DISCOVERING STATISTICS USING SPSS

THIRD EDITION

Comparing Two Means

Lecture 08

ANDY FIELD

Aims

- T-tests
  - Dependent (aka paired, matched)
  - Independent
- Rationale for the tests
  - Assumptions
- Interpretation
- Reporting results
- Calculating an Effect Size
- *T*-tests as a GLM

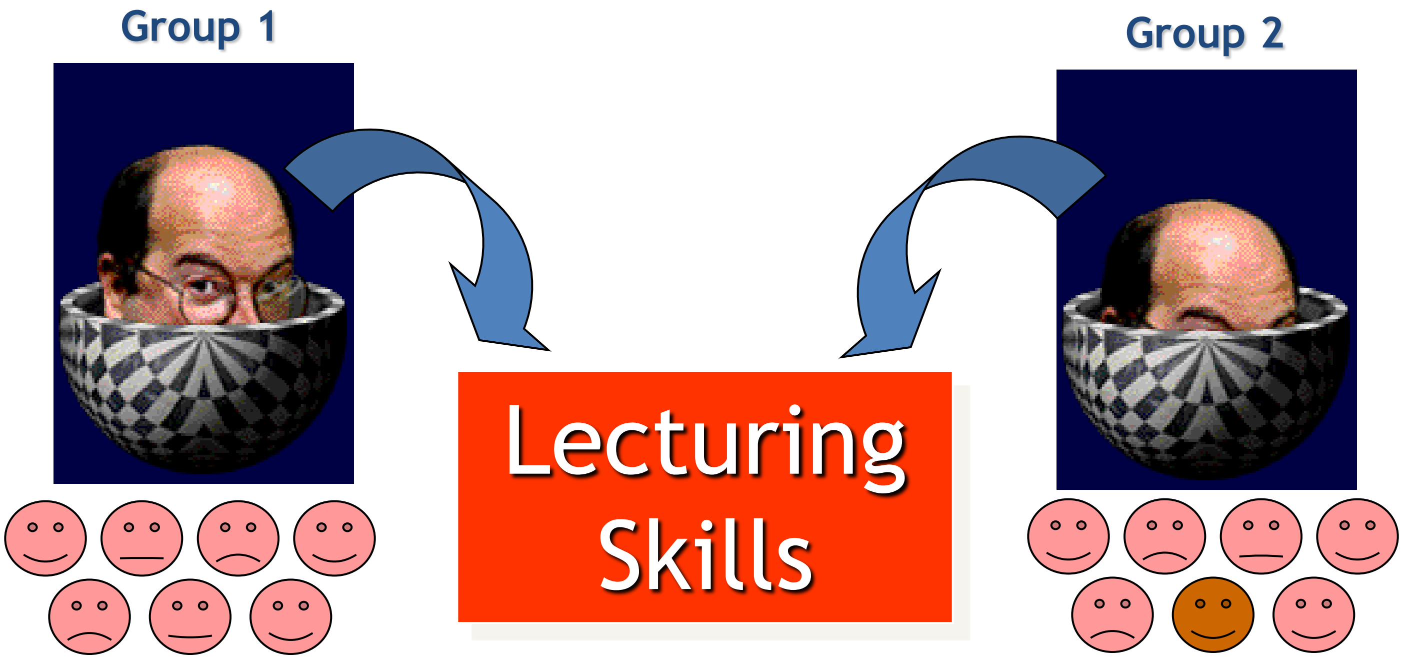
Experiments

- The simplest form of experiment that can be done is one with only one independent variable that is manipulated in only two ways and only one outcome is measured.
  - More often than not the manipulation of the independent variable involves having an experimental condition and a control.
  - E.g., Is the movie *Scream 2* scarier than the original *Scream*? We could measure heart rates (which indicate anxiety) during both films and compare them.
- This situation can be analysed with a *t*-test

T-test

- Dependent *t*-test
  - Compares two means based on related data.
  - E.g., Data from the same people measured at different times.
  - Data from 'matched' samples.
- Independent *t*-test
  - Compares two means based on independent data
  - E.g., data from different groups of people
- Significance testing
  - Testing the significance of *Pearson's correlation coefficient*
  - Testing the significance of *b* in regression.

Rationale to Experiments



- Variance created by our manipulation
  - Removal of brain (systematic variance)
- Variance created by unknown factors
  - E.g. Differences in ability (unsystematic variance)

Rational for the *t*-test

- Two samples of data are collected and the sample means calculated. These means might differ by either a little or a lot.
- If the samples come from the same population, then we expect their means to be roughly equal. Although it is possible for their means to differ by chance alone, we would expect large differences between sample means to occur very infrequently.
- We compare the difference between the sample means that we collected to the difference between the sample means that we would expect to obtain if there were no effect (i.e. if the null hypothesis were true). We use the standard error as a gauge of the variability between sample means. If the difference between the samples we have collected is larger than what we would expect based on the standard error then we can assume one of two:
  - There is no effect and sample means in our population fluctuate a lot and we have, by chance, collected two samples that are atypical of the population from which they came.
  - The two samples come from different populations but are typical of their respective parent population. In this scenario, the difference between samples represents a genuine difference between the samples (and so the null hypothesis is incorrect).
- As the observed difference between the sample means gets larger, the more confident we become that the second explanation is correct (i.e. that the null hypothesis should be rejected). If the null hypothesis is incorrect, then we gain confidence that the two sample means differ because of the different experimental manipulation imposed on each sample.

