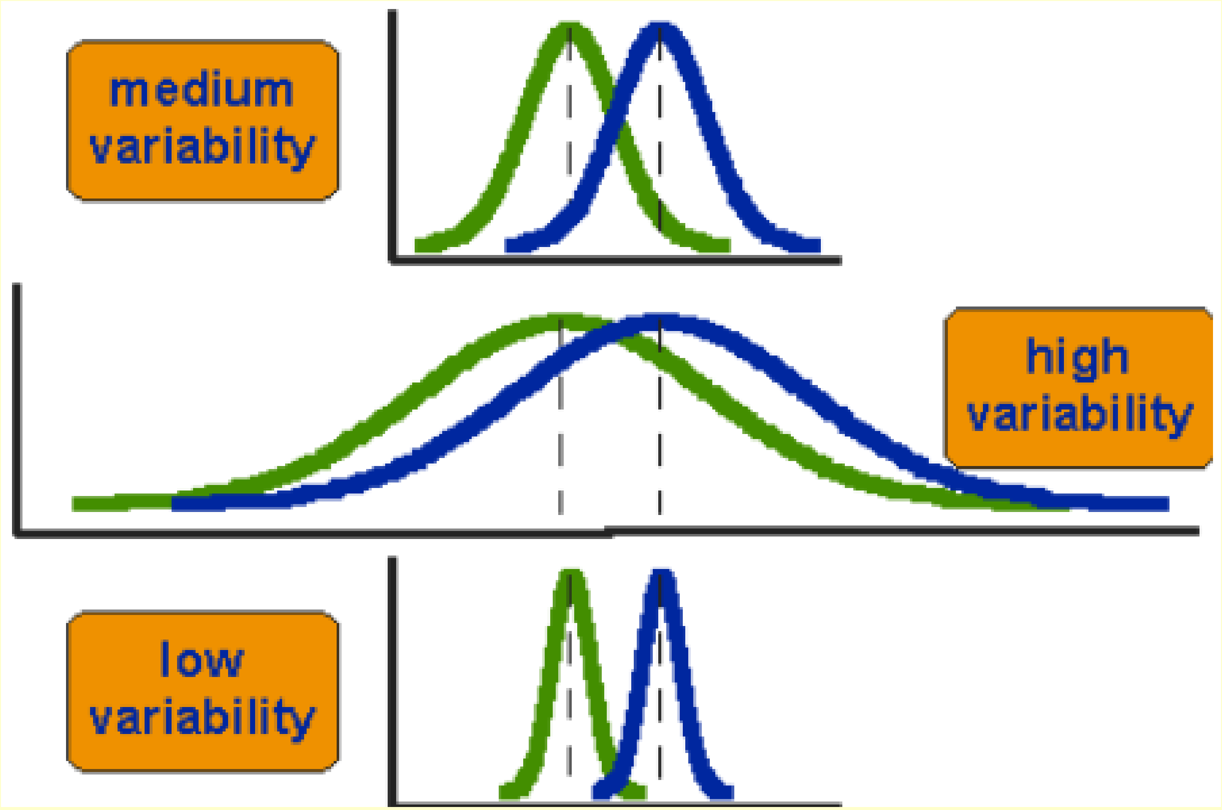
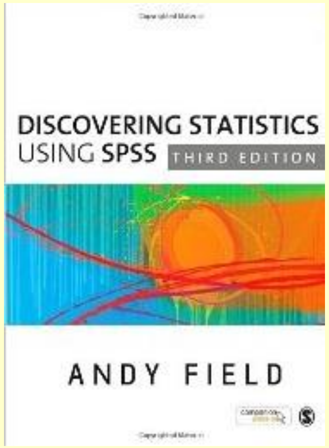
Comparing Two Means



http://www.pelagicos.net/classes_biometry_fa16.htm

Reading - Field: Chapter 9



AIMS

- Learn how to do T-tests
  - Dependent (aka paired, matched)
  - Independent

- Review the Test Assumptions
- Learn How to Interpret Test Results
- Learn How to Report Test Results

- Calculate the Effect Size

Correlational vs. Experimental Science

Correlations and Regressions quantify the associations between variables using the "natural variation" in the data (observations).

