

Lab 2 Part B : One-Sample T-Test using SPSS Statistics

1 Introduction

- The one-sample t-test is used to determine whether a sample comes from a population with a specific mean. This population mean is not always known, but is sometimes hypothesized.
- For example, you want to show that a new teaching method for pupils struggling to learn English grammar can improve their grammar skills to the national average. Your sample would be pupils who received the new teaching method and your population mean would be the national average score.
- Alternately, you believe that doctors that work in Accident and Emergency (A & E) departments work 100 hour per week despite the dangers (e.g., tiredness) of working such long hours. You sample 1000 doctors in A & E departments and see if their hours differ from 100 hours.
- This guide shows you how to carry out a one-sample t-test using SPSS Statistics, as well as interpret and report the results from this test.
- However, before we introduce you to this procedure, you need to understand the different assumptions that your data must meet in order for a one-sample t-test to give you a valid result. We discuss these assumptions later in the appendix.

2 Implementation

In the section, we illustrate the SPSS Statistics procedure required to perform a one-sample t-test assuming that no assumptions have been violated. First, we set out the example we use to explain the one-sample t-test procedure in SPSS Statistics.

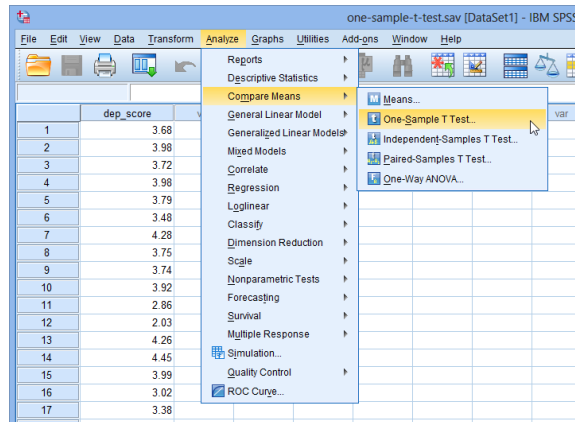
2.1 Example and Setup in SPSS Statistics

- A researcher is planning a psychological intervention study, but before he proceeds he wants to characterise his participants' depression levels.
- He tests each participant on a particular depression index, where anyone who achieves a score of 4.0 is deemed to have 'normal' levels of depression. Lower scores indicate less depression and higher scores indicate greater depression.
- He has recruited 40 participants to take part in the study. Depression scores are recorded in the variable `dep_score`.
- He wants to know whether his sample is representative of the normal population (i.e., do they score statistically significantly differently from 4.0).
- For a one-sample t-test, there will only be one variable's data to be entered into SPSS Statistics: the dependent variable, `dep_score`, which is the depression score.

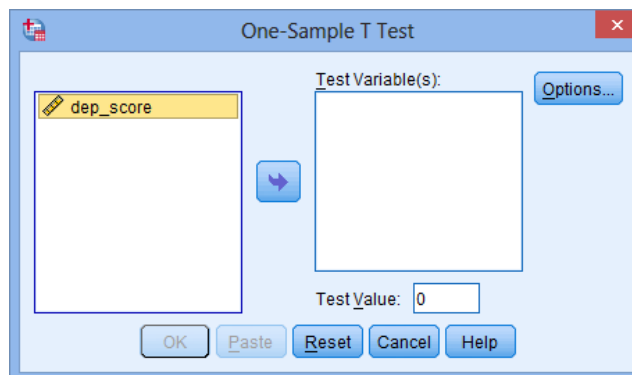
2.2 Test Procedure in SPSS Statistics

- The five steps below show you how to analyse your data using a one-sample t-test in SPSS Statistics when the four relevant assumptions have not been violated (see Appendix).
- At the end of these five steps, we show you how to interpret the results from this test.

