EDUR 8131
Chat 7

Notes 5c Correlated Samples t-test
Notes 6 Correlation

1. Notes 5c: Paired-samples t-test with APA presentation

Does an instrument designed to measure academic self-efficacy administered to a group of participants twice, one week apart, show similar mean scores on academic self-efficacy? Note that the scale provides a self-efficacy score that ranges from 1 = low to 7 = high.

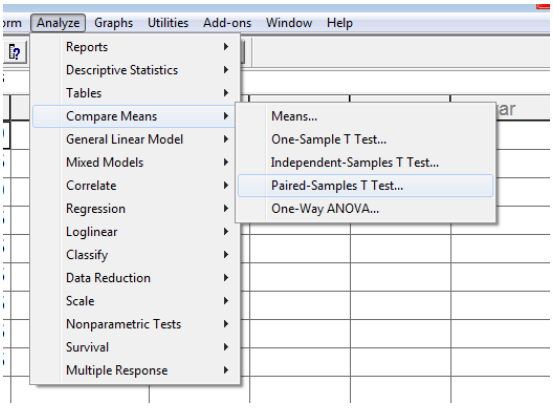
Student	Self-efficacy Time 1	Self-efficacy Time 2
A	6.00	6.25
B	5.25	5.75
C	3.50	4.50
D	4.25	5.00
E	2.75	2.75
F	4.75	5.25
G	5.25	5.00
H	6.25	6.00
I	3.75	4.25

SPSS Results (set alpha to .01, obtain 99% CI)

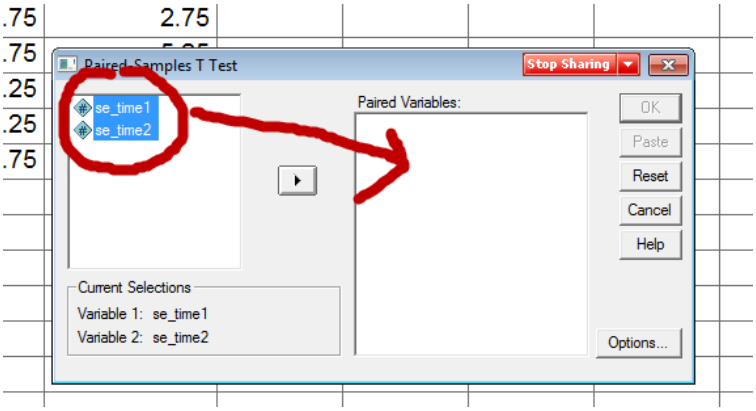
Data entry in SPSS (two columns, one for each time or data column)

	se_time1	se_time2
1	6.00	6.25
2	5.25	5.75
3	3.50	4.50
4	4.25	5.00
5	2.75	2.75
6	4.75	5.25
7	5.25	5.00
8	6.25	6.00
9	3.75	4.25
10		

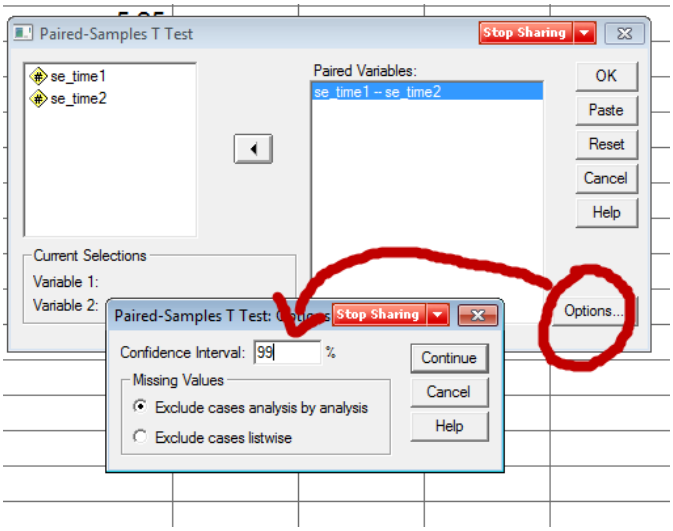
SPSS Commands



Select two columns of data to compare then move both over to paired variables box.



