Statistics with jamovi

Dana Wanzer

Last Update: 2020-10-24

Contents

Welcome

This is the website for PSYC 290 and PSYC 790 at the University of Wisconsin-Stout, taught by Dana Wanzer. These resources are aimed at teaching you how to use jamovi and null hypothesis significance testing (NHST) to answer research questions.

Portions of this book may have been adapted from "Learning statistics with jamovi: A tutorial for psychology students and other beginners" by Danielle J. Navarro and David R. Foxcroft, version 0.70. Furthermore, the template and style of this book is from PsyTeachR.

Chapter 1

Introduction

This chapter will walk you through how this website/book works.

1.1 Quiz Questions

Throughout this website, there will be questions to help you test your knowledge. When you type in or select the correct answer, the dashed box will change color and become solid.

For example:

- What is 2+2?
- We attend the University of Wisconsin- Stout Madison Green Bay
- True or false: Statistics is awesome. TRUE FALSE

1.2 Errors and mistakes

I am human, therefore I err. If you find an error in the textbook or something you think might be a mistake, please let me know ASAP so I can update this for everyone else. Let me know which section you find the error or mistake in and what the error or mistake is. For example, if there was an error here you could say, "There was an error in 1.2 that the first sentence should really be 'To err is human.' "

Chapter 2

Independent t-test

2.1 What is the independent t-test?

The independent t-test is used to test the difference in our dependent variable between two different groups of observations. Our grouping variable is our independent variable. In other words, we use the independent t-test when we have a research question with a **continuous dependent variable** and a **categorical independent variable with two categories in which different participants are in each category**.

The independent t-test is also the independent samples t-test and the Student's t-test. I will use these terms interchangeably.

2.2 Data set-up

To conduct the independent t-test, we first need to ensure our data is set-up properly in our dataset. This requires having two columns: one with our continuous dependent variable and one indicating which group the participant is in. Each row is a unique participant or unit of analysis. Here's what example data may look like if we were testing for differences in a test score by students in my fall or spring semesters of this course:

Table 2.1: Example data for the independent t-test

ID	Semester	TestScore
1	Fall	86
2	Fall	80
3	Fall	75

ID	Semester	TestScore
4	Fall	79
5	Fall	82
6	Spring	84
7	Spring	90
8	Spring	72
9	Spring	75
10	Spring	81

In the example data above, what is your **independent variable**? ID Semester TestScore

In the example data above, what is your **dependent variable**? ID Semester TestScore

2.3 The math behind the independent t-test

The basic math of the independent t-test the mean difference divided by the pooled standard error.

$t = \frac{\bar{X}_1 - \bar{X}_2}{SE(\bar{X}_1 - \bar{X}_2)}$

The denominator of the equation is more difficult to calculate and depends on whether the sample size between groups is equal.

2.4 Assumptions

As a parametric test, the independent t-test has the same assumptions as other parametric tests:

1. The dependent variable is **normally distributed**
2. Variances in the two groups are roughly equal (i.e., **homogeneity of variances**)
3. The dependent variable is **interval or ratio** (i.e., continuous)
4. Scores are **independent** between groups

We cannot <u>test</u> the third and fourth assumptions; rather, those are based on knowing your data.

However, we can and should test for the first two assumptions. Fortunately, the independent samples t-test in jamovi has two check boxes under “Assumption Checks” that lets us test for both assumptions.

2.5 In jamovi

Let's run an example with data from lsj-data. Open data from your Data Library in "lsj-data". Select and open "Harpo". This dataset is hypothetical data of 33 students taking Dr. Harpo's statistics lectures. We have two tutors for the class, Anastasia ($n = 15$) and Bernadette ($n = 18$). Our research question is "Which tutor results in better student grades?" We don't have a hypothesis that one does better than the other.

1. To perform an independent t-test in jamovi, go to the Analyses tab, click the T-Tests button, and choose "Independent Samples T-Test".
2. Move your dependent variable `grade` to the Dependent Variables box and your independent variable `tutor` to the Grouping Variable box.
3. Under Tests, select `Student's`
4. Under Hypothesis, because we have a two-sided hypothesis select a two-sided hypothesis (Group 1 does not equal Group 2).
5. Under Additional Statistics, select `Mean difference`, `Effect size`, and `Descriptives`.
6. Under Assumption Checks, select all three options: `Homogeneity test`, `Normality test`, and `Q-Q plot`.

