

One-sample t-test

Relevant research questions and data requirements

Research question: One-sample t-tests are mean difference tests. Specifically, we are seeking to compare if a population mean is the same as (or different from) a specific test value (i.e., $H_0: \mu_{\text{population of interest}} = \text{test value}$). Test values can come from previous research, certain legal requirements (e.g., the legal blood alcohol content or safe levels of toxins in water), standards (e.g., how many words a stenographer must be able to type in 10 seconds), or specific values of interest unique to your research. Because we cannot usually get data from everyone in our population of interest, we take a sample and calculate its mean and use it in our test. We make inferences from our sample to our population – hence, this is an example of Inferential Statistics.

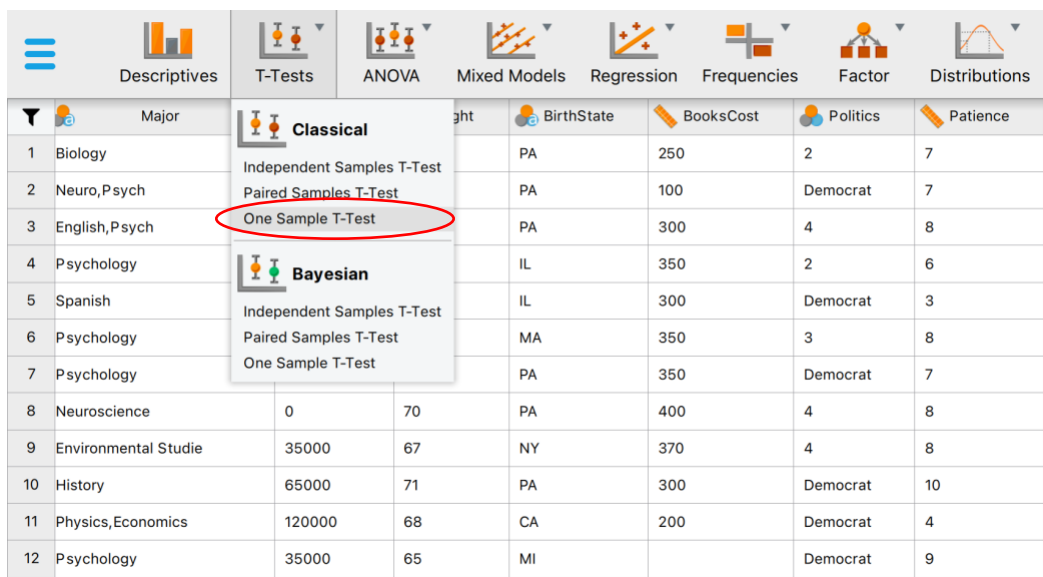
Let's say we know from a publication that the average height of adults in the US is 66.5 inches and we want to know if Allegheny students' height differs from that. Because we cannot measure every Allegheny student, we take a sample of 35 students and measure their height, which we then compare to the test value of 66.5 inches.

How many groups? One (the sample taken from your population of interest), will be compared to the test value

Data requirements? Interval/ratio outcome data (here: height), ideally the data are (roughly) normally distributed and independent (scores on height are not somehow all related, like sampling several participants from one family). Check for outliers.

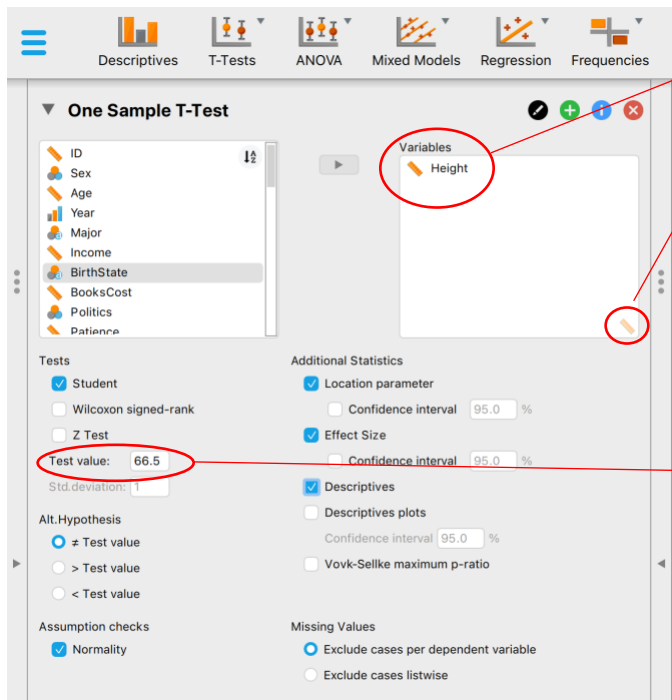
Running the test in JASP

Once you have opened your data in JASP, click on “t-tests” in the main menu and select “one-sample t” under “classical” (we will ignore all Bayesian analyses for now)



The screenshot shows the JASP software interface. The top menu bar includes Descriptives, T-Tests, ANOVA, Mixed Models, Regression, Frequencies, Factor, and Distributions. The 'T-Tests' menu is open, showing 'Classical' and 'Bayesian' sections. Under 'Classical', 'One Sample T-Test' is highlighted with a red circle. The background shows a data table with columns: Major, Height, BirthState, BooksCost, Politics, and Patience. The data table has 12 rows of student data.

