



The *t*-test in IBM SPSS Statistics

An Example: are invisible people mischievous?

In my SPSS book (Field, 2013) I imagine a future in which we have some cloaks of invisibility to test out. As a psychologist (with his own slightly mischievous streak) I might be interested in the effect that wearing a cloak of invisibility has on people's tendency for mischief. I took 24 participants and placed them in an enclosed community. The community was riddled with hidden cameras so that we could record mischievous acts. Half of them were given cloaks of invisibility: they were told not to tell anyone else about their cloak and that they could wear it whenever they liked. We measured how many mischievous acts they performed in a week. These data are in Table 1. The file **Invisibility.sav** shows how you should have entered the data: the variable **Cloak** records whether or not a person was given a cloak (**cloak** = 1) or not (**cloak** = 0), and **Mischief** is how many mischievous acts were performed.

Table 1: Data from **Invisibility.sav**

Participant	Cloak	Mischief
1	0	3
2	0	1
3	0	5
4	0	4
5	0	6
6	0	4
7	0	6
8	0	2
9	0	0
10	0	5
11	0	4
12	0	5
13	1	4
14	1	3
15	1	6
16	1	6
17	1	8
18	1	5
19	1	5
20	1	4
21	1	2
22	1	5
23	1	7
24	1	5

1.1. The independent t -test using SPSS

The general procedure

Figure 1 shows the general process for performing a t -test: as with fitting any model, we start by looking for the sources of bias. Having satisfied ourselves that assumptions are met and outliers dealt with, we run the test. We can also consider using bootstrapping if any of the test assumptions were not met. Finally, we compute an effect size.

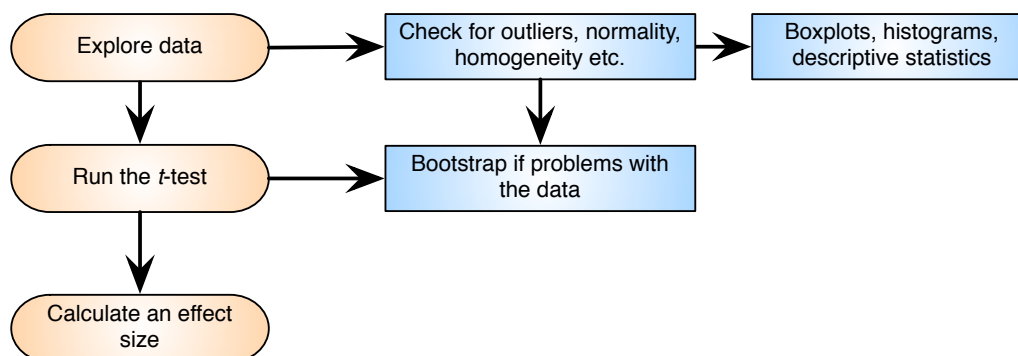


Figure 1: The general process for performing a t -test

Compute the independent t -test

To run an independent t -test, we need to access the main dialog box by selecting **Analyze** **Compare Means** **Independent-Samples T Test...** (see Figure 2). Once the dialog box is activated, select the dependent variable from the list (click on **Mischief**) and transfer it to the box labelled *Test Variable(s)* by dragging it or clicking on . If you want to carry out t -tests on several dependent variables then you can select other dependent variables and transfer them to the variables list. However, there are good reasons why it is not a good idea to carry out lots of tests.

Next, we need to select an independent variable (the grouping variable). In this case, we need to select **Cloak** and then transfer it to the box labelled *Grouping Variable*. When your grouping variable has been selected the **Define Groups...** button will become active and you should click on it to activate the *Define Groups* dialog box. SPSS needs to know what numeric codes you assigned to your two groups, and there is a space for you to type the codes. In this example, we coded our no cloak group as 0 and our cloak group as 1, and so these are the codes that we type.

When you have defined the groups, click on **Continue** to return to the main dialog box. If you click on **Options...** then another dialog box appears that gives you the chance to change the width of the confidence interval that is calculated. The default setting is for a 95% confidence interval and this is fine; however, if you want to be stricter about your analysis you could choose a 99% confidence interval but you run a higher risk of failing to detect a genuine effect (a Type II error). To run the analysis click on **OK**.