When you are done, your setup should look like this

2.5.1 Checking assumptions in jamovi

2.5.1.1 Testing normality

We test for normality using the Shapiro-Wilk test and the Q-Q plot. The Shapiro-Wilk test was not statistically significant (W = .98, p = .827); therefore, this indicates the data is normally distributed. Furthermore, the lines are fairly close to the diagonal line in the Q-Q plot. We can conclude that we satisfy the assumption of normality.

2.5.1.2 Testing homogeneity of variance

We test for homogeneity of variance using the Levene's test. The Levene's test was not statistically significant (F [1, 31] = 2.49, p = .125); therefore, this indicates our data satisfies the assumption of homogeneity of variance. However, I would add a caveat that we have a small sample of data (n = 15 for Anastasia and n = 18 for Bernadette) and the standard deviations are quite different from one another (SD = 9.00 vs 5.77, respectively). We should have tried to collect more data.

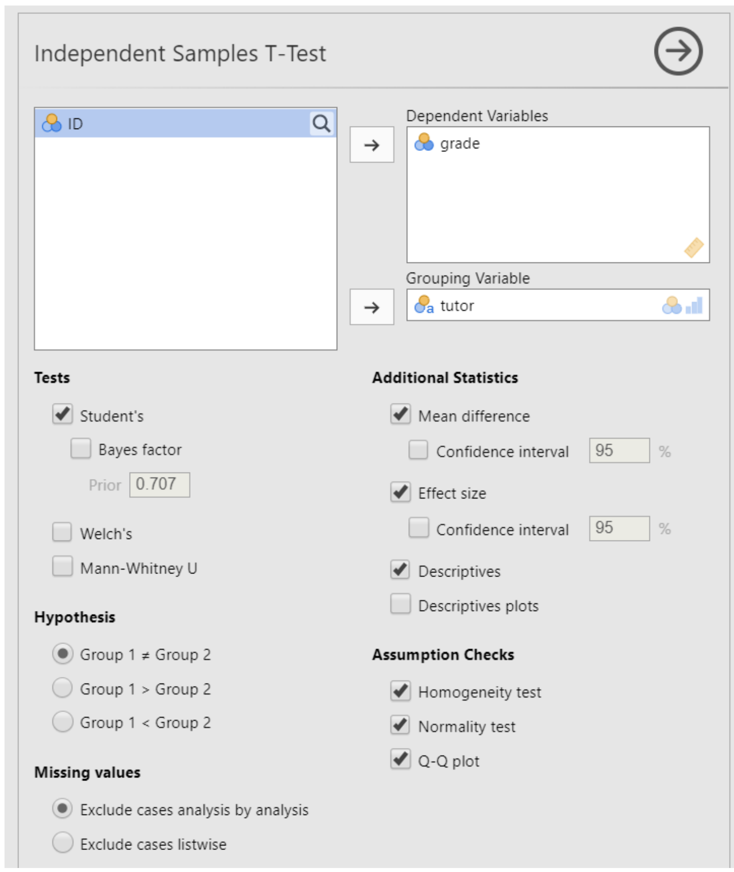

Figure 2.1: Independent t-test setup in jamovi

Assumptions

Normality Test (Shapiro-Wilk)

	W	p
grade	0.98	0.827

Note. A low p-value suggests a violation of the assumption of normality

Plots

grade

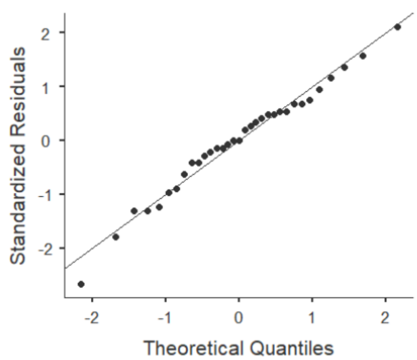

Figure 2.2: Testing normality in jamovi

Assumptions

Homogeneity of Variances Test (Levene's)

	F	df	df2	p
grade	2.49	1	31	0.125

Note. A low p-value suggests a violation of the assumption of equal variances

[3]

Figure 2.3: Testing homogeneity of variance in jamovi

2.5.2 Interpreting results

Once we are satisfied we have satisfied the assumptions for the independent t-test, we can interpret our results.