Rationale to the *t*-test

$$t = \frac{\begin{array}{c}\text{observed difference}\\ \text{between sample means}\end{array} - \begin{array}{l}\text{expected difference}\\ \text{between population means}\\ \text{(if null hypothesis is true)}\end{array}}{\begin{array}{c}\text{estimate of the standard error of the difference between two}\\ \text{sample means}\end{array}}$$

Assumptions of the *t*-test

- Both the independent *t*-test and the dependent *t*-test are *parametric tests* based on the normal distribution. Therefore, they assume:
  - The sampling distribution is normally distributed. In the dependent *t*-test this means that the sampling distribution of the *differences* between scores should be normal, not the scores themselves.
  - Data are measured at least at the interval level.
- The independent *t*-test, because it is used to test different groups of people, also assumes:
  - Variances in these populations are roughly equal (*homogeneity of variance*).
  - Scores in different treatment conditions are independent (because they come from different people).

The Dependent *t*-test

$$t = \frac{\overline{D} - \mu_D}{s_D / \sqrt{N}}$$

Example

- Is arachnophobia (fear of spiders) specific to real spiders or is a picture enough?
- Participants
  - 12 spider phobic individuals
- Manipulation
  - Each participant was exposed to a real spider and a picture of the same spider at two points in time.
- Outcome
  - Anxiety

Dependent *t*-test Output

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Picture of Spider	40.00	12	9.293	2.683
	Real Spider	47.00	12	11.029	3.184

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Picture of Spider & Real Spider	12	.545	.067

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper			
Pair 1	Picture of Spider - Real Spider	-7.000	9.807	2.831	-13.231	-.769	-2.473	11	.031

Reporting the Results

- On average, participants experienced significantly greater anxiety to real spiders (M = 47.00, SE = 3.18) than to pictures of spiders (M = 40.00, SE = 2.68), $t(11) = -2.47$, $p < .05$

The Independent *t*-test

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Example

- Is arachnophobia (fear of spiders) specific to real spiders or is a picture enough?
- Participants
  - 24 spider phobic individuals
- Manipulation
  - 12 participants were exposed to a real spider
  - 12 were exposed to a picture of the same spider.
- Outcome
  - Anxiety

Independent *t*-test Output

Group Statistics

	Spider or Picture?	N	Mean	Std. Deviation	Std. Error Mean
Anxiety	Picture	12	40.00	9.293	2.683
	Real Spider	12	47.00	11.029	3.184

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Anxiety	Equal variances assumed	.782	.386	-1.681	22	.107	-7.000	4.163	-15.634	1.634
	Equal variances not assumed			-1.681	21.385	.107	-7.000	4.163	-15.649	1.649

Calculating an Effect Size

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

$$r = \sqrt{\frac{-1.681^2}{-1.681^2 + 22}}$$

$$= \sqrt{\frac{2.826}{24.826}}$$

$$= 0.34$$

Reporting the Results

- On average, participants experienced greater anxiety to real spiders (*M* = 47.00, *SE* = 3.18), than to pictures of spiders (*M* = 40.00, *SE* = 2.68. This difference was not significant $t(22) = -1.68$, $p > .05$; however, it did represent a medium-sized effect $r = .34$.

The *t*-test as a GLM

$$A_i = b_0 + b_1 G_i + \varepsilon_i$$

$$\text{anxiety}_{\text{i}} = b_0 + b_1 \text{group}_{\text{i}} + \varepsilon_i$$

Picture group

- The group variable = 0
- Intercept = mean of baseline group

$$\bar{X}_{\text{Picture}} = b_0 + (b_1 \times 0)$$

$$b_0 = \bar{X}_{\text{Picture}}$$

$$b_0 = 40$$

Real Spider Group

- The group variable = 1
- b_1 = Difference between means

$$\bar{X}_{\text{Real}} = b_0 + (b_1 \times 1)$$

$$\bar{X}_{\text{Real}} = \bar{X}_{\text{Picture}} + b_1$$

$$b_1 = \bar{X}_{\text{Real}} - \bar{X}_{\text{Picture}}$$

$$= 47 - 40$$

$$= 7$$

Output from a Regression

When Assumptions are Broken

- Dependent *t*-test
  - Mann-Whitney Test
  - Wilcoxon rank-sum test
- Independent *t*-test
  - Wilcoxon Signed-Rank Test
- Robust Tests:
  - Bootstrapping
  - Trimmed means