If we have potential bias in the data we can reduce its impact by using bootstrapping to generate confidence intervals for the difference between means. We can select this option by clicking **Bootstrap...** in the main dialog box to access the bootstrap function. Select **Perform bootstrapping** to activate bootstrapping, and to get a 95% confidence interval click **Percentile** or

Bias corrected accelerated (BCa). For this analysis, let's ask for a bias corrected (BCa) confidence interval. Back in the main dialog box click on **OK** to run the analysis.

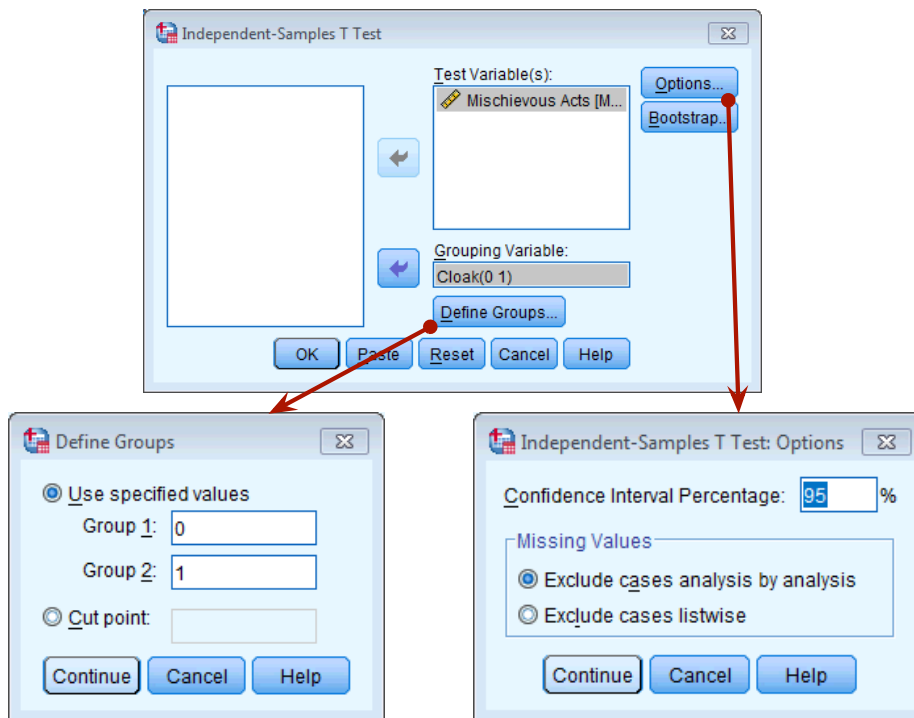


Figure 2: Dialog boxes for the independent-samples t -test

Output from the independent t -test (1)

The output from the independent t -test contains only three tables (two if you don't opt for bootstrapping). The first table (Output 1) provides summary statistics for the two experimental conditions (if you don't ask for bootstrapping this table will be a bit more straightforward). From this table, we can see that both groups had 12 participants (row labelled N). The group who had no cloak, on average, performed 3.75 mischievous acts with a standard deviation of 1.913. What's more, the standard error of that group is 0.552. The bootstrap SE estimate is .53, and the bootstrapped confidence interval for the mean ranges from 2.29 to 4.58. For those that were given an invisibility cloak, they performed, on average, 5 acts, with a standard deviation of 1.651, a standard error of 0.477. The bootstrap standard error is a bit lower at 0.46, and the confidence interval for the mean ranges from 4.33 to 5.67. Note that the confidence intervals for the two groups overlap, implying that they might be from the same population.

The second table of output (Output 2) contains the main test statistics. The first thing to notice is that there are two rows containing values for the test statistics: one row is labelled *Equal variances assumed*, while the other is labelled *Equal variances not assumed*. Parametric tests assume that the variances in experimental groups are roughly equal. The rows of the table relate to whether or not this assumption has been broken.