Independent Samples T-Test

Independent Samples T-Test

		Statistic	df	p	Mean difference	SE difference		Effect Size
grade	Student's t	2.12	31.00	0.043	5.48	2.59	Cohen's d	0.74

Group Descriptives

	Group	N	Mean	Median	SD	SE
grade	Anastasia	15	74.53	76.00	9.00	2.32
	Bernadette	18	69.06	69.00	5.77	1.36

Figure 2.4: Independent t-test results in jamovi

Our p-value is less than .05, so our results are statistically significant. We can write up our results in APA something like this:

> Anastasia's students ($M = 74.53$, $SD = 9.00$, $n = 15$) had significantly higher grades than Bernadette's students ($M = 69.06$, $SD = 5.77$, $n = 18$), t (31) = 2.12, $p = .043$, $d = .74$.

Sometimes, people like to put the statistics inside a parentheses. In that case, you need to change the parentheses around the degrees of freedom as brackets. Here's another example write-up of the results in APA style:

> I tested the difference in grades between Anastasia's students ($M = 74.53$, $SD = 9.00$, $n = 15$) and Bernadette's students ($M = 69.06$, $SD = 5.77$, $n = 18$). An independent samples t-test showed that the 5.48 mean difference between the tutor's student was statistically significant (t [31] = 2.12, $p = .043$, $d = .74$).

2.6 Additional information about the independent t-test

2.6.1 Positive and negative t values

Students often worry about positive or negative t-statistic values and are unsure how to interpret it. Positive or negative t-statistic values simply occur based on which group is listed first. Our t-statistic above is positive because we tested the difference between Anastasia and Bernadette: (Anastasia - Bernadette) = (74.53 - 69.06) = (5.48).

However, if we flipped it and tested the difference between Bernadette and Anastasia, our mean difference would be -5.48 and our t-statistic would be -2.12.

All that is to say, *your positive or negative t-statistic is arbitrary*. So do not fret!

However, it is important the sign of your t-statistic matches what you report. For example, notice the difference:

1. Anastasia's students had **higher** grades than Bernadette's, t (31) = **2.12**, p = .043, d = .74.
2. Bernadette's students had **lower** grades than Anastasia's, t (31) = **-2.12**, p = .043, d = .74.

One last note: this positive or negative t-statistic is only relevant for the independent and dependent t-test. You will not get negative values for the F-statistic or chi-square tests!

2.6.2 What if I violated assumptions?

The great news is that jamovi includes the Welch's t-statistic and the non-parametric version of the independent t-test (Mann-Whitney U)! The Welch's t-test has three main differences from the independent samples t-test: (a) the standard error (SE) is not a pooled estimate, (b) the degrees of freedom are calculated very different (not N - 2), and (c) it does not have an assumption of homogeneity of variance. The Mann-Whitney U is not calculated based on the mean but rather the median and compares ranks of values across the two groups: it has no assumptions about the distribution of data or homogeneity of variances.

Here's what statistic you should choose based on satisfying assumptions:

	Normality: satisfied	**Normality: not satisfied**
Homogeneity of Variance: satisfied	independent samples t-test	Mann-Whitney U
Homogeneity of Variance: not satisfied	Welch's t-test	Mann-Whitney U

Here is what the output for all three tests look like:

Independent Samples T-Test

Independent Samples T-Test

		Statistic	df	p	Mean difference	SE difference		Effect Size
grade	Student's t	2.12	31.00	0.043	5.48	2.59	Cohen's d	0.74
	Welch's t	2.03	23.02	0.054	5.48	2.69	Cohen's d	0.72
	Mann-Whitney U	79.50		0.046	6.00		Rank biserial correlation	0.41

Group Descriptives

	Group	N	Mean	Median	SD	SE
grade	Anastasia	15	74.53	76.00	9.00	2.32
	Bernadette	18	69.06	69.00	5.77	1.36

Figure 2.5: All independent t-test results in jamovi

2.6.2.1 Welch's t-test in jamovi

To conduct this in jamovi, under Tests select `Welch's`. You will interpret the results similarly to the independent t-test:

> Using a Welch's t-test, there was not a statistically significant difference in grades between Anastasia's students (M = 74.53, SD = 9.00, n = 15) and Bernadette's students (M = 69.06, SD = 5.77, n = 18), t (23.02) = 2.03, p = .054, d = .72.

Why is it no longer statistically significant? Which result should you trust? In reality, the difference in p-values is likely due to chance. However, the independent t-test and Welch's test have different strengths and weaknesses. If the two populations really do have equal variances, then the independent t-test is slightly more powerful (lower Type II error rate) than the Welch's test. However, if they *don't* have the same variances, then the assumptions of the independent t-test are violated and you may not be able to trust the results; you may end up with a higher Type I error rate. So it's a trade-off.

