Hypothesis testing

Examples:

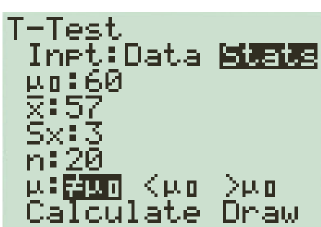
I. Claim: the mean temperature in the town of ALPHA for the month of January is 60 degrees Fahrenheit. A sample of 20 days is taken, the mean temp for those 20 days is 57 degrees, the standard deviation is 3.0 degrees. Use a 0.01 level of significance to test the claim.

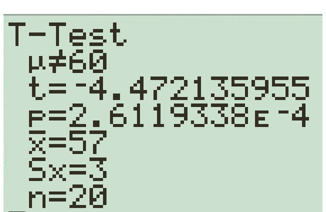
H_0: $\mu = 60$
H_1: $\mu \neq 60$

Note: the claim states a value; Temp (..) is equal to 60; it does not say greater or less, just a value. In this case the alternative is not equal to (≠).
The sample mean and standard deviations are given; therefore, this is a T-Test.
Level of significance is 0.01 ($\alpha = 0.01$); because the alternative hypothesis uses ≠, the test is a two tailed test. The t(alpha/2) critical value, for 20-1 = 19 degrees of freedom by the T-Table is 2.861. Remember, the critical value is needed in order to establish our conclusions, when we don't use a graphing calculator: we compare the T-test statistic, which we are calculating below, to the critical value: when the absolute value of the test-test is greater than the critical value we Reject the Null hypothesis.
Let's find the test stat:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{57 - 60}{\frac{3}{\sqrt{20}}} = -4.47$$

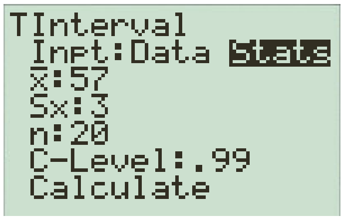
Calculator:

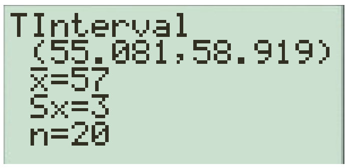




The output of the t-test shows a t-statistic, t = -4.47 and a p-value = 0.00026 or 2.6 x 10-4

$$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} = 57 \pm 2.861\frac{3}{\sqrt{20}} = (55.1, 58.9)$$





Conclusions: Reject the Null Hypothesis. There are two ways of drawing this conclusion: the t-stat absolute value is greater than the critical value or, the p-value is less than the stated value of alpha.
In technical terms: we have enough sample evidence to warrant rejection of the claim that the mean temperature is 60 degrees Fahrenheit.
Notice that the interval (55.1, 58.9) does not include the stated population parameter of 60.
That fact is consistent with the conclusion of rejecting H_0. The sample describes a population whose mean is between 55.1 and 58.9, range of values that exclude 60.

II. The mayor of the town of ALPHA claims that over 60% of the citizens approve his job as mayor. A random sample of 200 residents, all voters of the town, are interviewed; 130 of them agrees with the mayor. Use a 0.05 level of significance to test the claim.

H_0: p = 0.60
H_1: p > 0. 60

One proportion Z-Test.
α = 0.05; right tailed test.
$Z_{0.05}$= 1.645

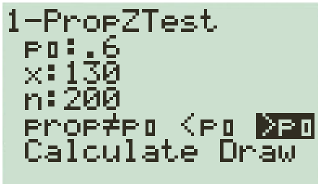
First, calculate the sample proportion:

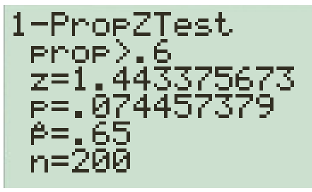
$$\hat{p} = \frac{x}{n} = \frac{130}{200} = 0.65$$

Z- test Statistic:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.65 - 0.60}{\sqrt{\frac{0.60 * 0.40}{200}}} = 1.44$$

Calculator:



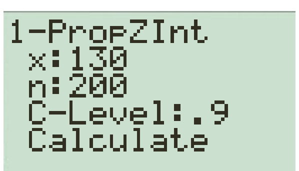


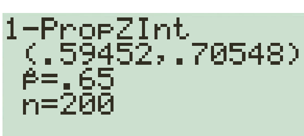
Calculator: test- statistic z =1.44; p-value = 0.074

Confidence interval:
This is a one tailed test, we construct the 1-2α conf. interval. This is the reason why: intervals include two tails, whenever the test is a one-tailed test, we need to create the other tail; so, we subtract two times alpha from one. In this case we construct a 1-2(0.05) = 0.90 Conf. Int.

