

HYPOTHESIS TESTING WITH SPSS:

A NON-STATISTICIAN'S GUIDE & TUTORIAL

by
Dr. Jim Mirabella

SPSS 14.0 screenshots reprinted with permission from SPSS Inc.
Published June 2006
Copyright Dr. Jim Mirabella

CHAPTER 2: Hypothesis Test for One Sample

In this chapter we shall walk through the steps for conducting a t-Test for One Sample. For this test, a single variable is chosen, a sample mean is computed, and that sample mean is compared to some specified value.

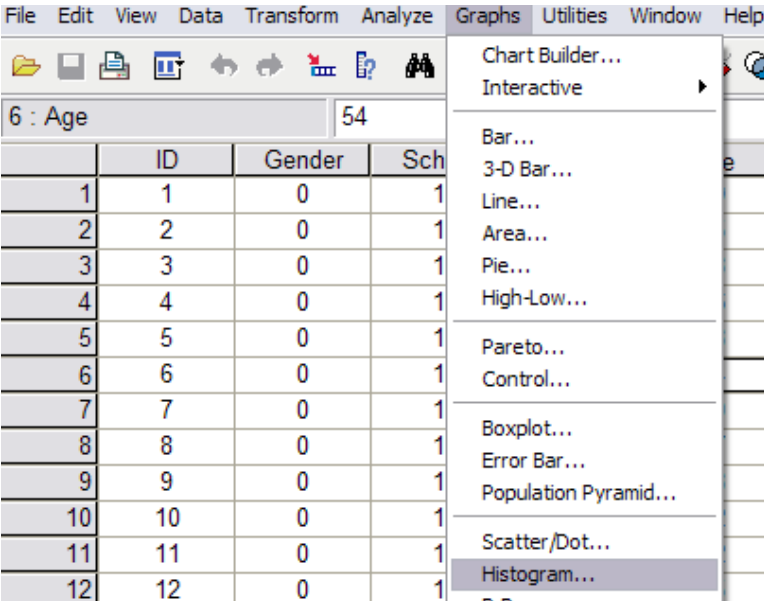
Suppose you wanted to know whether the PhD learners have a mean GPA that differs from 3.50. The null hypothesis (H_0) and alternate hypothesis (H_a) would be:

H_0 : The mean GPA of PhD learners equals 3.50

H_a : The mean GPA of PhD learners does not equal 3.50

The only variable being tested here is the GPA. Note that "PhD learners" is not a variable since all of the data being analyzed is from PhD learners (it is merely a descriptor for the target population in this case).

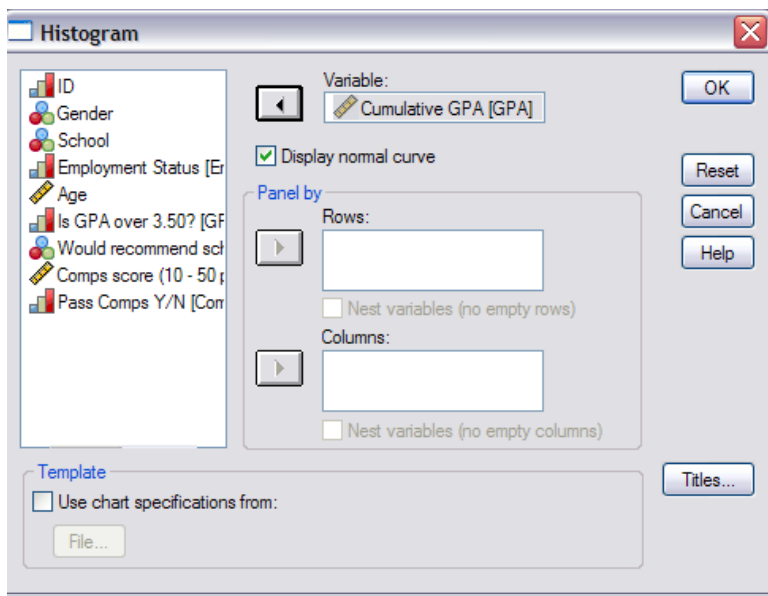
Since we are testing a sample mean against a hypothesized value, we shall use a t-Test for One Sample. To do so, there is an assumption that the GPA is normally distributed.