Step 1: Click **Analyze > Compare Means > One-Sample T Test...** on the main menu:

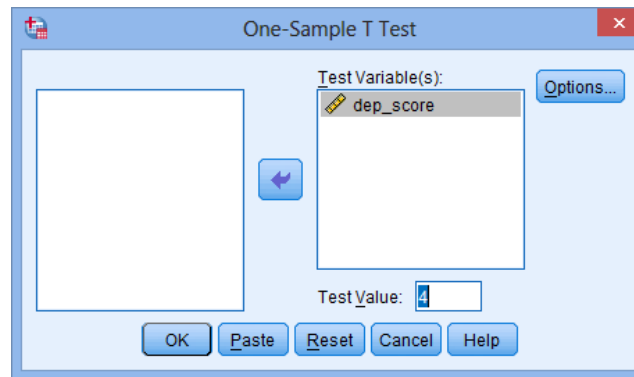


You will be presented with the One-Sample T Test dialogue box, as shown below:

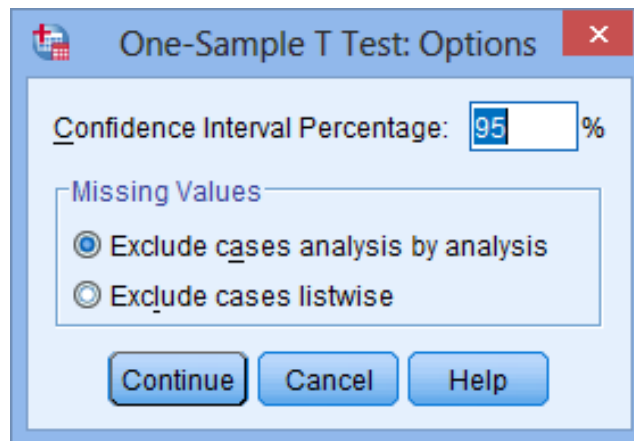


Transfer the dependent variable, `dep_score`, into the **Test Variable(s):** box by selecting it (by clicking on it) and then clicking on the button.

- Enter the population mean you are comparing the sample against in the **Test Value:** box, by changing the current value of “0” to “4”. You will end up with the following screen:



- Click on the **Options** button. You will be presented with the **One-Sample T Test: Options** dialogue box, as shown below:



For this example, keep the default 95% confidence intervals and Exclude cases analysis by analysis in the Missing Values area.

Note 1:

By default, SPSS Statistics uses 95% confidence intervals (labelled as the Confidence Interval Percentage in SPSS Statistics). This equates to declaring statistical significance at the $p < .05$ level.

If you wish to change this you can enter any value from 1 to 99. For example, entering "99" into this box would result in a 99% confidence interval and equate to declaring statistical significance at the $p < .01$ level. For this example, keep the default 95% confidence intervals.

Note 2:

If you are testing more than one dependent variable and you have any missing values in your data, you need to think carefully about whether to select Exclude cases analysis by analysis or Exclude cases listwise in the Missing Values area. Selecting the incorrect option could mean that SPSS Statistics removes data from your analysis that you wanted to include.

- Click the **continue** button. You will be returned to the One-Sample T Test dialogue box. Then click the OK button to generate the output.

2.3 Interpreting the SPSS Statistics output of the one-sample t-test

SPSS Statistics generates two main tables of output for the one-sample t-test that contains all the information you require to interpret the results of a one-sample t-test.

- If your data passed assumption No. 3 (i.e., there were no significant outliers) and assumption No. 4 (i.e., your dependent variable was approximately normally distributed for each category of the independent variable), which we explained earlier in the Assumptions section, you will only need to interpret these two main tables.
- However, since you should have tested your data for these assumptions, you will also need to interpret the SPSS Statistics output that was produced when you tested for them (i.e., you will have to interpret: (a) the boxplots you used to check if there were any significant outliers; and (b) the output SPSS Statistics produces for your Shapiro-Wilk test of normality to determine normality).
- Remember that if your data failed any of these assumptions, the output that you get from the one-sample t-test procedure (i.e., the tables we discuss below), will no longer be relevant, and you will need to interpret these tables differently.