Select Options to obtain 99% confidence interval (change interval from 95 which is default to 99)



First, what would be the null hypothesis for this study?

Second, if we set alpha = .01, would we reject or fail to reject Ho for this example?

SPSS Output

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	se_time1	4.6389	9	1.17334	.39111
	se_time2	4.9722	9	1.06393	.35464

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	se_time1 & se_time2	9	.930	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	99% Confidence Interval of the Difference Lower	Upper			
Pair 1	se_time1 - se_time2	-.33333	.43301	.14434	-.81764	.15098	-2.309	8	.050

What is the null for these data?

Written:

> There is no difference in mean self-efficacy scores between the first and second administration of the self-efficacy measure.

Symbolic:

> Ho: $\mu_1 = \mu_2$
> or
> Ho: $\mu_1 - \mu_2 = 0.00$

Would we reject or fail to reject given the SPSS output?

Since p = .05 and this is larger than alpha = .01 one would fail to reject.

How could we use the confidence interval to test the null hypothesis of no mean difference?

Since 0.00 lies within the 99% CI, one would fail to reject since 0.00 is one of the possible values for the mean difference between groups.

APA Style Presentation
Alpha = .01

Table 7
Results of t-test and Descriptive Statistics for Academic Self-efficacy Over Time

	First Admin.		Second Admin.			99% CI for Mean Difference			
Outcome	M	SD	M	SD	n		r	t	df
Aca. Self-efficacy	4.64	1.17	4.97	1.06	9	-0.82, 0.15	.93*	-2.31	8

* p < .01.

What wording would we use to indicate hypothesis testing results here; i.e., what wording would we use for **inference**?

> Results show that there is not a statistically significant mean difference, at the .01 level, between the first administration of the instrument and the second administration for academic self-efficacy.

What wording would we use for **interpretation** of results?

> Results show that there is not a statistically significant mean difference, at the .01 level, between the first administration of the instrument and the second administration for academic self-efficacy. Mean academic self-efficacy appears to be similar for participants in both administrations of the instrument.

What happens if we change alpha from .01 to alpha = .05?

Would we reach same conclusion as that provided above?
What information was used to decide if Ho rejected: p? CI? Something else?

SPSS Results for alpha = .01

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	selfefficacy1	4.6389	9	1.17334	.39111
	selfefficacy2	4.9722	9	1.06393	.35464

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	selfefficacy1 & selfefficacy2	9	.930	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	99% Confidence Interval of the Difference Lower	Upper			
Pair 1	selfefficacy1 - selfefficacy2	-.33333	.43301	.14434	-.81764	.15098	-2.309	8	.050

Recall formula for CI of mean difference:

(Mean difference) ± (critical t value) * (standard error of the mean difference)

If we change α from .01 to .05, does anything in the above CI formula change?

Answer
For α = .01 critical t = ± 3.36
For α = .05 critical t = ± 2.31

Looking at the CI formula, what effect will changing the critical t from 3.36 to 2.31 have on the CI calculated?

SPSS Results with Alpha = .05 (hence 95% CI)
[Show SPSS results, copy and paste]

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	se_time1	4.6389	9	1.17334	.39111
	se_time2	4.9722	9	1.06393	.35464

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	se_time1 & se_time2	9	.930	.000

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	se_time1 - se_time2	-.33333	.43301	.14434	-.66618	-.00049	-2.309	8	.050

Do we change our **inference** now with an alpha of .05?

With alpha = .05, one would reject Ho and conclude means are different, so efficacy scores do differ between administrations 1 and 2.

2. Notes 6: Pearson Correlation

Difference between Pearson Correlation and Two-independent Samples t-test and Correlated Samples t-test

Nature of the IV

Pearson correlation = IV quantitative
Two-group t-test = IV qualitative with only 2 groups
Correlated t-test = IV qualitative with only 2 groups

See chart in course notes for distinguishing among statistical tests.

