

Tutorial 08, Michaelmas Term

Research Methods for Political Science (PO3600)

Stefan Müller

05 December 2017

Trinity College Dublin

<http://muellerstefan.net/research-methods>

1. Calculate correlations by hand
2. Correlations in SPSS
3. Research proposal/paper

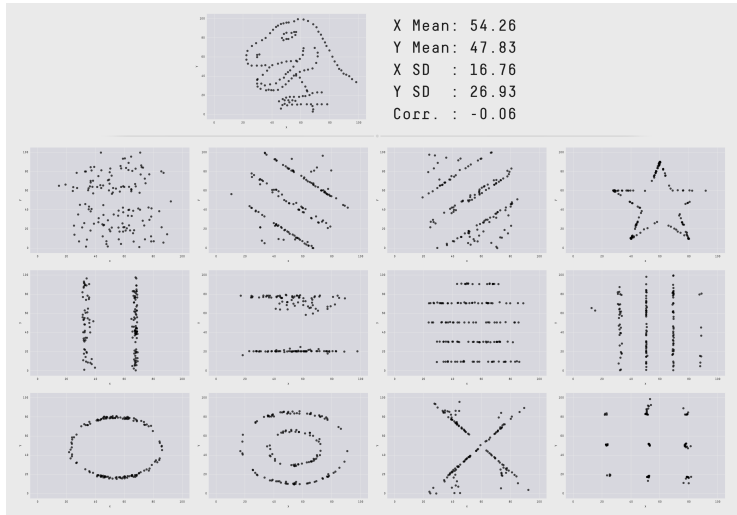
1. Mean, standard deviation, standard error
2. Confidence intervals
3. t-test (one-sample and two-samples)
4. SPSS: Plots, cross-tables, merging datasets, create new variable, summary statistics, methods from above

What is the right test for my hypothesis?

See links under Tutorial 08: <http://muellerstefan.net/research-methods>

- Before you start, always create a scatterplot!
- Correlation and covariation used to measure the association between two interval-ratio variables.

Correlation



Source: <https://www.autodeskresearch.com/publications/samestats>

One step back: variance and covariance

$$\text{variance} = \frac{\sum (x_i - \bar{x})^2}{N-1} = \frac{\sum (x_i - \bar{x})(x_i - \bar{x})}{N-1}$$

$$\text{covariance} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{N-1}$$

Interpreting covariance:

- $\text{cov}(x, y) = 0$: variables not related
- $\text{cov}(x, y) > 0$: variables vary in the same direction
- $\text{cov}(x, y) < 0$: variables vary in the opposite direction

Problem of covariation? It depends on the scales.

Solution? Divide the covariation by the standard deviations of x and y .

$$\text{correlation} = \frac{\text{covariance}}{\text{standard deviation}_x \times \text{standard deviation}_y}$$

Example: Calculate correlation by hand

x	y
1	1
2	3
4	5
5	7

Steps:

1. Calculate mean of each variable
2. Calculate standard deviations of x and y : $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$
3. Calculate covariance: $covariance = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{N-1}$
4. Divide covariance by the product of standard deviations

Example: Calculate correlation

1. Calculate mean of each variable: $\bar{x} = 3$, $\bar{y} = 4$
2. Calculate standard deviations: $\sigma_x = 1.83$ and $\sigma_y = 2.58$
3. Calculate covariance: $covariance = ((1 - 3) \times (1 - 4) + (2 - 3) \times (3 - 4) + (4 - 3) \times (5 - 4) + (5 - 3) \times (7 - 4)) / (4 - 1) = 4.66$
4. Divide covariance by the product of standard deviations:
 $4.66 / (1.83 \times 2.58)$

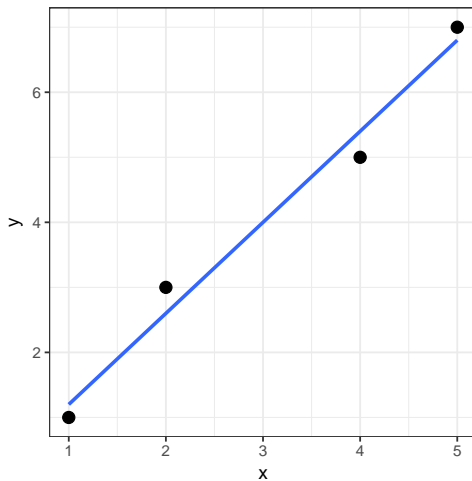
$$\text{Correlation } (r) = 0.99$$

Example: Reproduce analysis in SPSS

1. Insert dataset into SPSS
2. Create scatterplot of x and y
3. Estimate correlation

x	y
1	1
2	3
4	5
5	7

Example: Plot relationship



Example: SPSS output

Correlations

		x	y
x	Pearson Correlation	1	.990 [*]
	Sig. (2-tailed)		.010
	N	4	4
y	Pearson Correlation	.990 [*]	1
	Sig. (2-tailed)	.010	
	N	4	4

^{*}. Correlation is significant at the 0.05 level (2-tailed).

Example: Significance of correlation

$$t_r = \frac{r\sqrt{N-2}}{\sqrt{1-r^2}}$$

$$t_r = \frac{0.98\sqrt{4-2}}{\sqrt{1-0.98^2}}$$

$$t_r = 6.96$$

Conclusion: correlation is statistically significant

Team up and discuss (some of) the following aspects:

1. Research question + relevance
2. Theoretical argument + hypothesis
3. Type of data + operationalisation of variables
4. Ways of analysing your data

Discuss with your neighbour:

1. What is the puzzle you address?
2. Why is the question relevant?
3. What are your key dependent and independent variables?
4. What data will you use?