

# What is the necessary sample size for a one-sample t-test to detect a significant difference between the sample mean and the population mean, with a specified level of power and significance?

Authored by  
**stats writer**

June 29, 2024

## RECOMMENDED CITATION

stats writer (2024). *What is the necessary sample size for a one-sample t-test to detect a significant difference between the sample mean and the population mean, with a specified level of power and significance?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=159061>

A one-sample t-test is used to determine if there is a significant difference between the sample mean and the population mean. To ensure accurate and reliable results, it is important to have an appropriate sample size. The necessary sample size for a one-sample t-test depends on the desired level of power and significance. Power refers to the ability of the test to detect a true difference between the sample mean and the population mean, while significance refers to the threshold at which the difference is considered significant. Generally, a larger sample size is needed to achieve higher power and significance levels. Therefore, it is important to carefully consider these factors when determining the necessary sample size for a one-sample t-test.

## **Power Analysis for One-sample t-test**

### **Examples**

**Example 1. A company that manufactures light bulbs claims that a particular type of light bulb will last 850 hours on average with standard deviation of 50. A consumer protection group thinks that the manufacturer has overestimated the lifespan of their light bulbs by about 40 hours. How many light bulbs does the consumer protection group have to test in order to prove their point with reasonable confidence?**

**Example 2. It has been estimated that the average height of American white male adults is 70 inches. It has also been postulated that there is a positive correlation between height and intelligence. If**

**this is true, then the average height of a white male graduate students on campus should be greater than the average height of American white male adults in general. You want to test this theory out by random sampling a small group of white male graduate students. But you need to know how small the group can be or how few people that you need to measure such that you can still prove your point.**

#### **Prelude to the Power Analysis**

**For the power analysis below, we are going to focus on Example 1, testing the average lifespan of a light bulb. Our first goal is to figure out the number of light bulbs that need to be tested. That is, we will determine the sample size for a given a significance level and power. Next, we will reverse the process and determine the power, given the sample size and the significance**

**level.**

**We know so far that the manufacturer claims that the average lifespan of the light bulb is 850 with the standard deviation of 50, and the consumer protection group believes that the manufacturer has overestimated by about 40 hours. So in terms of hypotheses, our null hypothesis is  $H_0 = 850$  and our alternative hypothesis is  $H_a = 810$ .**

**The significance level is the probability of a Type I error, that is the probability of rejecting  $H_0$  when it is actually true. We will set it at the .05 level. The power of the test against  $H_a$  is the probability of that the test rejects  $H_0$ . We will set it at .90 level.**

**We are almost ready for our power analysis. But let's talk about the standard deviation a little bit. Intuitively, the number of light bulbs we need to test depends on the variability of the lifespan of these light bulbs. Take an extreme**

case where all the light bulbs have exactly the same lifespan. Then we just need to check a single light bulb to prove our point. Of course, this will never happen. On the other hand, suppose that some light bulbs last for 1000 hours and some only last 500 hours. We will have to select quite a few of light bulbs to cover all the ground. Therefore, the standard deviation for the distribution of the lifespan of the light bulbs will play an important role in determining the sample size.

### Power Analysis

In SPSS, it is fairly straightforward to perform a power analysis for comparing means. For example, we can use SPSS's power mean onesample command for our calculation as shown below using the information we discussed above.

power means onesample

/parameters test=nondirectional significance=0.05

**power=0.9 sd=50 mean=850 null=810.**

**Power Analysis - One-Sample Mean**

Power Analysis Table

	N	Actual Power <sup>b</sup>	Test Assumptions			
			Power	Std. Dev.	Effect Size	Sig.
Test for Mean <sup>a</sup>	19	.909	.9	50	.800	.05

a. Two-sided test.

b. Based on noncentral t-distribution.

The result tells us that we need a sample size at least 19 light bulbs to reject H0 under the alternative hypothesis Ha to have a power of 0.9.

Next, suppose we have a sample of size 10. How much power do we have keeping all of the other numbers the same? We can use the same command, power means onesample, to calculate it.

power means onesample

/parameters test=nondirectional significance=0.05 n=10 sd=50 mean=850 null=810.