Science can also use controlled experiments: whereby one or more factors are changed – while keeping the others unchanged – and the resulting changes in response variables are quantified.

The simplest experiment entails comparing two samples (representative of two populations).

Experiments: Comparing Groups

- The simplest experiment involves the following:
  one independent variable is modified in only one way and only one outcome is measured.
  - Comparisons involve two samples (populations).
    (e.g., Ant densities in a field or a forest ?)
  - Or the manipulation involves comparing an experimental condition and a control group.
    (e.g., give two groups of pigs diet A or diet B)
- These situations can be analysed with a *t*-test

T-test

- Independent *t*-test (or just plain T-test)
  - Compares two means from independent observations – values are not connected.

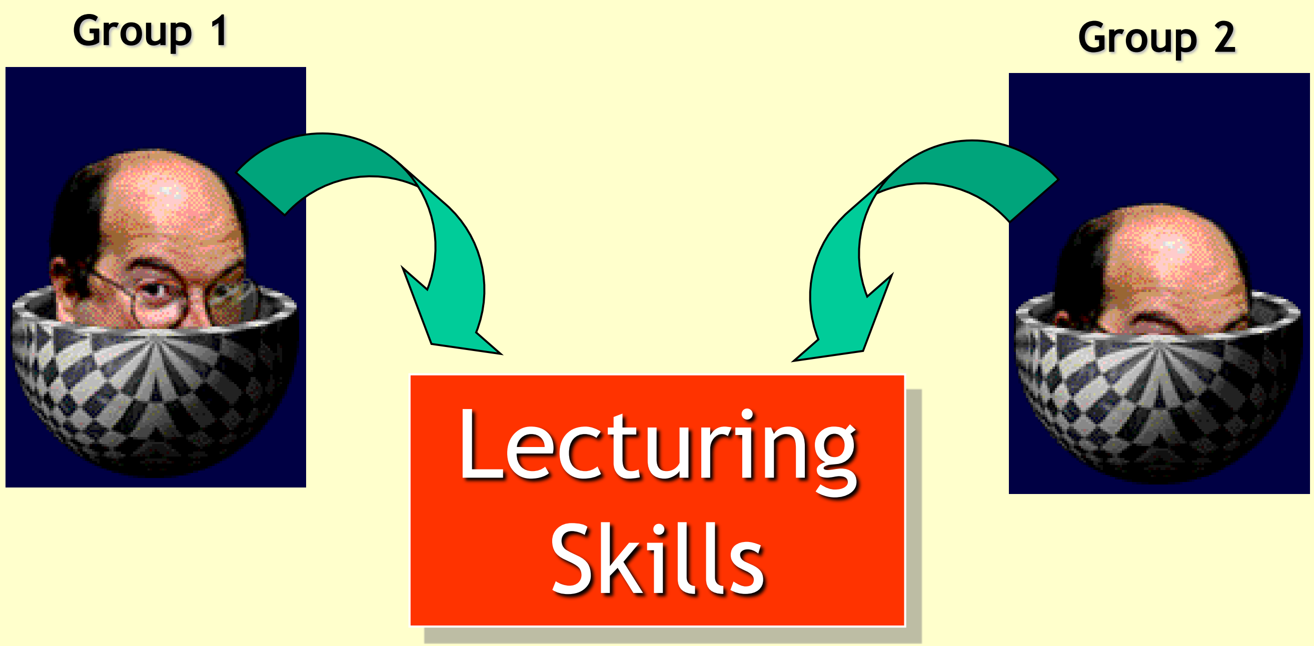
    (e.g., data from different groups of people).

- Dependent *t*-test (or paired T-test)
  - Compares two means based on related observations – they are not independent.

    (e.g., data from same individuals).

    Note: These data are from paired samples.