See Course Index, section

"12. Types of Statistical Procedures and Their Characteristics: PDF Table"

General Interpretation

a. Found r = -.77 between car horsepower and MPG. What does this tell us; what is the interpretation of this correlation in terms of the variables examined?

Negative relationship ---- The greater the car horsepower, the lower will be expected MPG.

b. Found r = .40 between reading self-efficacy and reading test scores. What does this tell us; what is the interpretation of this correlation in terms of the variables examined?

Positive relationship --- The higher reading self-efficacy, the higher will be reading test scores, on average.

c. Found r = .00 between student weight and interest in mathematics. What does this tell us?

No linear relationship --- student weight and interest in mathematics does not appear to be linearly related; one cannot predict interest in mathematics based upon one's weight.

Worked Example

Data file with three variables

(a) math_sat

Average mathematics SAT scores in each state.

(b) pupil_teacher_ratio

Average ratio of students to teacher in each state; a higher number indicates more students per teacher.

(c) average_teacher_salary

Average salary per teacher in each state in thousands of dollars, thus a figure of 25.000 means the average salary per teacher is $25,000 per year.

Data can be found here:

http://www.bwgriffin.com/gsu/courses/edur8132/tests/math_sat.sav

In regard to Math SAT, how might teacher salary and teacher-student ratio be related to math SAT scores?

Predicted relationships?

a. Math SAT and Teacher Salary --- how related?

Positive – as salary increases math SAT increases, but maybe no relationship. Is a negative relationship likely?

b. Math SAT and Student-Teacher Class Size/Ratio --- how related?

Negative – as class size increases math SAT declines; or maybe no relationship.

What would be the null hypotheses for the above predictions?

Written:

No correlation between Math SAT and Teacher Salary.
No correlation between Math SAT and Class size (Student-teacher ratio).

Symbolic:

Ho: $\rho_{(SAT,Salary)} = 0.00$
Ho: $\rho_{(SAT,Class)} = 0.00$

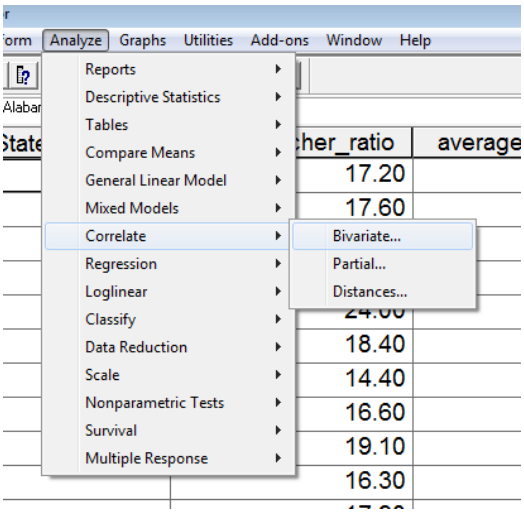
Ho: ρ = 0.00 -> symbolic way of saying that the correlation between two variables in the population is zero (population correlation is symbolized by Greek ρ [rho] rather than r)

H_1: ρ ≠ 0.00 -> symbolic way of saying that the correlation between two variables in the population is not zero, hence there is a correlation between the two variables

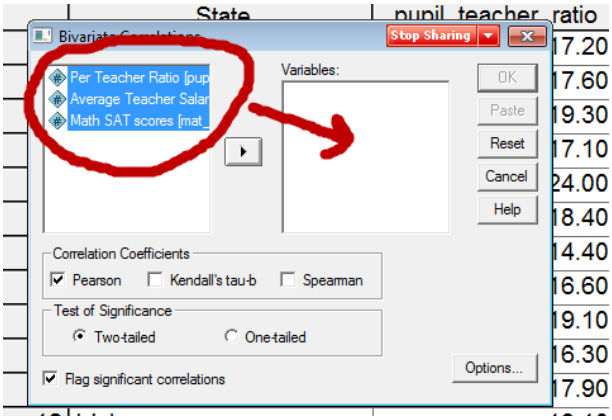
r = correlation

For each of the questions below, set alpha = .05
[Run in SPSS, copy and paste from SPSS]