**Power Analysis - One-Sample Mean**

Power Analysis Table

	Power <sup>b</sup>	Test Assumptions			
		N	Std. Dev.	Effect Size	Sig.
Test for Mean <sup>a</sup>	.616	10	50	.800	.05

a. Two-sided test.

b. Based on noncentral t-distribution.

You can see that the power is about .62 for a sample size of 10.

What then is the power for sample size of 15?

power means onesample

/parameters test=nondirectional significance=0.05 n=15  
sd=50 mean=850 null=810.

#### Power Analysis - One-Sample Mean

Power Analysis Table

	Power <sup>b</sup>	Test Assumptions			
		N	Std. Dev.	Effect Size	Sig.
Test for Mean <sup>a</sup>	.821	15	50	.600	.05

a. Two-sided test.

b. Based on noncentral t-distribution.

So now the power is about .82. You could also do it again to find out the power for a sample size of 20. You'll probably expect that the power will be greater.

power means onesample

/parameters test=nondirectional significance=0.05 n=20  
sd=50 mean=850 null=810.

**Power Analysis - One-Sample Mean**

Power Analysis Table

	Power <sup>b</sup>	Test Assumptions			
		N	Std. Dev.	Effect Size	Sig.
Test for Mean <sup>a</sup>	.924	20	50	.800	.05

a. Two-sided test.

b. Based on noncentral t-distribution.

**We can also expect that if we specified a lower power or the standard deviation is smaller, then the sample size should also be smaller. We can experiment with different values of power and standard deviation as shown below.**

**power means onesample /parameters test=nondirectional significance=0.05 power=0.8 sd=50 mean=850 null=810.**

**Power Analysis - One-Sample Mean**

Power Analysis Table

	N	Actual Power <sup>b</sup>	Test Assumptions			
			Power	Std Dev	Effect Size	Sig.
Test for Mean <sup>a</sup>	15	.821	.8	50	.800	.05

a. Two-sided test.

b. Based on noncentral t-distribution.

**If the standard deviation is lower, then the sample size should also go down, as we discussed before.**

**power means onesample**

**/parameters test=nondirectional significance=0.05  
power=0.8 sd=30 mean=850 null=810.**

**Power Analysis - One-Sample Mean**

Power Analysis Table

	N	Actual Power <sup>b</sup>	Test Assumptions			
			Power	Std. Dev.	Effect Size	Sig.
Test for Mean <sup>a</sup>	7	.834	.8	30	1.333	.05

a. Two-sided test.

b. Based on noncentral t-distribution.

## Discussion

There is another technical assumption, the normality assumption. If the variable is not normally distributed, a small sample size usually will not have the power indicated in the results, because those results are calculated using the common method based on the normality assumption. It might not even be a good idea to do a t-test on such a small sample to begin with if the normality assumption is in question.

Here is another technical point. What we really need to know is the difference between the two means, not the individual values. In fact, what really matters is the difference of the means

over the standard deviation. We call this the effect size. For example, we would get the same power if we subtracted 800 from each mean, changing 850 to 50 and 810 to 10.

power means onesample

/parameters test=nondirectional significance=0.05 power=0.8 sd=50 mean=50 null=10.

**Power Analysis - One-Sample Mean**

	N	Actual Power <sup>b</sup>	Test Assumptions			
			Power	Std. Dev.	Effect Size	Sig.
Test for Mean <sup>a</sup>	15	.821	.8	50	.800	.05

a. Two-sided test.

b. Based on noncentral t-distribution.

If we standardize our variable, we can calculate the means in terms of change in standard deviations.

power means onesample

/parameters test=nondirectional significance=0.05 power=0.8 sd=1 mean=1 null=0.2.

**Power Analysis - One-Sample Mean**

	N	Actual Power <sup>b</sup>	Test Assumptions			
			Power	Std. Dev.	Effect Size	Sig.
Test for Mean <sup>a</sup>	15	.821	.8	1	.800	.05

a. Two-sided test.

b. Based on noncentral t-distribution.

**It is usually not an easy task to determine the "true" effect size. We make our best guess based upon the existing literature, a pilot study or the smallest effect size of interest. A good estimate of the effect size is the key to a successful power analysis.**

**See Also**

**D. Moore and G. McCabe, Introduction to the Practice of Statistics, Third Edition, Section 6.4.**

ARABPSYCHOLOGY.COM