Paired Samples Statistics

		Statistic	Bootstrap ^a				
			Bias	Std. Error	BCa 95% Confidence Interval		
					Lower	Upper	
Pair 1	Mischief (No Invisibility Cloak)	Mean	3.75	.01	.53	2.92	4.58
		N	12				
		Std. Deviation	1.913	-.127	.354	1.311	2.229
		Std. Error Mean	.552				
	Mischief (Invisibility Cloak)	Mean	5.00	.00	.46	4.33	5.67
		N	12				
		Std. Deviation	1.651	-.117	.307	1.168	1.881
		Std. Error Mean	.477				

a. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples

Output 1

We can use Levene's test to see whether variances are different in different groups (although there are problems with this test discussed in my book), and SPSS produces this test for us. Levene's test tests the hypothesis that the variances in the two groups are equal. Therefore, if Levene's test is significant at $p \leq .05$, it suggests that the assumption of homogeneity of variances has been violated. If, however, Levene's test is non-significant (i.e., $p > .05$) then we can assume that the variances are roughly equal and the assumption is tenable. For these data, Levene's test is non-significant (because $p = .468$, which is greater than $.05$) and so we should read the test statistics in the row labelled *Equal variances assumed*. Had Levene's test been significant, then we would have read the test statistics from the row labelled *Equal variances not assumed*.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Mischievous Acts	Equal variances assumed	.545	.468	-1.713	22	.101	-1.250	.730	-2.763	.263
	Equal variances not assumed			-1.713	21.541	.101	-1.250	.730	-2.765	.265

Output 2

Having established that the assumption of homogeneity of variances is met, we can look at the t -test itself. We are told the mean difference ($\bar{X}_{NC} - \bar{X}_C = 3.75 - 5 = -1.25$) and the standard error of the sampling distribution of differences. The t -statistic is -1.71 , which is assessed against the value of t you might expect to get if there was no effect in the population when you have certain degrees of freedom. For the independent t -test, degrees of freedom are calculated by adding the two sample sizes and then subtracting the number of samples ($df = N_1 + N_2 - 2 = 12 + 12 - 2 = 22$). SPSS produces the exact significance value of t , and typically we are interested in whether this value is less than or greater than $.05$. In this case the two-tailed value of p is $.107$, which is greater than $.05$, and so we would have to conclude that there was no significant difference between the means of these two samples. In terms of the experiment, we can infer that having a cloak of invisibility did not significantly affect the amount of mischief a person got up to.

Bootstrap for Independent Samples Test

		Mean Difference	Bootstrap ^a			
			Bias	Std. Error	BCa 95% Confidence Interval	
					Lower	Upper
Mischievous Acts	Equal variances assumed	-1.250	.003	.726	-2.606	.043
	Equal variances not assumed	-1.250	.003	.726	-2.606	.043

a. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples

Output 3

Output 3 shows the results of the bootstrapping (if you selected it). You can see that the bootstrapping procedure has been applied to re-estimate the standard error of the mean difference (which is estimated as .726 rather than .730). Using this bootstrapped standard error confidence intervals for the difference between means are computed. The difference between means was -1.25 , and the confidence interval ranged from -2.606 to 0.043 . The confidence interval implies that the difference between means in the population could be negative, positive or even zero (because the interval ranges from a negative value to a positive one). In other words, it's possible that the true difference between means is zero—no difference at all. Therefore, this bootstrap confidence interval confirms our conclusion that having a cloak of invisibility seems not to affect acts of mischief.

Reporting the independent t -test (1)

You usually state the finding to which the test relates and then report the test statistic, its degrees of freedom and the probability value of that test statistic. We could write:

- ✓ On average, participants given a cloak of invisibility engaged in more acts of mischief ($M = 5$, $SE = 0.48$), than those not given a cloak ($M = 3.75$, $SE = 0.55$). This difference, -1.25 , BCa 95% CI [-2.606 , 0.043], was not significant $t(22) = -1.71$, $p = .101$; however, it did represent a medium-sized effect $d = .65$.




1.2. Matched-samples t -test using SPSS

Entering Data

Let's imagine that we had collected the cloak of invisibility data using a repeated measures design; so, the data are identical to before. In this scenario we might have recorded everyone's natural level of mischievous acts in a week, then given them an invisibility cloak and counted the number of mischievous acts in the next week.