Which should you use? I tend to prefer always using Welch's t-test because if the variances are equal, then there will be practically no difference between the independent and Welch's t-test. But if the variances are not equal, then Welch's t-test will outperform the independent t-test. For that reason, defaulting to the Welch's t-test makes most sense to me.

2.6.2.2 Mann-Whitney U test

If you do not satisfy the assumption of normality (regardless of whether you satisfy the assumption of homogeneity of variance), you should either try to transform your data to be normally distributed or you will need to use a non-parametric test. In this case, if you originally wanted to perform an independent t-test, the non-parametric equivalent test is the Mann-Whitney U test.

I will not go into specifics, but the idea behind the Mann-Whitney U test is that you take all the values (regardless of group) and rank them. You then sum the ranks across groups and calculate your U statistic and p-value. You interpret the p-value like you normally would, but there are differences in how we report the results because this statistic is based on the *median* not the *mean*.

> Using the Mann-Whitney U test, there was a statistically significant difference in grades between Anastasia's students (Mdn = 76, n = 15) and Bernadette's students (Mdn = 69, n = 18), $t\,(23.02) = 2.03$, $p = .054$, $d = .72$.

2.7 Your turn!

Open the `Sample_Dataset_2014.xlsx` file that we will be using for all Your Turn exercises.

Perform independent t-tests based on the following research questions. Think critically about whether you should be using a one-tailed or two-tailed hypothesis and check your assumptions so you know which test to use!

To get the most out of these exercises, try to first find out the answer on your own and then use the drop-down menus to check your answer.

1. **Does height differ by gender (Gender: male = 0, female = 1)?**

    - Should you use a one-tailed or two-tailed hypothesis? one-tailed two-tailed
    - Which statistic should you use based on your assumptions? independent t-test Welch's t-test Mann Whitney U
    - Does height differ by gender? yes no

2. **Do athletes (Athlete: athletes = 1, non-athlete = 0) have faster sprint times than non-athletes?**

    - Should you use a one-tailed or two-tailed hypothesis? one-tailed two-tailed
    - Which statistic should you use based on your assumptions? independent t-test Welch's t-test Mann Whitney U
    - Do athletes have faster sprint times than non-athletes? yes no

3. **Do students who live on campus (LiveOnCampus: on campus = 1, off campus = 0) have higher English scores than students who live off campus?**

    - Should you use a one-tailed or two-tailed hypothesis? one-tailed two-tailed
    - Which statistic should you use based on your assumptions? independent t-test Welch's t-test Mann Whitney U
    - Does students who live on campus have higher English scores? yes no

4. **Does athletic status relate to math scores?**

    - Should you use a one-tailed or two-tailed hypothesis? one-tailed two-tailed
    - Which statistic should you use based on your assumptions? independent t-test Welch's t-test Mann Whitney U
    - Does athletic status relate to math scores? yes no

Chapter 3

Dependent t-test

3.1 What is the dependent t-test?

The dependent t-test is used to test the difference in our dependent variable between two categories in which participants are the *same* across categories. Our category variable is our independent variable. In other words, we use the independent t-test when we have a research question with a **continuous dependent variable** and a **categorical independent variable with two categories in which the same participants are in each category**.

The dependent t-test is also called a dependent samples t-test or paired samples t-test.

3.2 Data set-up

To conduct the dependent t-test, we first need to ensure our data is set-up properly in our dataset. This requires having two columns: one is our dependent variable score for the participant in one category and the other column is our dependent variable score for the participant in the other category. Each row is a unique participant or unit of analysis. Here's what example data may look like if we were testing for differences in test scores across the same participants in the fall and spring:

Table 3.1: Example data for the dependent t-test

ID	TestScore_Fall	TestScore_Spring
1	75	86
2	79	80

ID	TestScore_Fall	TestScore_Spring
3	65	75
4	81	79
5	73	82
6	72	84
7	69	90
8	60	72
9	75	75
10	74	81

In the example data above, what is your **independent variable**? ID Semester TestScore

In the example data above, what is your **dependent variable**? ID Semester Test Score

3.3 The math behind the independent t-test

The basic math of the dependent t-test is the mean difference divided by the standard error, which is estimated based on the standard deviation and sample size (N).