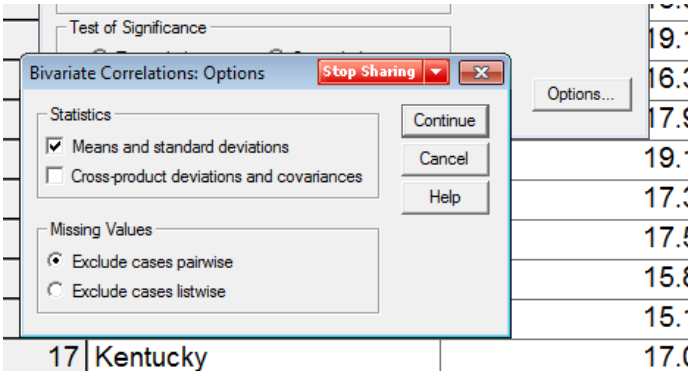
SPSS Commands



Move variables to be correlated to variables box



Select Options then place mark next to Means and Standard deviations to obtain descriptive statistics for each variable



SPSS Results

Descriptive Statistics

	Mean	Std. Deviation	N
Per Teacher Ratio	16.8580	2.26635	50
Average Teacher Salary in Thousands of Dollars	35.2800	5.96603	50
Math SAT scores	508.7800	40.20473	50

Correlations

		Per Teacher Ratio	Average Teacher Salary in Thousands of Dollars	Math SAT scores
Per Teacher Ratio	Pearson Correlation	1	.010	.095
	Sig. (2-tailed)		.945	.510
	N	50	50	50
Average Teacher Salary in Thousands of Dollars	Pearson Correlation	.010	1	-.403**
	Sig. (2-tailed)	.945		.004
	N	50	50	50
Math SAT scores	Pearson Correlation	.095	-.403**	1
	Sig. (2-tailed)	.510	.004	
	N	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

a. What is the correlation between Math SAT and student-teacher ratio (find correlation with SPSS). Do we reject or fail to reject Ho (α = .05)?

Recall decision rule for p-values:

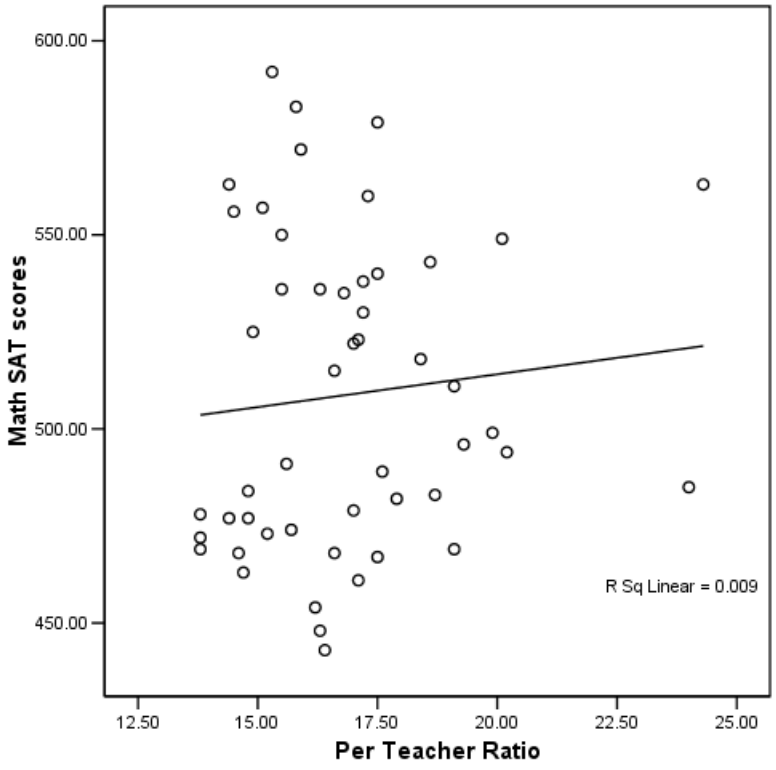
If p ≤ α reject Ho; if p > α fail to reject Ho

r = .095 (p = .51, .51 is larger than alpha = .05 and alpha = .01, so fail to reject)
Pearson Correlation r = .095
Pearson Correlation p-value for that r is p = .51

How would this be interpreted?