However, in this guide, we take you through each of the two main tables in turn, assuming that your data met all the relevant assumptions:

3 Descriptive statistics

You can make an initial interpretation of the data using the One-Sample Statistics table, which presents relevant descriptive statistics:

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
dep_score	40	3.7225	.73709	.11654

It is more common than not to present your descriptive statistics using the mean and standard deviation ("Std. Deviation" column) rather than the standard error of the mean ("Std. Error Mean" column), although both are acceptable.

You could report the results, using the standard deviation, as follows:

Mean depression score (3.72 ± 0.74) was lower than the population 'normal' depression score of 4.0.

- However, by running a one-sample t-test, you are really interested in knowing whether the sample you have (`dep_score`) comes from a 'normal' population (which has a mean of 4.0). This is discussed in the next section.

3.1 One-sample t-test

The One-Sample Test table reports the result of the one-sample t-test. The top row provides the value of the known or hypothesized population mean you are comparing your sample data to, as highlighted below:

One-Sample Test						
	Test Value = 4					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
dep_score	-2.381	39	.022	-.27750	-.5132	-.0418

Figure 1:

- In this example, you can see the 'normal' depression score value of "4" that you entered in earlier. You now need to consult the first three columns of the One-Sample Test table, which provides information on whether the sample is from a population with a mean of 4 (i.e., are the means statistically significantly different), as highlighted below:
- Moving from left-to-right, you are presented with the observed t-value ("t" column), the degrees of freedom ("df"), and the statistical significance (p-value) ("Sig. (2-tailed)") of the one-sample t-test.

One-Sample Test						
	Test Value = 4					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
dep_score	-2.381	39	.022	-.27750	-.5132	-.0418

Figure 2:

$$t(39) = -2.381, p = .022$$

	Part	Meaning
1	t	Indicates that we are comparing to a t -distribution (t -test).
2	(39)	Indicates the degrees of freedom, which is $N - 1$
3	-2.381	Indicates the obtained value of the t -statistic (obtained t -value)
4	$p = .022$	Indicates the probability of obtaining the observed t -value if the null hypothesis is correct.

Table 4.1: Breakdown of a one-sample t -test statistical statement.

- In this example, $p < .05$ (it is $p = .022$). Therefore, it can be concluded that the population means are statistically significantly different.
- If $p > .05$, the difference between the sample-estimated population mean and the comparison population mean would not be statistically significantly different.
- Note: If you see SPSS Statistics state that the "Sig. (2-tailed)" value is ".000", this actually means that $p < .0005$. It does not mean that the significance level is actually zero.
- SPSS Statistics also reports that $t = -2.381$ ("t" column) and that there are 39 degrees of freedom ("df" column). You need to know these values in order to report your results, which you could do as follows:

Depression score was statistically significantly lower than the population normal depression score, $t(39) = -2.381, p = .022$.

The breakdown of the last part (i.e., $t(39) = -2.381, p = .022$) is as follows:

Table 4.1: Breakdown of a one-sample t -test statistical statement. You can also include measures of the difference between the two population means in your written report. This information is included in the columns on the far-right of the One-Sample Test table, as highlighted below:

This section of the table shows that the mean difference in the population means is -0.28 ("Mean Difference" column) and the 95% confidence intervals (95% CI) of the difference are -0.51 to -0.04 ("Lower" to "Upper" columns). For the measures used, it will be sufficient to report the values to 2 decimal places. You could write these results as:

One-Sample Test

	Test Value = 4					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
dep_score	-2.381	39	.022	-.27750	-.5132	-.0418

Depression score was statistically significantly lower by 0.26 (95% CI, 0.04 to 0.51) than a normal depression score of 4.0, $t(39) = -2.381$, $p = .022$.

4 Reporting Your Results

4.1 Reporting the SPSS Statistics output of the one-sample t-test

You can report the findings, without the tests of assumptions, as follows:

General Mean depression score (3.73 0.74) was lower than the normal depression score of 4.0, a statistically significant difference of 0.26 (95% CI, 0.04 to 0.51), $t(39) = -2.381$, $p = .022$.