$t = \frac{\bar{X}_1 - \bar{X}_2}{s_d/\sqrt{N}}$

3.4 Assumptions

As a parametric test, the independent t-test has the same assumptions as other parametric tests minus the homogeneity of variance assumption because we are dealing with the same people across categories

1. The *differences in scores* in the dependent variable are **normally distributed**
2. The dependent variable is **interval or ratio** (i.e., continuous)
3. Scores are **independent** across participants

We cannot test the second and third assumptions; rather, those are based on knowing your data.

However, we can and should test for the first assumption. Fortunately, the dependent samples t-test in jamovi has two check boxes under "Assumption Checks" that lets us test normality.

3.5 In jamovi

Let's run an example with data from lsj-data. Open data from your Data Library in "lsj-data". Select and open "Chico". This dataset is hypothetical data from Dr. Chico's class in which students took two tests: one early in the semester and one later in the semester. Dr. Chico thinks that the first test is a "wake up call" for students. When they realise how hard her class really is, they'll work harder for the second test and get a better mark. Is she right? Let's test it!

1. To perform an dependent t-test in jamovi, go to the Analyses tab, click the T-Tests button, and choose "Paired Samples T-Test".
2. Move both measurements of your dependent variable (`grade_test1` and `grade_test2`) to the Paired Variables box.
3. Under Tests, select `Student's`
4. Under Hypothesis, choose the correct hypothesis: Measure 1 is not equal to Measure 2 Measure 1 > Measure 2 Measure 1 < Measure 2
5. Under Additional Statistics, select `Mean difference`, `Effect size`, and `Descriptives`.
6. Under Assumption Checks, select both options: `Normality test` and `Q-Q plot`.

When you are done, your setup should look like this

3.5.1 Checking assumptions in jamovi

3.5.1.1 Testing normality

We test for normality using the Shapiro-Wilk test and the Q-Q plot. The Shapiro-Wilk test was not statistically significant (W = .97, $p = .678$); therefore, this indicates the data is normally distributed. Furthermore, the lines are fairly close to the diagonal line in the Q-Q plot (although it's a bit hard to tell because our sample size is small). We can conclude that we satisfy the assumption of normality.

3.5.2 Interpreting results

Once we are satisfied we have satisfied the assumptions for the dependent t-test, we can interpret our results.

Our p-value is less than .05, so our results are statistically significant. We can write up our results in APA something like this:

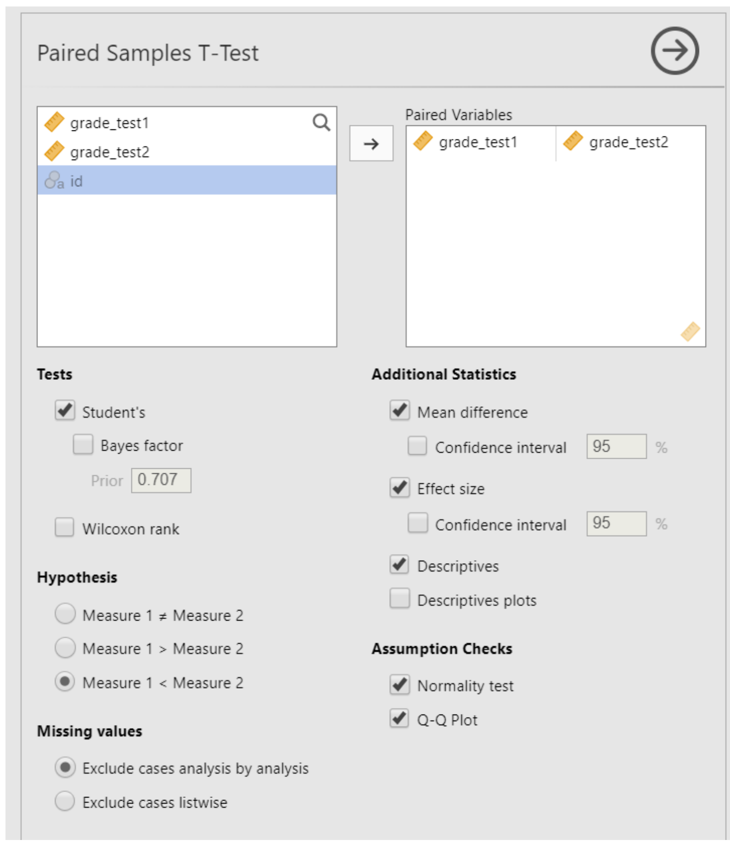

Figure 3.1: Dependent t-test setup in jamovi

Normality Test (Shapiro-Wilk)