Example of An Experiment

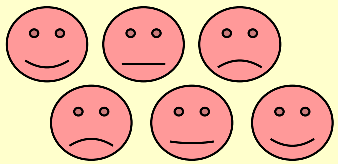


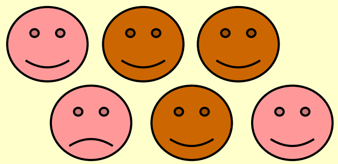
- Need to Balance two Sources of Variation:
  - Variance created by our manipulation
  - Brain Transplant (systematic variance)
  - Variance created by unknown factors
  - Differences in ability (unsystematic variance)

First Approach- Independent Samples

1. Randomly assign individuals to one treatment:

Group 1





Group 2

2. Measure the Response Variable:

Standardized metrics
Same audience

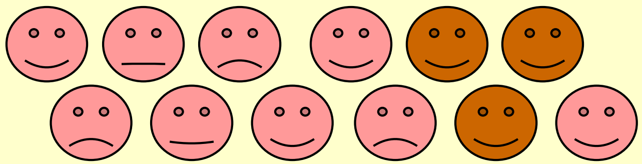
Lecturing Skills

3. Compare Response Variable in Two Groups:

6 measurements (group 1)
6 measurements (group 2)

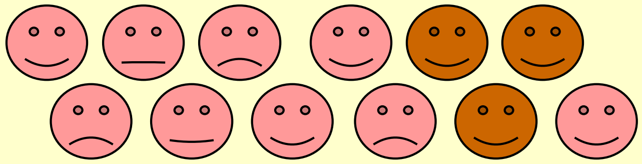
Second Approach- Dependent Samples

1. Each individual is measured (control) :



Lecturing Skills

2. Each individual is manipulated (experiment):



Lecturing Skills

3. Compare Response Variable in Each Individual:
12 individual measurements (control / manipulation)

Rationale for the T-test

- Two samples collected and means calculated.
- If samples come from the same population, we expect their means to be roughly equal.
- While these means may differ by chance alone, large differences occur infrequently.
- Compare the observed difference between the sample means that we collected and the difference between the sample means we would expect if there were no effect (i.e., if the null hypothesis were true).

Rationale for the T-test

- We use the standard error to gauge the variability between sample means. If the difference between the samples is larger than what we would expect based on the standard error then we can assume one of two:
  - There is no effect and we have, by chance, collected two samples that are atypical of the origin population (type I error)
  - The two samples come from two different populations. The difference between samples represents a genuine difference (reject the incorrect null hypothesis)

Rationale for the T-test

- As the observed difference between the sample means gets larger, we are more confident that second explanation is correct (i.e., the null hypothesis should be rejected).

- In an experiment, if we reject the null hypothesis, we infer that the two sample means differ because of the different experimental manipulation we imposed on each sample.

- NOTE: This requires random assignment of individuals to ensure there are no biases.

Generic T-test Statistic

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

$$t = \frac{\text{Observed (Mean1 – Mean 2)} - \text{Expected (Mean1 – Mean 2)}}{\text{Estimated Standard Error of the Difference (Mean1 – Mean 2)}}$$

Assumptions of the T-test

- The independent and dependent t-tests are *parametric tests* based on normal distributions.

- Therefore, they assume:

  Observations are randomly sampled.

  Data are on the interval or ratio scale .

  Sampling distribution is normally distributed.

  NOTE: In the dependent t test, the sampling distribution *of the differences* between scores is normal, not the scores themselves.

Assumptions of the T-test

- Because the independent *t*-test compares different groups of observations, it assumes:
  - Variances in these populations are equal (*homogeneity of sample variances*).
  - Scores in different treatment conditions are independent (because they come from different individuals – not paired observations).

Dependent (Paired) T-test

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

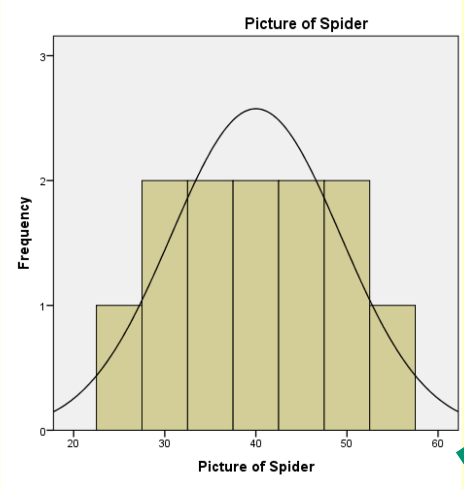
$$\mathbf{t} = \frac{\text{Observed Mean Difference} - \text{Expected Mean Difference}}{\text{Standard Error}}$$