APA

Adding in the information about the statistical test you ran, including the assumptions, you have:

A one-sample t-test was run to determine whether depression score in recruited subjects was different to normal, defined as a depression score of 4.0. Depression scores were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$) and there were no outliers in the data, as assessed by inspection of a boxplot.

Mean depression score (3.73 0.74) was lower than the normal depression score of 4.0, a statistically significant difference of 0.26 (95% CI, 0.04 to 0.51), $t(39) = -2.381$, $p = .022$.

Null hypothesis significance testing You can write the result in respect of your null and alternative hypothesis as:

GeneralAPA There was a statistically significant difference between means ($p < .05$) and, therefore, we can reject the null hypothesis and accept the alternative hypothesis.

4.2 Practical vs. statistical significance

- Although a statistically significant difference was found between the depression scores in the recruited subjects vs. the normal depression score, it does not necessarily mean that the difference encountered, 0.26 (95% CI, 0.04 to 0.51), is enough to be practically significant.
- Indeed, the researcher might accept that although the difference is statistically significant (and would report this), the difference is not large enough to be practically significant (i.e., the subjects can be treated as normal).

5 Assumptions

When you choose to analyse your data using a one-sample t-test, part of the process involves checking to make sure that the data you want to analyse can actually be analysed using a one-sample t-test.

You need to do this because it is only appropriate to use a one-sample t-test if your data "passes" four assumptions that are required for a one-sample t-test to give you a valid result.

In practice, checking for these four assumptions just adds a little bit more time to your analysis, requiring you to click a few more buttons in SPSS Statistics when performing your analysis, as well as think a little bit more about your data, but it is not a difficult task.

Before we introduce you to these four assumptions, do not be surprised if, when analysing your own data using SPSS Statistics, one or more of these assumptions is violated (i.e., is not met).

This is not uncommon when working with real-world data rather than textbook examples, which often only show you how to carry out a one-sample t-test when everything goes well! However, don't worry. Even when your data fails certain assumptions, there is often a solution to overcome this. First, let's take a look at these four assumptions:

Assumption No. 1: Correct Scale Your dependent variable should be measured at the interval or ratio level (i.e., continuous). Examples of variables that meet this criterion include revision time (measured in hours), intelligence (measured using IQ score), exam performance (measured from 0 to 100), weight (measured in kg), and so forth. You can learn more about interval and ratio variables in our article: [Types of Variable](#).

Assumption No. 2: Independence The data are independent (i.e., not correlated/related), which means that there is no relationship between the observations. This is more of a study design issue than something you can test for, but it is an important assumption of the one-sample t-test.

Assumption No. 3: No Outliers There should be no significant outliers. Outliers are data points within your data that do not follow the usual pattern (e.g., in a study of 100 students' IQ scores, where the mean score was 108 with only a small variation between students, one student had a score of 156, which is very unusual, and may even put her in the top 1% of IQ scores globally). The problem with outliers is that they can have a negative effect on the one-sample t-test, reducing the accuracy of your results. Fortunately, when using SPSS Statistics to run a one-sample t-test on your data, you can easily detect possible outliers.

Assumption No. 4: Normality Your dependent variable should be approximately normally distributed. We talk about the one-sample t-test only requiring approximately normal data because it is quite "robust" to violations of normality, meaning that the assumption can be a little violated and still provide valid results. You can test for normality using the Shapiro-Wilk test of normality, which is easily tested for using SPSS Statistics.

5.1 Checking Assumptions

- You can check assumptions No. 3 and No. 4 using SPSS Statistics. Before doing this, you should make sure that your data meets assumptions No. 1 and No. 2, although you don't need SPSS Statistics to do this. When moving on to assumptions No. 3 and No. 4, we suggest testing them in this order because it represents an order where, if a violation to the assumption is not correctable, you will no longer be able to use a one-sample t-test.

- Just remember that if you do not run the statistical tests on these assumptions correctly, the results you get when running a one-sample t-test might not be valid.