			W	p
grade_test1	-	grade_test2	0.97	0.678

Note. A low p-value suggests a violation of the assumption of normality

Plots

grade_test1 - grade_test2

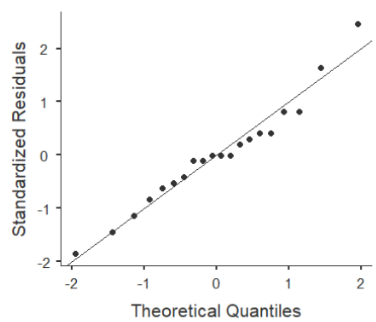

Figure 3.2: Testing normality in jamovi

Paired Samples T-Test

Paired Samples T-Test

			statistic	df	p	Mean difference	SE difference		Effect Size
grade_test1	grade_test2	Student's t	−6.48	19.00	< .001	−1.40	0.22	Cohen's d	−1.45

Note. H_a Measure 1 < Measure 2

Descriptives

	N	Mean	Median	SD	SE
grade_test1	20	56.98	57.70	6.62	1.48
grade_test2	20	58.38	59.70	6.41	1.43

Figure 3.3: Dependent t-test results in jamovi

> The 20 students in Dr. Chico's class performed worse on the first test (M = 56.98, SD = 6.62) than they did on the second test (M = 58.38, SD = 6.41), $t(19)$ = -6.48, p < .001, d = -1.45.

Remember in the previous chapter that our t-test can be negative but we can always flip the interpretation. Here's another example of how we could write-up our results in APA style:

> Dr. Chico's hypothesis was correct in that her 20 students performed better on the second test (M = 58.38, SD = 6.41) than they did on the first test (M = 56.98, SD = 6.62), $t(19)$ = 6.48, p < .001, d = 1.45.

3.6 What if I violated assumptions?

If you violated the assumption of normality and no transformation fixed your data, then you can perform the non-parametric version of the dependent t-test called the Wilcoxon Rank test. As a reminder, non-parametric tests do not make assumptions about the distribution of data because it deals with the *median* not the *mean*.

Here is the output for both the dependent t-test and the Wilcoxon rank test:

Paired Samples T-Test

Paired Samples T-Test

			Statistic	df	p		Effect Size
grade_test1	grade_test2	Student's t	−6.48	19.00	< .001	Cohen's d	−1.45
		Wilcoxon W	2.00[a]		< .001	Rank biserial correlation	−0.98

Note. H_a Measure 1 < Measure 2

[a] 1 pair(s) of values were tied

Descriptives

	N	Mean	Median	SD	SE
grade_test1	20	56.98	57.70	6.62	1.48
grade_test2	20	58.38	59.70	6.41	1.43

Figure 3.4: All independent t-test results in jamovi

3.6.0.1 Wilcoxon rank in jamovi

To conduct this in jamovi, under Tests select `Wilcoxon rank`. You will interpret the results similarly to the dependent t-test:

> Using Wilcoxon rank test, students' test scores were significantly higher at the second test (*Mdn* = 59.70) than at the first test (*Mdn* = 57.70), W = 2.00, $p < .001$.

The note about tied values is not necessary to discuss. It is just telling us one participant had identical values for both test1 and test2 (student15).

3.7 Your turn!

Open the `Sample_Dataset_2014.xlsx` file that we use for all Your Turn exercises.

Perform dependent t-tests based on the following research questions. Think critically about whether you should be using a one-tailed or two-tailed hypothesis and check your assumptions so you know which test to use!

To get the most out of these exercises, try to first find out the answer on your own and then use the drop-down menus to check your answer.

Note: Technically, none of our data is suitable for a dependent t-test in this dataset. We will pretend that the four test score variables (`English`, `Reading`, `Math`, and `Writing`) are really four measurements of the same underlying test. In reality, we would analyze this data using correlation.

1. **Do students perform better on the English test than they do the Writing test?**
   - Should you use a one-tailed or two-tailed hypothesis? one-tailed two-tailed
   - Which statistic should you use based on your assumptions? dependent t-test Wilcoxon rank
   - Do students perform better on the English test than they do the Writing test? yes no
2. **Does students' English scores relate to their Reading scores?**
   - Should you use a one-tailed or two-tailed hypothesis? one-tailed two-tailed
   - Which statistic should you use based on your assumptions? dependent t-test Wilcoxon rank
   - Does students' English scores relate to their Reading scores? yes no

Appendix A

References