	ID	Gender	Sch
1	1	0	1
2	2	0	1
3	3	0	1
4	4	0	1
5	5	0	1
6	6	0	1
7	7	0	1
8	8	0	1
9	9	0	1
10	10	0	1
11	11	0	1
12	12	0	1

We really only need to visually check the data to see if it "appears" normal.

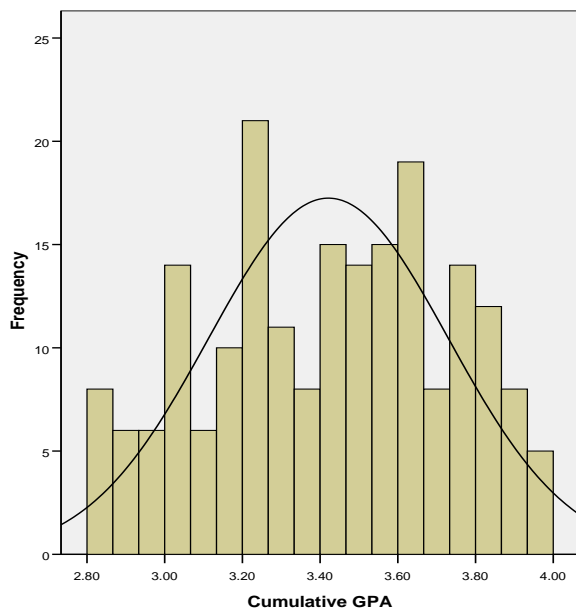
Go to *Graphs* → *Histogram*



Choose GPA as the Variable.

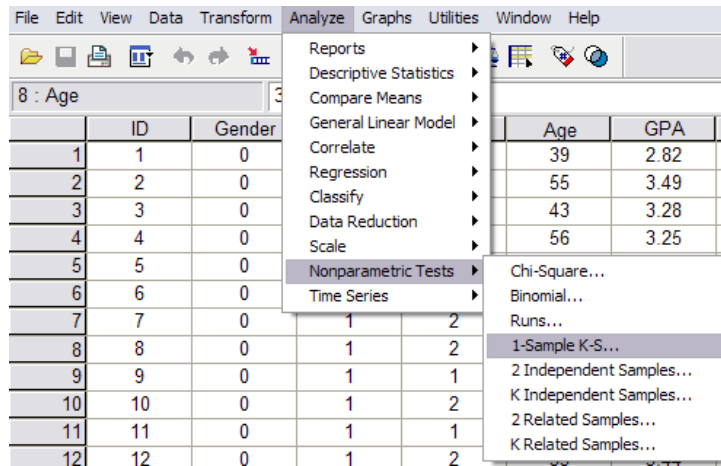
Check the box next to *Display normal curve*.