There is no association between class size ratio and SAT math scores ---- on average class size does not appear to predict or be related to SAT math scores.

Scatterplot of data.



Descriptive Statistics

	Mean	Std. Deviation	N
Per Teacher Ratio	16.8580	2.26635	50
Average Teacher Salary in Thousands of Dollars	35.2800	5.96603	50
Math SAT scores	508.7800	40.20473	50

Correlations

		Per Teacher Ratio	Average Teacher Salary in Thousands of Dollars	Math SAT scores
Per Teacher Ratio	Pearson Correlation	1	.010	.095
	Sig. (2-tailed)		.945	.510
	N	50	50	50
Average Teacher Salary in Thousands of Dollars	Pearson Correlation	.010	1	-.403**
	Sig. (2-tailed)	.945		.004
	N	50	50	50
Math SAT scores	Pearson Correlation	.095	-.403**	1
	Sig. (2-tailed)	.510	.004	
	N	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

b. What is the correlation between Math SAT and teacher salary?

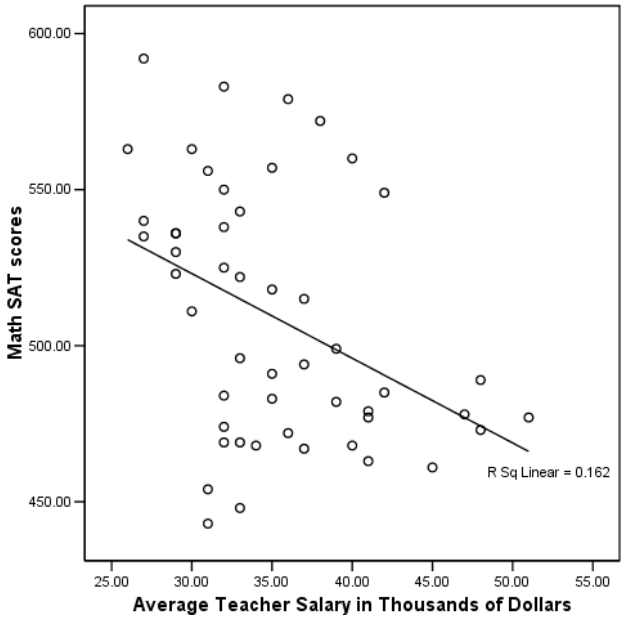
Do we reject or fail to reject Ho?

r = -.403 (p = .004, which is less than alpha = .05 and less than alpha = .01, so reject Ho)

How would this be interpreted?

As teacher salary increases, SAT math scores decline

Scatterplot of scores.



SPSS Results

Descriptive Statistics

	Mean	Std. Deviation	N
Math SAT scores	508.7800	40.20473	50
Per Teacher Ratio	16.8580	2.26635	50
Average Teacher Salary in Thousands of Dollars	35.2800	5.96603	50

APA Style Presentation

[Note location of APA styled example presentation on course website]

Table 1

Correlations and Descriptive Statistics for State-level Mean Scores for Math SAT, Student-teacher Ratio, and Teacher Salary

	1	2	3
1. Math SAT	---		
2. Student-Teacher Ratio	.095	---	
3. Teacher Salary	-.403*	.010	---
M	508.78	16.86	35.28
SD	40.20	2.27	5.97

Note. n = 50.

* $p < .05$.

There a statistically significant association between state-level mean mathematics SAT scores and teacher salary. There is not, however, an association at the .05 level between state-level mean mathematics SAT scores and student-teacher class ratio. Results show that states with higher salaried teachers tend to have lower mathematics SAT scores, while states with lower salaried teachers tend to have higher mathematics SAT scores; stated differently, there is a negative association between mathematics SAT scores and mean teacher salary. Results also show that mathematics SAT scores are unrelated to student-teacher class ratio, and this suggests that mathematics SAT scores are similar for both large and small sized classes across the states.