The data would now be arranged differently in SPSS. Instead of having a coding variable, and a single column with mischief scores in, we would arrange the data in two columns (one representing the **Cloak** condition and one representing the **No_Cloak** condition). The data are in **Invisibility RM.sav** if you had difficulty entering them into SPSS yourself.

Compute the paired-samples t -test

To conduct a paired-samples t -test, we need to access the main dialog box by selecting **Analyze** **Compare Means** **▶ Paired-Samples T Test...** (Figure 3). Once the dialog box is activated, you need to select pairs of variables to be analysed. In this case we have only one pair (**Cloak** vs. **No_Cloak**). To select a pair you should click on the first variable that you want to select (in this case **No_Cloak**), then hold down the *Ctrl* key (*Cmd* on a Mac) and select the second (in this case **Cloak**). To transfer these two variables to the box labelled *Paired Variables* click on . (You can also select each variable individually and transfer it by clicking on , but selecting both variables as just described is quicker.) If you want to carry out several t -tests then you can select another pair of variables, transfer them to the variables list, then select another pair and so on. If you click on  then

another dialog box appears that gives you the same options as for the independent t -test. Similarly we can click on **Bootstrap...** to access the bootstrap function (section **Error! Reference source not found.**). As with the independent t -test, select **Perform bootstrapping** and **Bias corrected accelerated (BCa)**. Back in the main dialog box click on **OK** to run the analysis.

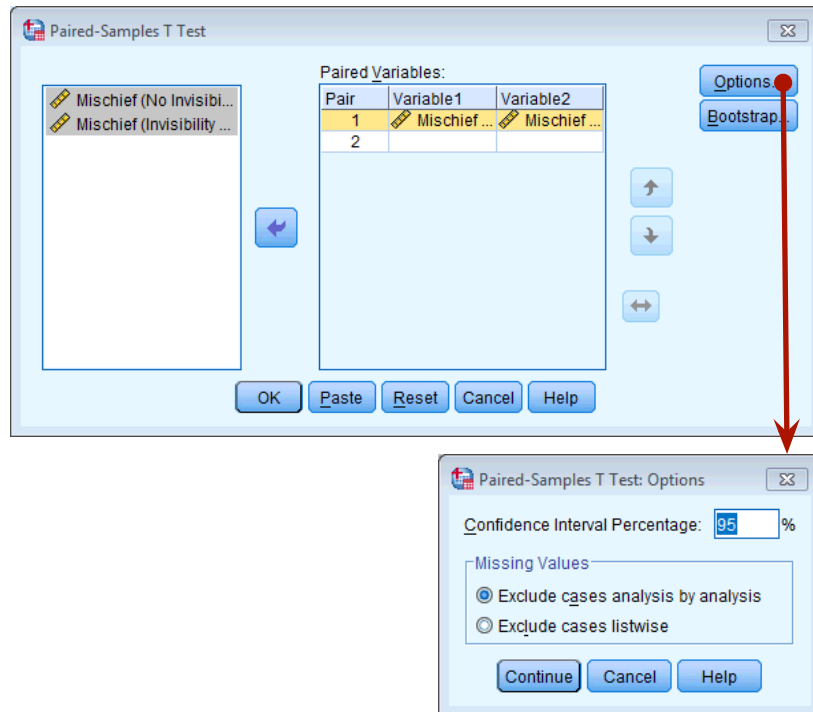


Figure 3: Main dialog box for paired-samples t -test

Output from the paired-samples t -test

The resulting output produces four tables (3 if you don't select bootstrapping). Output 4 shows a table of summary statistics for the two experimental conditions (if you don't ask for bootstrapping this table will be a bit more straightforward). For each condition we are told the mean, the number of participants (N), the standard deviation and standard error. These values are the same as when we treated the data as an independent design.

Output 4 also shows the Pearson correlation between the two conditions. When repeated measures are used it is possible that the experimental conditions will correlate (because the data in each condition come from the same people and so there could be some constancy in their responses). SPSS provides the value of Pearson's r and the two-tailed significance value. For these data the experimental conditions yield a very large correlation coefficient, $r = .806$, which is highly significant, $p = .002$ and has a bootstrap confidence interval that doesn't include zero, BCa 95% CI [.185, .965].