Click *OK*

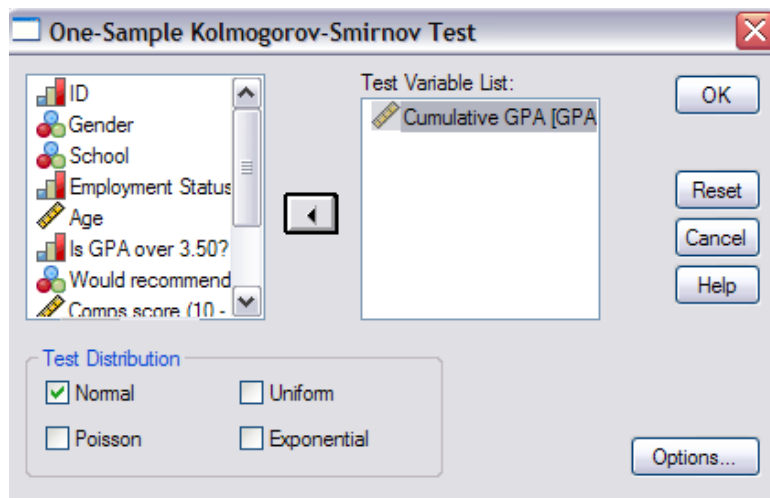


Mean = 3.42
Std. Dev. = 0.52
N = 200

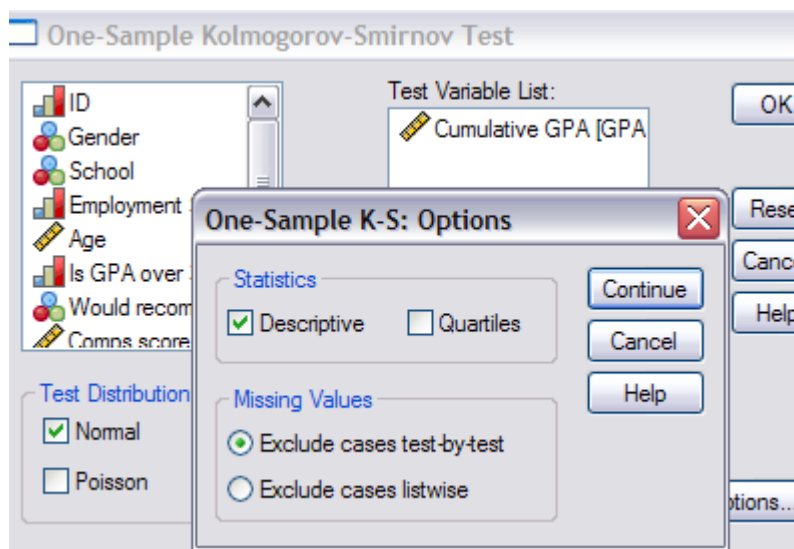
Now a truly normal curve is shaped like a bell that peaks in the middle and is perfectly symmetrical. This histogram does not appear to have a perfect bell-shaped pattern, but I wouldn't expect it to (remember that you are only looking at a sample, not the entire population). There is more data in the middle, and less toward the two extremes. Additionally, it appears that about half of the data are above and half are below the mean. Also, there aren't any unusually large or small numbers at either extreme (known as outliers). Based on these observations, the assumption of normality appears reasonable, so we can proceed with the t-test. If you weren't sure, you could conduct a Kolmogorov-Smirnov test to evaluate the normality assumption.



To conduct the Kolmogorov-Smirnov test, go to *Analyze* → *Nonparametric Tests* → *1-Sample K-S*



Choose *GPA* as the *Test Variable* and check the *Normal* box since that is what we wish to test for. I recommend you click *Options* next.



Check the *Descriptives* option to generate descriptive statistics (usually includes the mean /average, standard deviation, sample size, minimum & maximum values, and sometimes a few other useful statistics). Then click *Continue* and *OK*.

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Cumulative GPA	200	3.4215	.30841	2.80	3.99

One-Sample Kolmogorov-Smirnov Test

		Cumulative GPA
N		200
Normal Parameters ^{a,b}	Mean	3.4215
	Std. Deviation	.30841
Most Extreme Differences	Absolute	.076
	Positive	.058
	Negative	-.076
Kolmogorov-Smirnov Z		1.071
Asymp. Sig. (2-tailed)		.202

a. Test distribution is Normal.

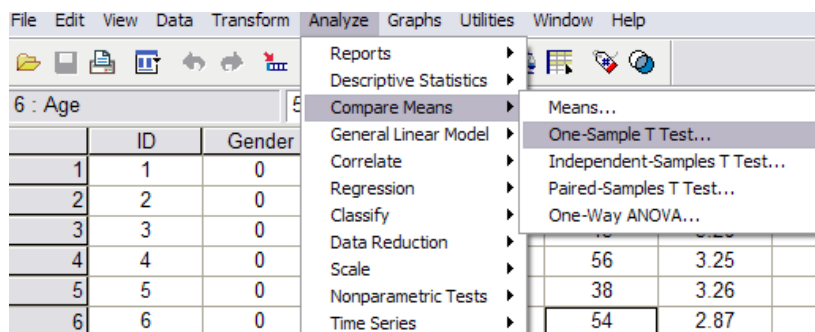
b. Calculated from data.