Dependent (Paired) T-test

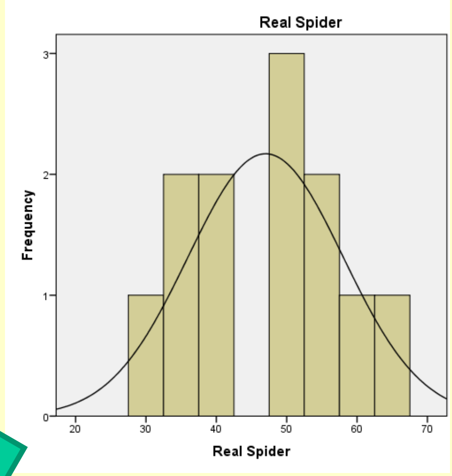
Example: Is arachnophobia specific to real spiders or is seeing a picture enough?

- Participants: 12 spider phobic individuals
- Manipulation: Participants exposed to two treatments: (at two points in time): a real spider and a picture of that same spider
- Outcome: Anxiety (Heart rate)

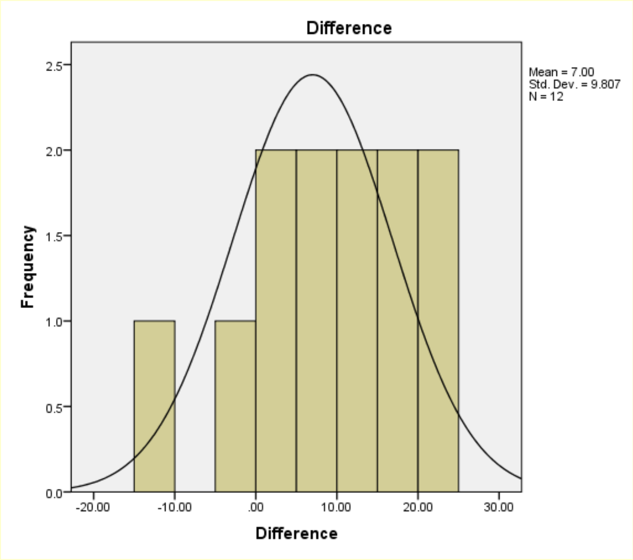
Paired T-test



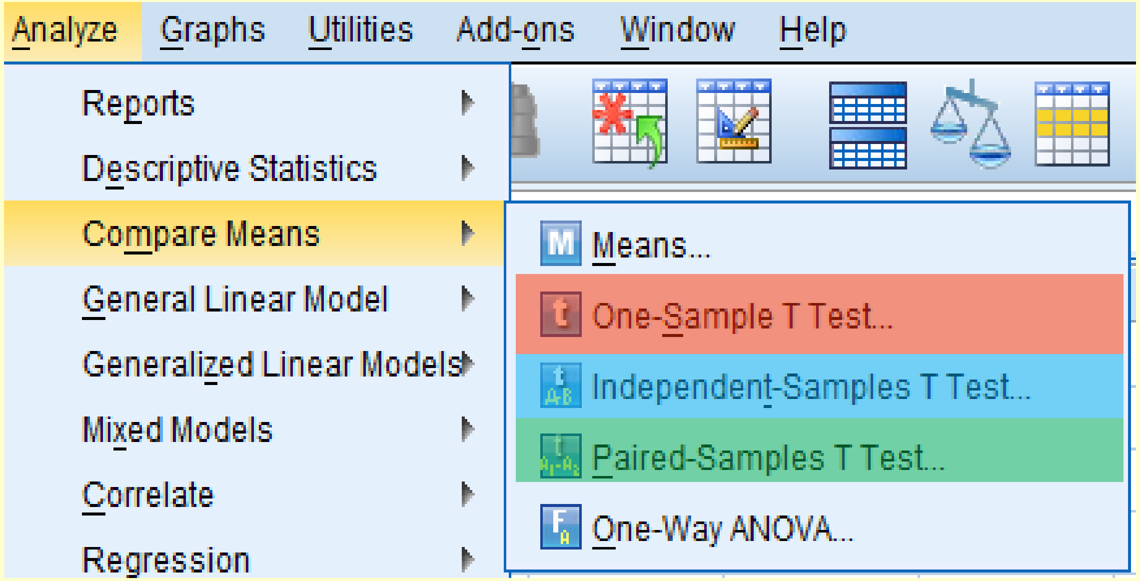
picture	real	Difference
30	40	10.00
35	35	.00
45	50	5.00
40	55	15.00
50	65	15.00
35	55	20.00
55	50	-5.00
25	35	10.00
30	30	.00
45	50	5.00
40	60	20.00
50	39	-11.00