Major	Height	BirthState	BooksCost	Politics	Patience	
1 Biology		PA	250	2	7	
2 Neuro,Psych		PA	100	Democrat	7	
3 English,Psych		PA	300	4	8	
4 Psychology		IL	350	2	6	
5 Spanish		IL	300	Democrat	3	
6 Psychology		MA	350	3	8	
7 Psychology		PA	350	Democrat	7	
8 Neuroscience	0	70	PA	4	8	
9 Environmental Studie	35000	67	NY	4	8	
10 History	65000	71	PA	300	Democrat	10
11 Physics,Economics	120000	68	CA	200	Democrat	4
12 Psychology	35000	65	MI		Democrat	9



Move your variable of interest into the “variable” window. Notice the little ruler in the bottom right reminds you this variable must be identified as a scale variable in your data set.

Importantly, you need to tell JASP here what your test value is. This should be the value specified in your null hypothesis. In our example, it is 66.5 inches – the average height of a US adult

Tests: Select “Student” (Fun fact: This is not a special test for students, but named after a Statistician who published about the t-distribution by the synonym “Student”). Wilcoxon signed-rank is the nonparametric equivalent, i.e., the test to use when your data are not normally distributed (see the output below to tell whether data are normally distributed)

Alt. hypothesis: Here, the two-tailed option is chosen automatically as the alternative hypothesis, i.e. “Allegheny students’ height is *different from* 66.5 inches”. (Recall a one-tailed test would specify a direction, e.g. we might hypothesize “Allegheny students are *taller than* 66.5 in.” Unless you have good reason to run a one-tailed test, you should always run a two-tailed test because it requires slightly more extreme cut-offs for the critical region, i.e. it is more stringent.)

Assumption Checks: Check “Normality” to test the hypothesis that height in your sample is, in fact, (roughly) normally distributed. Note: You can also get a visual impression by running the distribution plot under Descriptives as outlined in the handout on “Descriptive Statistics”

Additional Statistics:

- Check “Location Parameter” – this is the difference between the sample mean and the test value; then click “confidence interval”
- Check “effect size” – this will calculate Cohen’s d (i.e., the distance between the sample mean and the test value in SD units)
- Descriptives will calculate M and SD as well as standard error (SE) for your sample

Reading and understanding the output

Results

One Sample T-Test

One Sample T-Test

	t	df	p	Sample Mean	95% CI for Sample Mean		Cohen's d
					Lower	Upper	
Height	1.683	34	0.101	67.457	66.302	68.613	20.055

Note. For the Student t-test, effect size is given by Cohen's d.

Note. For the Student t-test, location estimate is given by the sample mean d.

Note. For the Student t-test, the alternative hypothesis specifies that the mean is different from 66.5.

Note. Student's t-test.

Assumption Checks

Test of Normality (Shapiro-Wilk)

	W	p
Height	0.951	0.119

Note. Significant results suggest a deviation from normality.

Descriptives

Descriptives

	N	Mean	SD	SE
Height	35	67.457	3.364	0.569

This is our hypothesis test. Shown in order are the t-value, the df (for a one-sample t this is n-1), the p-value, the sample mean, the 95% CI around the sample mean, and the effect size Cohen's d.

Here we learn that the average height of our sample is **67.46** inches. Even though our sample is almost one inch taller on average than the average US adult, this is not enough for us to reject the null $H_0: \mu_{\text{Allegheny}} = 66.50$ because **$p > .05$** . Our sample mean is therefore still within sampling error, i.e. not different enough to be significantly different from our test value.

This is the **95% CI** around the sample mean. Note that the value of the null hypothesis ($\mu = 66.50$) is contained in the interval. Thus, 66.50 is a likely value of the true height of the population of interest (Allegheny students) and therefore not statistically different from the US population.

Cohen's d is 20.055, which is huge. *In this version of JASP (0.14.1), this is a miscalculation, as a small yellow warning window will tell you (see below). To get the actual value, you need to calculate $(M - \mu) / SD$, i.e. $(67.457 - 66.50) / 3.364 = 0.28$. This tells us our sample mean is 0.28 SD above 66.50 inches. This makes it a small effect ($d = 0.20$ is small, $d = 0.50$ is medium, $d = 0.80$ is a large effect)*

Known bug: computed effect sizes ignore the test value

Under **Assumption Checks** we can see that the test of normality is not significant ($p = .119$, which is $> .05$). This means the null hypothesis that normality can be assumed is not rejected, so there are no serious deviations from normality.

Descriptives tell us sample size, mean, standard deviation, and standard error (SD / \sqrt{n})

Writing up results in APA style

A one-sample t-test indicated that Allegheny students' height ($n = 35$, $M = 67.46$, $SD = 3.36$) did not differ from the average US adult's height ($\mu = 66.50$), $t(34) = 1.68$, $p = .101$, 95% CI [-0.20; 2.11], Cohen's $d = 0.29$.

*Note that there are spaces before and after equal signs; M, SD, t, p are italicized; everything is rounded to two digits *except* for p-values, which should be reported exactly as given in the output. Only report leading zeros for values that can exceed 1 (hence, p-values should *not* be reported with leading 0s)