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.65 \pm 1.645\sqrt{\frac{0.65 * 0.35}{200}} = (0.59, 0.71)$$

1-Prop ZInt by Calculator:





Notice that the z-test statistic (1.44) is less than the critical value (1.645); also, the p-value, 0.074 > alpha (0.01).

Conclusions: We fail to reject the Null hypothesis.
In technical terms: there is no sufficient evidence to support the claim that over 60% of the citizens approve the mayor's job.

The confidence interval includes 0.60 (0.59, 0.71); that is, the sample data is consistent with the Null hypothesis; therefore, the is no support for the claim that indeed the proportion is greater than 60%.

III. 1,500 women followed the Atkin's diet for a month. A random sample of 29 women gained an averag e of 6.7 pounds. Use a 0.05 level of significance to test the claim that the average weight gain per woman for the month was over 5 pounds. The standard deviation for all women in the group was 7.1

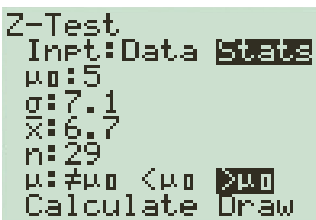
H_0: $\mu = 5$
H1: $\mu > 5$

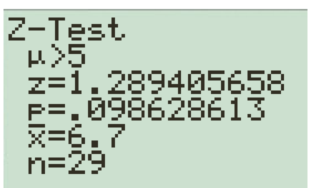
Significance level, α =0.05; right tailed test.
Since the problem states that *the standard deviation of all women was 7.*1; that is, sigma (σ) is given, so we use a Z-Test.
The $z_{0.05} = 1.645$, which is the critical value.

Test Statistic:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{6.7 - 5}{\frac{7.1}{\sqrt{29}}} = 1.29$$

Calculator: Z-Test:

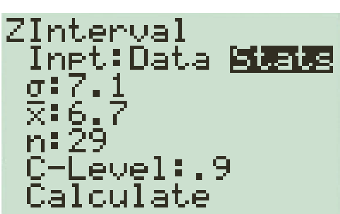


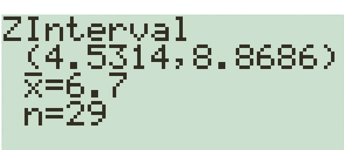


The 1-2alpha CI: 1-2(0.05) = 0.90

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 6.7 \pm 1.645\frac{7.1}{\sqrt{29}} = (4.5, 8.9)$$

Calculator: Z-Int: 90% Conf. Int:





Conclusions: Fail to reject the Null hypothesis. According to the sample data there is no sufficient evidence to support the claim that the average weight gain per woman was over five pounds.
Notice that the confidence interval includes five; therefore, the true value of the population parameter may be five, not a value significantly greater than five. This fact is consistent with the decision of not rejecting the Null hypothesis ($\mu = 5$).

IV: Weight Loss for Diet vs Exercise
Diet Only: sample mean = 5.9 kg; sample standard deviation = 4.1 kg; sample size: n = 42 Exercise Only: sample mean = 4.1 kg; sample standard deviation = 3.7 kg; n = 47
Did dieters lose more fat than the exercisers? Set $\alpha = 0.01$ and assume that the populations standard deviations are not equal.

2 Samples T Test

H_0: $\mu_1 = \mu_2$
H_1: $\mu_1 > \mu_2$

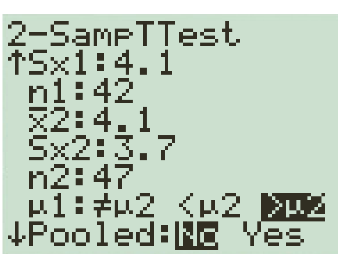
The question (claim) is: Does dieters lose more fat than the exercisers? Since we label first group (dieters) as 1, and exercisers as 2, the alternative hypothesis is symbolized: $\mu_1 > \mu_2$.

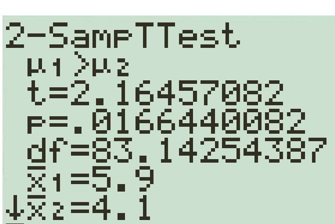
It is a right tailed test; the critical value is t = 2.423. It is the corresponding value for alpha = 0.01 and 41 df – degrees of freedom: smaller n-1. Notice that graphing calculators and statistical software generate different degrees of freedom by using Satterthwaite formula. Our textbook, Elementary Statistics by M. Triola, follows the rule smaller n-1. One method or the other does not change the final conclusion.