(Real - Photo)



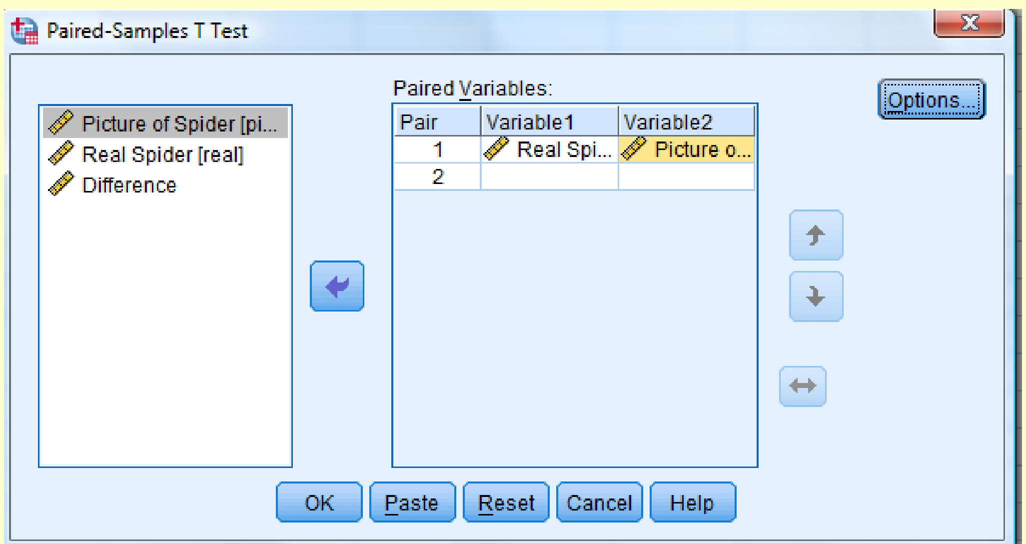
Performing a T-test on SPSS



Compare sample mean against theoretical mean

Independent t-test

Dependent t-test



Move variables to paired observations "drop zones"

Dependent T-test Output

Paired Samples Statistics

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

Means
SDs
SEs

Paired Samples Correlations

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

Correlation

Paired Samples Test

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

Difference: Mean, SD, SE

95% CI

df = n - 1

$t = (7 - 0) / 2.831 = 2.471$

($p = 0.031$)

Reporting Dependent T-test 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), as revealed by a paired t-test (*t* = 2.47, df = 11, *p* = 0.031).

$t = (+7 - 0) / 2.831 = +2.471$

$t = (-7 - 0) / 2.831 = -2.471$

t	df	Sig. (2-tailed)
2.473	11	.031

NOTE: Sign of t test statistic does not matter in the two tailed test. However, it does matter for the one tailed test.

Independent T-test

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

$$\frac{\text{Observed Mean 1} - \text{Observed Mean 2}}{\text{Sqrt (Pooled Sample Variance)}}$$

Sample Pooled Variance:

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

Independent (Unpaired) T-test

Example: Is arachnophobia specific to real spiders or is seeing a picture enough?

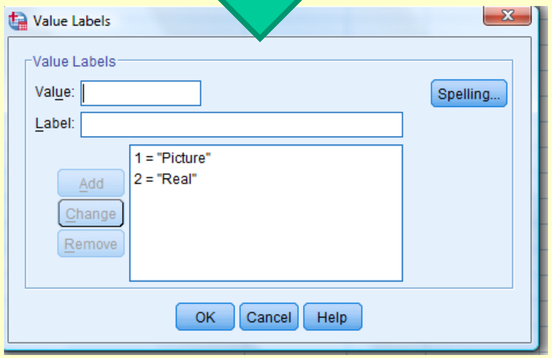
- Participants: 24 spider phobic individuals
- Manipulation: Participants exposed to one of two treatments: 12 to a real spider and a 12 to a picture of that same spider
- Outcome: Anxiety (Heart rate)

