Research Design - - Topic 2
Inferential Statistics: The t-test

- General Rationale Underlying the t-test (Gardner & Tremblay, 2007, Ch. 2)
- The Independent t-test
- The Correlated (paired) t-test
- Effect Size and Power (Kirk, 1995, pp 58-64; Cohen, 1988, Ch. 2)

1

Independent t-test

- Single Sample t-test (Gosset, "Student",1908)

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

- Two sample t-test (Fisher,1925)

$$t = \frac{(\bar{X}_1 - \mu_1) - (\bar{X}_2 - \mu_2)}{\text{standard error of the difference}}$$

$$t = \frac{(\bar{X}_1 - \mu_1) - (\bar{X}_2 - \mu_2)}{\sqrt{S^2_{\bar{X}_1 - \bar{X}_2}}}$$

When H_o: True

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S^2_{\bar{X}_1^2 - \bar{X}_2^2}}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S^2_{\bar{X}_1} + S^2_{\bar{X}_2}}}$$

2

If variances are heterogeneous

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

and degrees of freedom are estimated using the Welch estimate

If variances are homogeneous, compute a pooled estimate

$$S_p^2 = \frac{\sum(X_1 - \bar{X}_1)^2 + \sum(X_2 - \bar{X}_2)^2}{n_1 + n_2 - 2}$$

Then:

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

with degrees of freedom = n1 + n2 - 2

3

Tests for Heterogeneity of Variance

Levene's (1960) test of Heterogeneity of Variance involves an analysis of variance of the absolute deviations of each score from its group mean. If the mean absolute deviations differ significantly in this two group case, it suggests that the variances differ significantly. Under this condition, one should use the t-test for independent variances; otherwise, the t-test with pooled estimates should be used.

Degrees of Freedom

- Welch (1938) degrees of freedom for independent variance estimates

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{S_1^4}{n_1^2(n_1 - 1)} + \frac{S_2^4}{n_2^2(n_2 - 1)}}$$

- Degrees of freedom for pooled variance estimate

$$df = n_1 + n_2 - 2$$

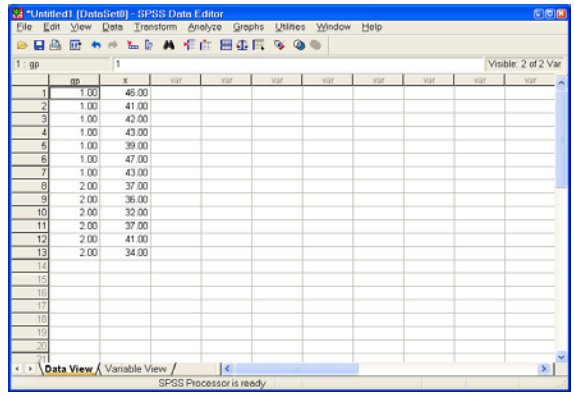
4

Data for the Independent t-test

	Group	
	1	2
	46	37
	41	36
	42	32
	43	37
	39	41
	47	34
	43	
Mean	43.00	36.17
Standard Deviation	2.77	3.06

5

Data Editor for the Independent t-test

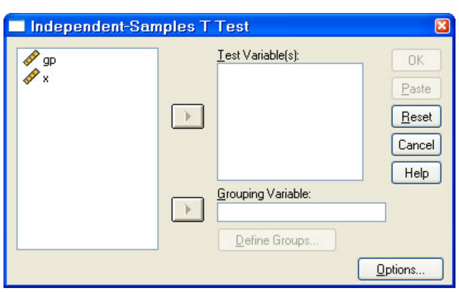


6

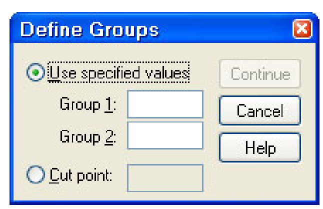
Using CLOPE to run SPSS t-test

Clope = Click and hope that you do what you want to do.

Enter SPSS, Put data in the Data Editor, Click on: Analyze → Compare Means →Independent-Samples T test. This presents the following window



7



When you specify the values identifying data in groups 1 and 2, the Continue box will darken, and when you click it, the program returns to the previous window. Clicking on OK produces the following results.

8

SPSS Run for an Independent t-test

```
GET
  FILE='F:\PSYCH540\dataforindependentttest.sav'.
DATASET NAME DataSet1 WINDOW=FRONT.
T-TEST
  GROUPS = gp(1 2)
  /MISSING = ANALYSIS
  /VARIABLES = x
  /CRITERIA = CI(.95) .
```

Group Statistics

	gp	N	Mean	Std. Deviation	Std. Error Mean
x	1.00	7	43.0000	2.76887	1.04654
	2.00	6	36.1667	3.06050	1.24944

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
x	Equal variances assumed	.027	.873	4.228	11	.001	6.83333	1.61623	3.27604	10.39063
	Equal variances not assumed			4.193	10.266	.002	6.83333	1.62983	3.21456	10.45211

Assumptions: Independent t-test

- Independent Random Sampling: The samples are independently and randomly obtained from the populations of interest.
- Normality: The two populations are each normally distributed.
- Homogeneity of variance: The variances are equal in the two populations.