Test Statistic: We use this formula when the populations' variances are different for both groups. In the calculator menu this is achieved by choosing pooled No. This is called the Welch T test, the test by default in statistical software like R.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{5.9 - 4.1}{\sqrt{\frac{4.1^2}{42} + \frac{3.7^2}{47}}} = 2.16$$

Calculator:

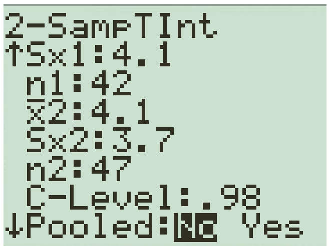


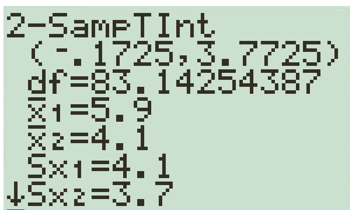


The output reads: t statistic = 2.16, p-value = 0.0166.

Again, we construct the 1-2α conf. interval: 1-2(0.01)=0.98

Two samples T Interval. TI calculator:





Notice that the test statistic is less than the t critical value, and the p-value is larger than alpha; also, the confidence interval includes zero. All this information suggests that the two populations' means are equal.

Conclusions: We fail to reject the Null Hypothesis. There is no sufficient evidence to answer yes to the question that dieters lose more weight than exercisers.

V: Two medications:

A medication for blood pressure was administered to a group of 13 randomly selected patients with ele vated blood pressure while a group of 15 was given a placebo. At the end of 3 months, the following dat a was obtained on their Systolic Blood Pressure. Control group: n=15, sample mean = 180, s=50. Trea ted group: n=13, sample mean =150, s=30. Test if the treatment has been effective, use α=0.01. Assume that the populations variances are not equal.

2 samples T Test.

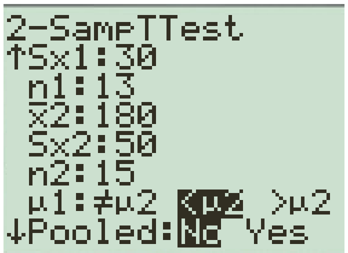
Label treatment group as 1; and control group as 2 (it could be the other way and we will draw the same conclusions): the researcher test whether the treatment is effective, in that case the mean systolic blood pressure of group 1 would be lower than the mean for group two.

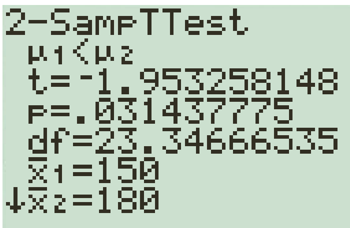
H_0: $\mu_1 = \mu_2$
H_1: $\mu_1 < \mu_2$

Left tailed test, alpha = 0.01; t critical = -2.681 (see T-Table).
Test Statistic:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{150 - 180}{\sqrt{\frac{30^2}{13} + \frac{50^2}{15}}} = -1.95$$

Calculator:

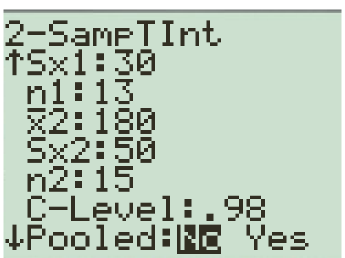


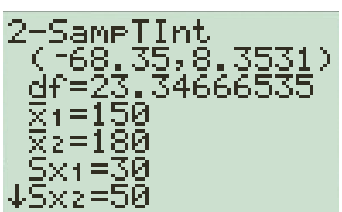


Test stat: t = -1.95; p-value: 0.0314

The $1-2\alpha$ conf. interval: 1-2(0.01)=0.98

TI Calculator: 2 sample T interval:





Notice that the test statistic absolute value is less than the critical value; accordingly, the p-value is of 0.0314 is greater than the stated significance level, and the confidence interval includes zero.

Conclusions: Fail to reject the Null. There is no sufficient evidence to support the researcher's claim that the treatment has been effective in lowering the systolic blood pressure.

VI: Another case of two medications.

Two types of medication for hives are being tested to determine if there is a difference in the percentage of adult patient reactions. Twenty out of a random sample of 200 adults given medication A still had hives 30 mins after taking the medication. Twelve out of another random sample of 200 adults given medication B still had hives 30 mins after taking the medication. Test at a 1% level of significance.

2-Proportions Z Test:

H_0: $p_1 = p_2$

H_1: $p_1 \neq p_2$

The research consists of testing whether there is a difference or not between the two medications. That is the reason why we choose the symbols to be equal and not equal to (Is there a difference or not?). This is a two-tailed test with significance level of 0.01; the Z critical value is 2.575.