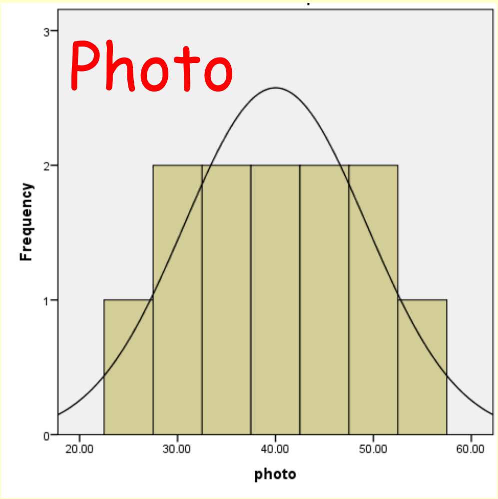
Independent T-test

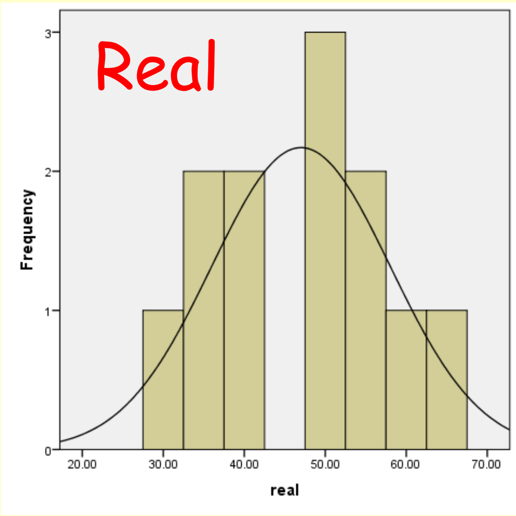
Name	Type	Width	Decimals	Label	Values
Group	Numeric	8	0	Spider or Picture?	{0, Picture}...

Group	Anxiety
1	30
1	35
1	45
1	40
1	50
1	35
1	55
1	25
1	30
1	45
1	40
1	50
2	40
2	35
2	50
2	55
2	65
2	55
2	50
2	35
2	30
2	50
2	60
2	39