The output shows an *N* of 200 (this is the sample size) and a *Mean* of 3.4215 (this is the sample average for the 200 learners), with a *Minimum GPA* of 2.80 and a *Maximum* of 3.99. It also shows an *Asymp. Sig. (2-tailed)* value of .202 (this is also known as the p-value). The p-value tells you the probability of getting the results you got if the null were actually true (i.e., it is the probability you would be in error if you rejected the null hypothesis). While it is never stated in your analysis, the hypotheses for this test of normality are:

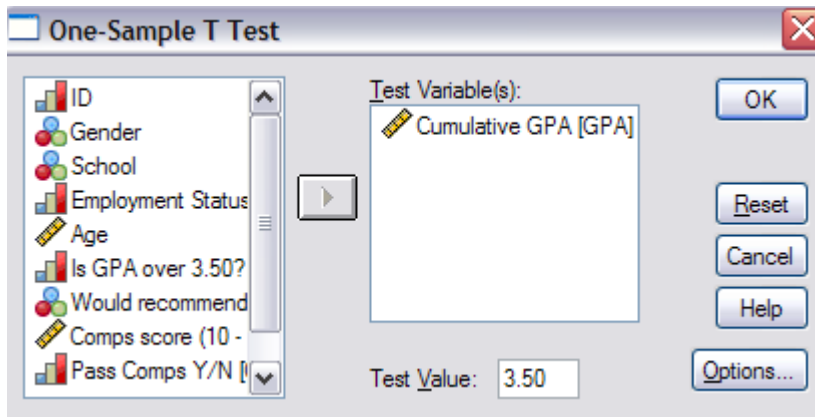
Ho: The distribution of GPAs is normal.

Ha: The distribution of GPAs is not normal.

If the p-value is less than .05, you reject the normality assumption, and if the p-value is greater than .05, there is insufficient evidence to suggest the distribution is not normal (meaning that you can proceed with the assumption of normality). Since the p-value is .202, there is no reason to doubt the distribution is normal, so you can safely proceed with the t-test.



To conduct the t-test, go to *Analyze* → *Compare Means* → *One-Sample T Test*



Choose GPA as the *Test Variable* and 3.50 as the *Test Value*. Then click *OK*.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Cumulative GPA	200	3.4215	.30841	.02181

One-Sample Test

	Test Value = 3.50					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Cumulative GPA	-3.602	199	.000	-.07855	-.1216	-.0355

The above output shows an *N* of 200 and a *Mean* of 3.4215, as we already knew from the prior test. The t-test output has a *Sig. (2-tailed)* / p-value of .000. A p-value of .000 means that the probability of a randomly drawing a sample of 200 from a population with a mean of 3.50 and getting a sample mean as low as 3.42 purely by chance is 0.00%. In other words, it is unlikely to have occurred by chance and is more likely the case that the mean is not as hypothesized.

