

Running and Reading a One-Sample t test in SPSS versions 21-22

The following example uses the **Rikers 1989.sav** dataset from SPSS Work in Blackboard.

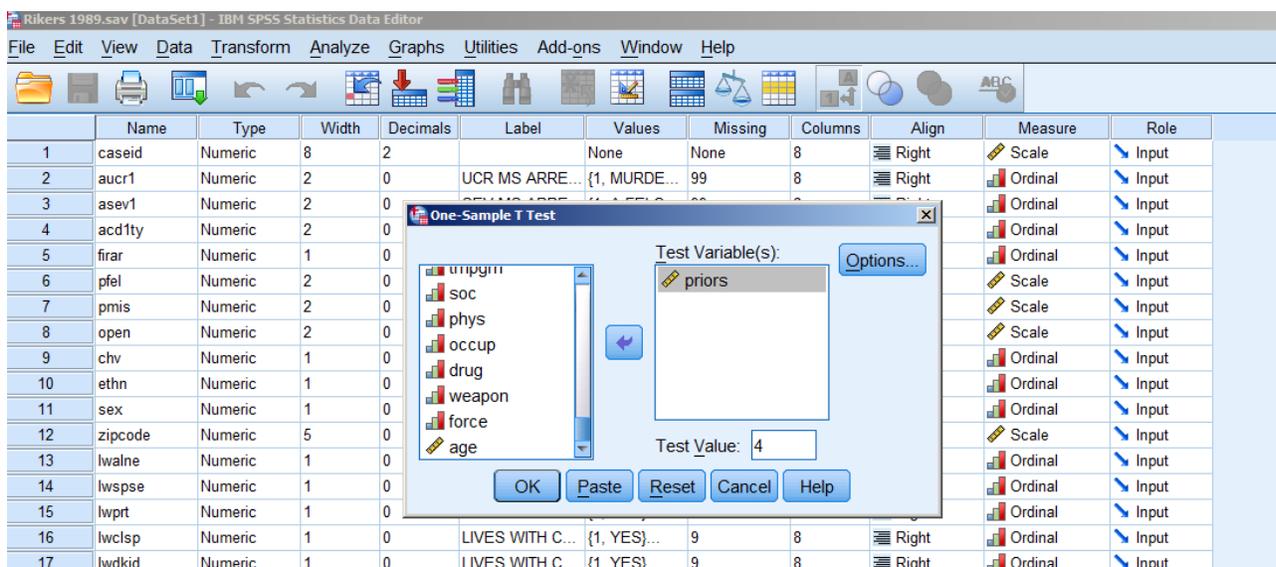
When we run a one-sample t test, we are testing the mean for a single sample against a particular value, called the **test value**. The test value could be determined in many ways – it could be the mean of the population from which the sample came (to test whether the sample is representative of the population for example), or it could be the known mean value of another group, or it could be a passing value on a particular test, etc.

So, for example I might want to test whether the average number of prior arrests reported by the arrestees in the Rikers dataset was statistically significantly different from 4 (the national average found in a study done in 1975). The variable in the Rikers 1989 dataset is **priors**. The null hypothesis for our test is:

$$H_0: \mu = 4$$

Essentially, our null hypothesis states that the average number of prior arrests for all arrestees reported in New York in 1989 was the same as the average of 4 prior arrests reported in 1975 and any result in our sample that is different from 4 is simply the result of sampling error. The Null Hypothesis is always a statement about a population parameter (in this example, that the average number of prior arrests for all arrestees in New York in 1989 was 4) and is tested using data from a sample.

In SPSS, we want to load the data set for the sample we are going to be testing. Once the data set is loaded, we then go to the Analyze menu → Compare Means → One Sample T Test. This action brings up the One-Sample T Test window. **Priors** is the **Test Variable** – the variable whose mean we are comparing to the population mean of 4. And that value (4) is added to the **Test Value** space.



The screenshot shows the IBM SPSS Statistics Data Editor window with the 'One-Sample T Test' dialog box open. The dialog box has a list of variables on the left, including 'priors', which is selected and moved to the 'Test Variable(s):' field. The 'Test Value:' field is set to 4. The background shows the variable list with columns for Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, Measure, and Role.

When you press OK to run the **One Sample T Test** procedure, you will see the output shown on the next page.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
# PRIOR ARRESTS	1038	5.11	7.958	.247

One-Sample Test

	Test Value = 4					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
# PRIOR ARRESTS	4.477	1037	.000	1.106	.62	1.59

The first box gives me the descriptive statistics for the variable **priors**. The average number of prior arrests for the 1038 arrestees according to the sample data is 5.11. That appears to be different from 4 but remember that I could get a sample of arrestees that might differ on average from the population mean of 4 because of sampling error. The *t* test results determine if the result is statistically significantly different from 4 – is the probability that the result is due to sampling error low enough that I can reject sampling error as the reason my sample has a mean that is different from 4?

The *t* test results give you the value for *t* (**the test statistic**), which is $t = 4.477$, the degrees of freedom for the test (*df*) and the Sig. value. The *t* test value in a test with this large a sample is exactly the same as a *z* value – in other words, the mean of 5.11 is four and a half standard units above the population mean of 4. Remember that the population mean is the “expected value” of the mean of all the samples of $n = 1038$ (the sampling distribution of $n=1038$).

The Sig. value in SPSS represents our *p* value (*p* meaning probability). Specifically, the Sig. value tells us the probability of getting the sample mean value that we’re testing (in this case, 5.11 prior arrests) just because of sampling error. So, we can see that the probability of getting a sample mean value of 5.11 because of sampling error alone is $<.001$ or less than 1%. The standard cut-off for statistical significance is .05, or 5%. If our Sig. value falls *below* .05, then we reject the null hypothesis that the difference is caused by sampling error and conclude that the number of prior arrests for a jail sample in NYC in 1989 is statistically significantly different from the population mean of 4 (from 1975).

When we write this result up, it might look something like this:

The average number of prior arrests for detainees in New York in 1989 was statistically significantly greater than 4, $t(1037) = 4.48, p = .000$ (or $p < .01$).

First, you write a sentence that describes the results of the test (in our example, that it’s statistically significant) and the nature of the difference if there is one (our sample is *greater* than 4), then you include the statistics from the test itself. The sentence is your conclusion, the statistics are the evidence to back up your conclusion.

The One Sample T test results also provide the 95% confidence interval of the difference between the observed mean (5.11) and the population mean (4). That interval is .62 at the **lower bound** and 1.59 at the **upper bound**. This confidence interval can be understood to predict that 95% of the samples of 1038 that we could **possibly** test would result in an average number of prior arrests that would be at least 4.62 (.62 higher than 4) and as much as 5.62 (1.62 higher than 4). And our observed mean falls into this interval estimate.