Output 5 shows us whether the difference between the means of the two conditions was significant. First, the table tells us the mean difference between the mean scores of each condition: $3.75 - 5 = -1.25$. The table also reports the standard deviation of the differences between the means and more important the standard error of the differences between participants' scores in each condition. The size of the test statistic, t , is compared against known values based on the degrees of freedom. When the same participants have been used, the degrees of freedom are the sample



size minus 1 ($df = N - 1 = 11$). SPSS uses the degrees of freedom to calculate the exact probability that a value of t at least as big as the one obtained could occur if there was no difference between population means. This probability value is in the column labelled *Sig.* SPSS provides the two-tailed probability, which is the one I recommend using. Typically, We are interested in whether this value is less than or greater than .05, and because the value of p is less than .05 we can conclude that there was a significant difference between the means of these two samples. In terms of the experiment, we can infer that having a cloak of invisibility significantly affected the amount of mischief a person got up to, $t(11) = -3.80, p = .003$.

Paired Samples Statistics

			Statistic	Bootstrap ^a			
				Bias	Std. Error	BCa 95% Confidence Interval	
						Lower	Upper
Pair 1	Mischief (No Invisibility Cloak)	Mean	3.75	-.02	.52	2.83	4.50
		N	12				
	Mischief (Invisibility Cloak)	Mean	5.00	.00	.46	4.17	5.75
		N	12				
		Std. Deviation	1.913	-.110	.341	1.311	2.234
		Std. Error Mean	.552				
		Std. Deviation	1.651	-.087	.309	1.114	1.992
		Std. Error Mean	.477				

a. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples

Paired Samples Correlations

				Bootstrap for Correlation ^a				
				Bias	Std. Error	BCa 95% Confidence Interval		
						Lower	Upper	
Pair 1	Mischief (No Invisibility Cloak) & Mischief (Invisibility Cloak)	N	Correlation	Sig.	Bias	Std. Error	Lower	Upper
		12	.806	.002	-.019	.160	.185	.965

a. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples

Output 4

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	Mischief (No Invisibility Cloak) - Mischief (Invisibility Cloak)	-1.250	1.138	.329	-1.973	-.527	-3.804	11	.003

Output 5

Finally, this output provides a 95% confidence interval for the mean difference. However, a more robust confidence interval, estimated using bootstrapping, is produced in Output 6. remember that confidence intervals are constructed such that in 95 of samples the intervals contain the true value of the mean difference. So, assuming that this sample's confidence interval is one of the 95 out of 100 that contains the population value, we can say that the true mean difference lies between -1.67 and -0.83 . The importance of this interval is that it does not contain zero (both limits are negative), which tells us that the true value of the mean difference is unlikely to be zero. In other words, there is an effect in the population reflecting more mischievous acts performed when someone is given an invisibility cloak.



Bootstrap for Paired Samples Test

		Mean	Bootstrap ^a				
			Bias	Std. Error	Sig. (2-tailed)	BCa 95% Confidence Interval	
					Lower	Upper	
Pair 1	Mischief (No Invisibility Cloak) – Mischief (Invisibility Cloak)	-1.250	-.019	.323	.004	-1.667	-.833

a. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples

Output 6

Reporting the paired-samples t-test

We can basically report the same information for matched-samples *t*-test as the independent *t*-test, but obviously the confidence intervals, degrees of freedom and values of *t* and *p* have changed:

- ✓ On average, participants given a cloak of invisibility engaged in more acts of mischief ($M = 5, SE = 0.48$), than those not given a cloak ($M = 3.75, SE = 0.55$). This difference, -1.25 , BCa 95% CI $[-1.67, -0.83]$, was significant $t(11) = -3.80, p = .003$ and represented a medium-sized effect $d = .65$.

References

Field, A. P. (2013). *Discovering statistics using IBM SPSS Statistics: And sex and drugs and rock 'n' roll* (4th ed.). London: Sage.

Terms of Use

This handout contains material from:

Field, A. P. (2013). *Discovering statistics using SPSS: and sex and drugs and rock 'n' roll* (4th Edition). London: Sage.

This material is copyright Andy Field (2000-2016).

This document is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), basically you can use it for teaching and non-profit activities but not meddle with it without permission from the author.