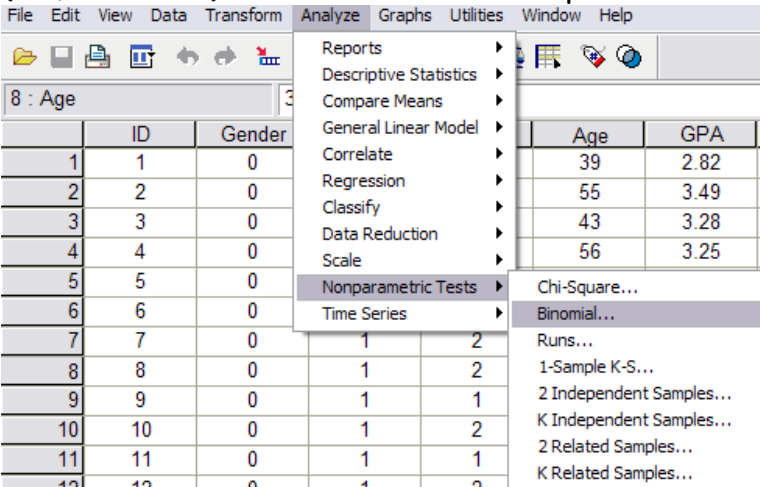
We typically set a significance level at .05, but sometimes we adjust it to as little as .01 or as much as .10. Our decision to adjust it is based on our tolerance for the two types of error (i.e., rejecting the null hypothesis that is true vs. not rejecting a null that is false). For now let's go with .05, which has become a default for most studies. Since the p-value is less than .05 (our chosen significance level), we reject the null. When we reject the null, we are basically declaring the alternate hypothesis to be true; when we fail to reject the null, we state that there is insufficient evidence to declare the alternate hypothesis to be true (but you need to write in accordance with the wording of the specific hypothesis instead of just making a generic statement about the null or alternate, as shown in the next paragraph).

Had we not rejected the null hypothesis, we would state that there is insufficient evidence to conclude the mean GPA differs from 3.50. Under no circumstances should you ever "accept" the null and/or conclude the mean equals 3.50. It is impossible to prove the null is true. Hypothesis testing is all about gathering evidence to suggest the null is not true, and the lack of such evidence warrants a "Do not reject" decision. Think about the courtroom where we find a defendant "guilty" or "not guilty". We never declare the defendant "innocent", and a "not guilty" verdict means there was insufficient evidence to find the defendant guilty (that is, evidence was likely presented, but it just was not enough to convince the jurors to deliver a guilty verdict).

As for this case of testing the GPA, however, we did reject the null hypothesis, so we should conclude that the mean GPA of PhD learners differs from 3.50. Since the sample mean is 3.42, we can get more specific and state that the mean GPA is less than 3.50.

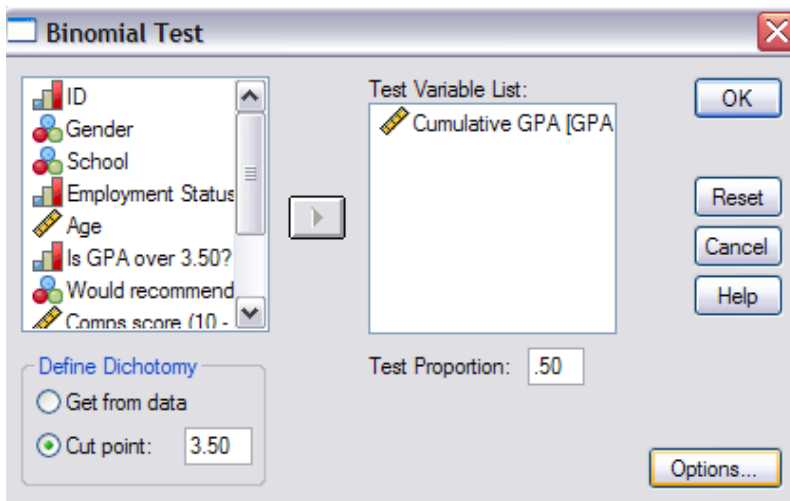
At this point, I recommend going one extra step with the conclusion. Does it make sense? What does it really mean? In this case, you might state that, contrary to rumors, grades do not appear to be inflated, as the mean GPA leans closer to the B range than to the A range. Anything more should be saved for Chapter 5 of your dissertation, when you tie the results to the literature and make suggestions for future research.

Now what if the normality assumption did not hold up, or the sample size was relatively small (often considered to mean less than 30 per sample)? In such cases, you can conduct a nonparametric test that essentially tests the same principle without the parameter (i.e., the mean) and without the assumptions. In this case, we shall use a Binomial test.

