columns of the two-way table.

Outline

Exercise 1

1. Does increased calcium reduce blood pressure?

Treatment	n	$\bar{x}$	$\hat{\sigma}$
Calcium	10	5.000	8.743
Placebo	11	-0.273	5.901

   Also calculate a 95% confidence interval.

2. A movie is rated by male and female visitors to a website.

Rating:	1	2	3	4	5
Female	5	44	48	183	188
Male	5	29	30	91	66

   1. Calculate $\bar{x}$ and $\hat{\sigma}_x$ for each group.
   2. Perform a t-test whether the means differ.

(Moore, McCabe and Craig, 2012, 444, 450)

Exercise 2

Here is data from a study on binge drinking among college students:

	n_{sample}	$n_{drinking}$	$\hat{p} = \frac{n_{drinking}}{n_{sample}}$
Men	5,348	1,392	0.260
Women	8,471	1,748	0.206

Do men binge drink more than women in this sample?

Exercise 3

Fifty census tracts in a city are selected at random. It is found that 20 of these are serviced by community centers, the remainder are not. You compare the delinquency rates for these two types of tracts and obtain the following (per 1,000 adolescents):

	With center	Without center
Sample size	20	30
Mean	27	31
Standard deviation	6	8

Test for the significance of the difference between these two types of tracts.

(Blalock, 1979, 243) (verbatim with minor edits)

Exercise 4

A random sample of married women still living with their husbands is selected and the women are classified as being either "satisfied" or "unsatisfied" with their marital lives. The two groups of women are then compared with respect to the duration of their marriages with the following results:

Duration of marriage	Satisfied	Unsatisfied
0–2	34	10
3–4	41	16
5–9	50	23
10–14	39	25
15–19	18	14
20–39	15	16
total	197	104

Is there a significant difference between the two groups?

(Blalock, 1979, 243) (verbatim)

Exercise 5

For the data on the next slide, create cross-tables and perform χ^2-tests to evaluate:

1. The experimental treatment did reduce the unemployment rate.
2. The above is true regardless of geographic location,
3. and regardless of size.

Exercise 5 – continued

Freq	Treatment	Region	Size	Unemployment
10	Experimental	West	Medium	Increased
30	Experimental	West	Medium	Decreased
50	Experimental	West	Small	Increased
10	Experimental	West	Small	Decreased
15	Experimental	East	Medium	Increased
35	Experimental	East	Medium	Decreased
25	Experimental	East	Small	Increased
5	Experimental	East	Small	Decreased
20	Control	West	Medium	Increased
40	Control	West	Medium	Decreased
20	Control	West	Small	Increased
20	Control	West	Small	Decreased
15	Control	East	Medium	Increased
15	Control	East	Medium	Decreased
25	Control	East	Small	Increased
25	Control	East	Small	Decreased

(Blalock, 1979, 330)

comparing means

One-sample t-test

Two-sample t-test

χ^2-test for crosstables

Exercises

References

Blalock, Hubert M. 1979. *Social statistics*. 2nd ed. Tokyo: McGraw-Hill Kogakusha.

Moore, David S., George P. McCabe and Bruce A. Craig. 2012. *Introduction to the practice of statistics*. 7th international edition ed. New York: W.H. Freeman.