Test Statistic:

First, calculate the two samples proportions using the formula:

$$\hat{p} = \frac{x}{n}$$

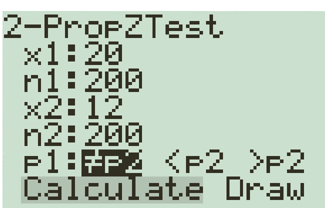
And p-bar as follows:

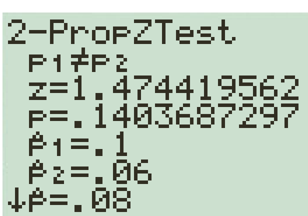
$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$\hat{p_1} = 0.10; \hat{p_2} = 0.06; \bar{p} = 0.08$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.10 - 0.06}{\sqrt{0.08(0.92)\left(\frac{1}{200} + \frac{1}{200}\right)}} = 1.47$$

TI Calculator:

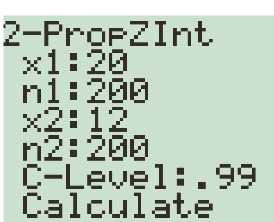


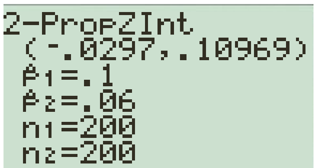


Test Stat is z=1.47; p-value = 0.1404

The 1-α conf. interval: 1-(0.01)=0.99

Conf Interval, TI Calculator:





Notice that the test statistic is less than the critical value, and the p-value is greater than alpha; accordingly, the corresponding confidence interval includes zero.

Conclusions: Fail to reject the Null hypothesis. There is no sufficient evidence to determine that indeed there is a difference in the percentage of adult patients' reactions.

VII: T Test with raw data:

College-aged adults need at least 7 hours of sleep each night to stay healthy. Sleep deprivation can lead to decreased immune system function, lack of concentration, and poor memory. In the data set a simple random sample of 27 college students reports the number of hours of sleep they had last night.

a) What is a 90% confidence interval for the population average sleeping time based on the sample?

b) Is there evidence that the true population mean hours of sleep for college students in the population is different from the 7 hours that are recommended

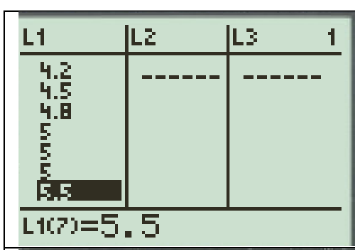
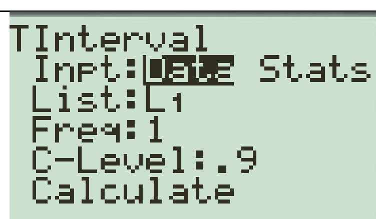
Data:

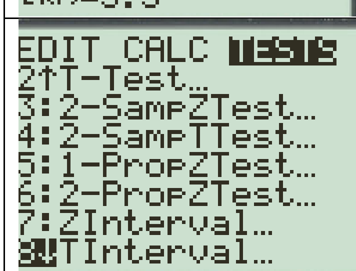
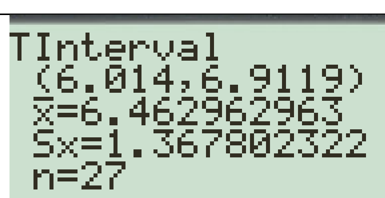
4.2	4.5	4.8	5	5	5	5.5	5.5	5.5
6	6	6	6	6.5	6.5	6.5	6.5	6.5
7.5	7.5	7.5	8	8	8.5	8.5	8.5	9

Taken and modified from:
https://www.stat.purdue.edu/~tqin/system101/method/method_one_t_sas.htm

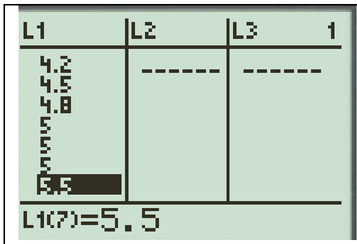
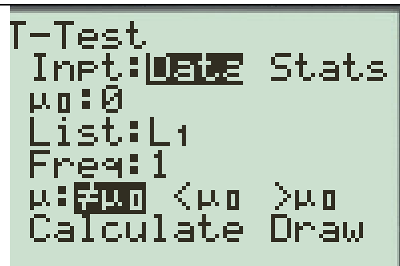
a) 90% Conf Interval. T interval:

Enter data:

<table>
<tr>
<td>
</td>
<td>
</td>
<td>90% Conf Interval

$6.0 < \mu < 6.9$</td>
</tr>
<tr>
<td>
</td>
<td>
</td>
<td>90% conf interval does not include 7; therefore, we are 90 % that the population mean does not include 7 hours of sleep.
In short, the sample data support the claim that the population mean is different from 7.</td>
</tr>
</table>

b) H$_0$: $\mu = 7$
H$_1$: $\mu \neq 7$

		P value = $1.63 \times 10^{-19} \approx 0$ P value very low, lower than any alpha value we may select. Therefore, there is sufficient evidence to support the claim that the population mean is different from 7 hours of sleep time.