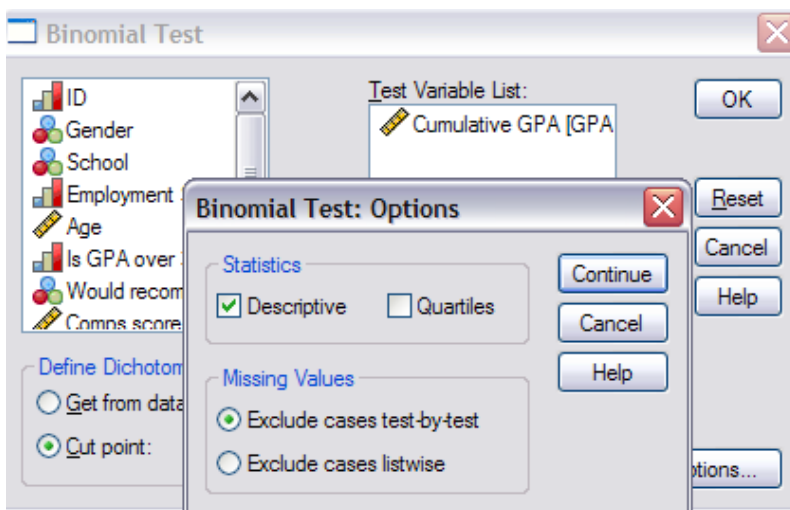


	Age	GPA
1	39	2.82
2	55	3.49
3	43	3.28
4	56	3.25

To conduct the Binomial test, go to *Analyze* → *Nonparametric Tests* → *Binomial*



Choose *GPA* as the *Test Variable*, select a *Cut point* of 3.50 and a *Test Proportion* of .50. This will test if 50% of the GPAs are above 3.50 and 50% are below. In a normal distribution where the mean is 3.50, you would expect to see a perfect 50/50 split, so this is essentially the equivalent without testing the mean.



If you choose *Options*, you will see an opportunity to get the *Descriptive Statistics*. This is the same output you saw in the Kolmogorov-Smirnov test. Click *Continue* and *OK*.

Binomial Test

	Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (2-tailed)
Cumulative GPA	Group 1 <= 3.5	111	.56	.50	.137 ^a
	Group 2 > 3.5	89	.45		
Total		200	1.00		

a. Based on Z Approximation.

The above output shows that the *GPA* was broken down into two groups: ≤ 3.5 and > 3.5 . 56% of the GPAs are less than 3.5 and 44% are greater. If the mean were 3.50 and the distribution were normal, you would expect to see 50% above / below 3.50. The p-value of .137 is greater than .05, so the null hypothesis is not rejected and there is insufficient evidence to conclude that the percentage of GPAs above 3.50 is not 50%.

Notice that the t-test resulted in rejecting H_0 , while the Binomial test did not. This is not unusual. A parametric test (like the t-test) is more powerful and more capable of detecting significant differences, while a nonparametric test (like the Binomial test) is weaker and more conservative in its likelihood of finding a difference to be significant. This is why we try to use parametric tests whenever possible, but when it is not an option, at least we have a viable alternative, albeit a little weaker.

So let's review.

State the hypotheses:

H_0 : The mean GPA of PhD learners equals 3.50

H_a : The mean GPA of PhD learners does not equal 3.50

Choose a significance level \rightarrow .05

State the assumption(s):

PhD GPAs are normally distributed.

\rightarrow evaluate graphically or with a Kolmogorov-Smirnov test

Conduct t-Test for One Sample

Compare the sig.value / p-value to .05. If greater than .05, do not reject H_0 and then state that there is insufficient evidence to conclude the mean GPA differs from 3.50. If less than .05, reject H_0 and conclude that the mean GPA in the PhD program is not 3.50 (specify whether it is greater or less than 3.50). Then feel free to add some insights in English, but be careful not to overstate beyond what you tested.

If the normality assumption does not hold, conduct a Binomial test with the following hypothesis \rightarrow

H_0 : The GPAs of PhD learners are distributed with 50% above 3.50