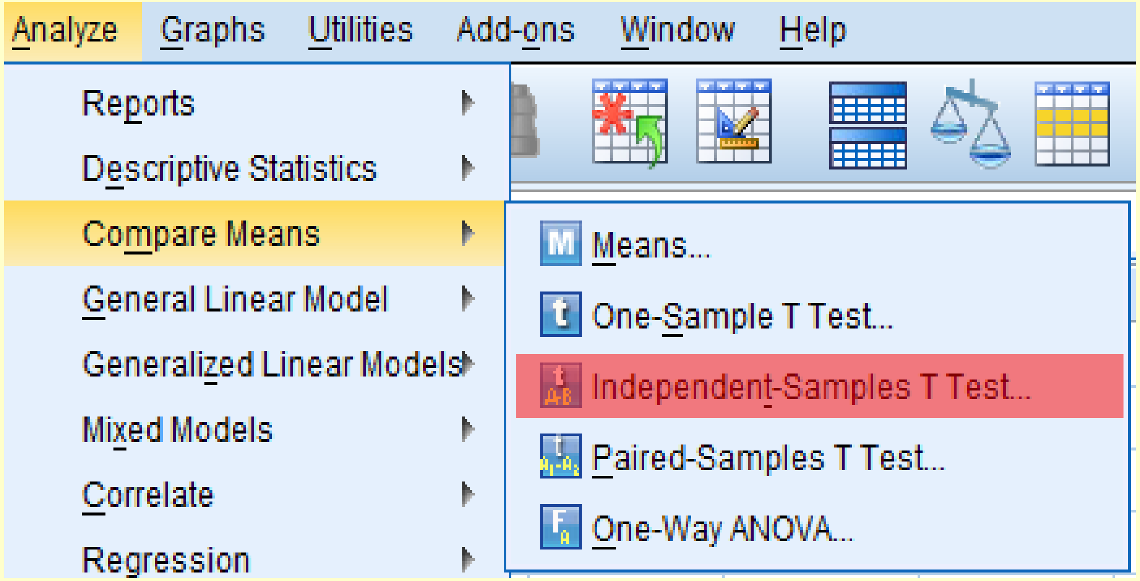
Groups
1 & 2



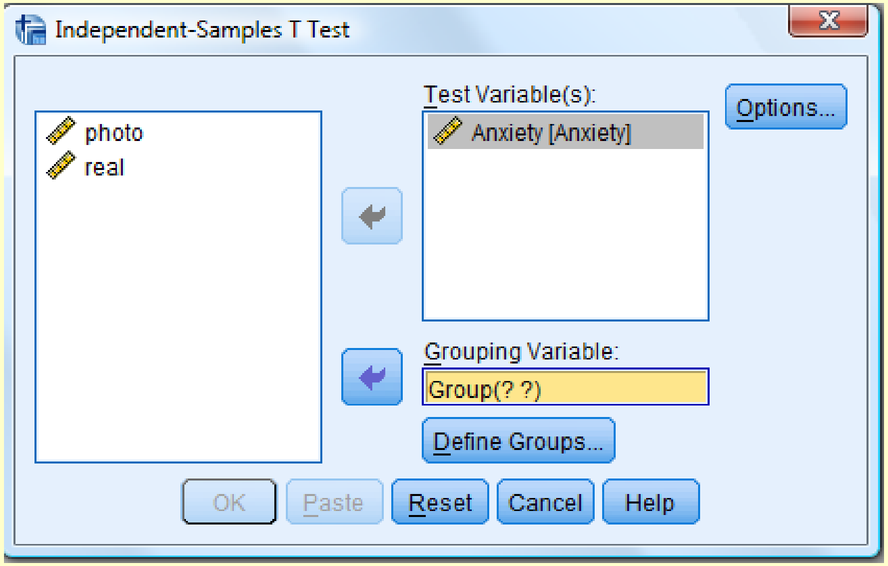




Performing a T-test on SPSS



Independent t-test



Define test variable (dependent variable) and grouping variable (group or treatment)

Independent (Paired) T-test

Example: Is arachnophobia specific to real spiders or is seeing a picture enough?

- Participants: 24 spider phobic individuals
- Manipulation: Participants exposed to one of two treatments: 12 to a real spider and a 12 to a picture of that same spider
- Outcome: Anxiety (Heart rate)

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	12	47.00	11.029	3.184

Means
SDs
SEs

Equality of Variances

		Levene's Test for Equality of Variances	
		F	Sig.
Anxiety	Equal variances assumed	.782	.386
	Equal variances not assumed		

p value = 0.386
(Not Significant)

(Variances Equal)

Independent Samples Test

Equality of Means

		t-test for Equality of Means						
							95% Confidence Interval of the Difference	
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Anxiety	Equal variances assumed	-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 the Effect Size

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

t = t statistic (difference of means)

df = degrees of freedom

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

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

$$= 0.34$$

If t = 0 (no difference)
Then r ~ 0

If t = ∞ (large difference)
Then r ~ 1

Reporting Dependent T-test 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). While a t- test did not reveal a statistically significant difference (*t* = −1.68, df = 22, *p* = 0.107); had a medium-sized effect (*r* = 0.34).

$t = (+7 - 0) / 4.163 = +1.681$

$t = (-7 - 0) / 4.163 = -1.681$

Independent Samples Test

t	df	Sig. (2-tailed)
-1.681	22	.107
-1.681	21.385	.107

NOTE: Sign of t test statistic does not matter in the two tailed test. However, it does matter for the one tailed test.

T-test Summary

- T-tests allow us to analyze simple experiments involving only one independent variable manipulated in two ways and only one measured outcome variable.
- T-tests can be independent or dependent (paired); on the basis of whether the same individuals are measured one or two times.
- Independent T-tests should allocate individuals to treatments randomly to avoid individual biases.
- Whenever possible, use paired tests because they allow you to control for individual-level variation.

When Assumptions are Not Met

- Dependent *t*-test
  - Mann-Whitney Test (medians)

- Independent *t*-test
  - Wilcoxon Signed-Rank Test (medians)