Null Hypothesis: The population means are identical. That is:

H_0: $\mu_1 = \mu_2$

10

Paired t-test

$$t = \frac{(\overline{X}_1 - \mu_1) - (\overline{X}_2 - \mu_2)}{\sqrt{S^2_{\overline{X}_1 - \overline{X}_2}}}$$

But the data are correlated, thus:

$$S^2_{\overline{X}_1 - \overline{X}_2} = S^2_{\overline{X}_1} + S^2_{\overline{X}_2} - 2r_{X_1X_2} S_{\overline{X}_1} S_{\overline{X}_2}$$

Therefore

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{n} - \frac{2r_{12}S_1S_2}{n}}} = \frac{\overline{d}}{\frac{S_d}{\sqrt{n}}}$$

where $df = n - 1$

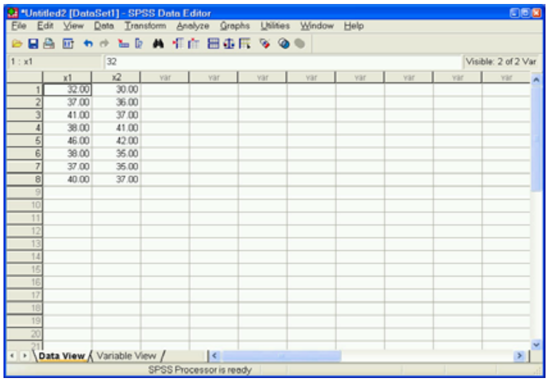
11

Data for the Paired t-test

	X1	X2	d
	32	30	2
	37	36	1
	41	37	4
	38	41	-3
	46	42	4
	38	35	3
	37	35	2
	40	37	3
Mean	38.63	36.63	2.00

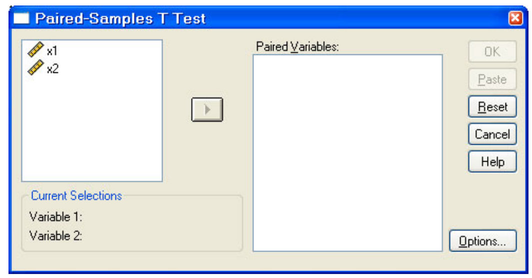
12

Data Editor for the Paired t-test



Using CLOPE to run the Paired t-test

Enter SPSS, Put data in the Data Editor, Click on: Analyze → Compare Means →Paired-Samples T test. This presents the following window



SPSS Run for a Paired t-test

```
GET
  FILE='F:\PSYCH540\dataforpairedttest.sav'.
DATASET NAME DataSet2 WINDOW=FRONT.
T-TEST
  PAIRS = x1  WITH x2 (PAIRED)
  /CRITERIA = CI(.95)
  /MISSING = ANALYSIS.
```

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	x1	38.6250	8	3.99777	1.41342
	x2	36.6250	8	3.73927	1.32203

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	x1 & x2	8	.830	.011

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	x1 - x2	2.00000	2.26779	.80178	.10408	3.89592	2.494	7	.041

Assumptions: Paired t-test

- Independent random sampling: The pairs of observations are independently and randomly obtained from the population of interest.
- Normality: The differences between the pairs of observations are normally distributed in the population of differences.

Null Hypothesis: The mean difference in the population is 0. That is:

H_0: $\mu_d = 0$

Or its equivalent:

H_0: $\mu_1 - \mu_2 = 0$

17

Effect Size and Power

- Cohen (1988) states "… it is convenient to use the term "effect size" to mean the "degree to which the phenomenon is present in the population" or "the degree to which the null hypothesis is false"." (p.10-11). With respect to the t-test, he proposed:

$$d = \frac{|\mu_1 - \mu_2|}{\sigma}$$

Where: **Small** d = .20; **Medium** d = .50; **Large** d = .80

Estimate for the Independent t-test

$$d = \frac{|\bar{X}_1 - \bar{X}_2|}{S_{pooled}} = |t| \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Estimate for the paired t-test

$$d = \frac{|\bar{X}_1 - \bar{X}_2|}{S_d} = |t| \frac{1}{\sqrt{n}}$$

18

Power estimates can be obtained using the Cohen (1988) Text or computed using the GPower3.1 program which can be downloaded from:

http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/

GPower 3.1 calculates power estimates for most statistics of interest to psychologists. It has two types of application:

1. Posthoc permits one to determine the power associated with a given sample and effect size.
2. A priori permits one to determine the sample size for a given power and effect size (not available for all procedures).

19

References

Cohen, J. (1988). *Statistical Power for the Behavioral Sciences* (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum.

Fisher, R.A. (1925). Applications of "Student's" distribution. *Metron*, 5, 90-104.

Levene, H. (1960). Robust tests for equality of variances. In I. Olkins (ed.) *Contributions to probability and statistics.* Stanford, CA: Stanford University Press.

"Student" (1908) The probable error of a mean. *Biometrika, 6*, 1-25.

Welch, B.L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika, 29,* 350-362.

20