MORE HYPOTHESIS TESTING

PART 1
Testing a Claim about a Mean, Standard Deviation Unknown

If we are testing a claim about a mean and the standard deviation is unknown, then we do a ***t-test*** using the ***t distribution***.

If we are testing a claim about a mean and the standard deviation is unknown, then we do a ***t-test*** using the ***t distribution***.

Our main requirement is that either our sample size is large ($n > 30$) or that the parent population be normally distributed.

If we are testing a claim about a mean and the standard deviation is unknown, then we do a ***t-test*** using the ***t distribution***.

Our main requirement is that either our sample size is large ($n > 30$) or that the parent population be normally distributed.

However, be aware that the *t-test* is ***robust***. This means that it tends to work well even when our population is not normally distributed.

Below is the formula for our test statistic.

Test Statistic Formula: $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

Below is the formula for our test statistic.

$$\text{Test Statistic Formula: } t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$\text{degrees of freedom} = n - 1$$

Below is the formula for our test statistic.

$$\text{Test Statistic Formula: } t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$\text{degrees of freedom} = n - 1$$

And as usual, we will be able to do the test on our TI calculator by selecting *T-Test.*

Example 1: A simple random sample of 50 adults is obtained, and each person's red blood cell count is measured. The sample mean is 5.23 and the sample standard deviation is 0.54. At the .01 level of significance, is our result significantly greater than 5.1?

Example 1: A simple random sample of 50 adults is obtained, and each person's red blood cell count is measured. The sample mean is 5.23 and the sample standard deviation is 0.54. At the .01 level of significance, is our result significantly greater than 5.1?

$$\bar{x} = 5.23$$

$$s = 0.54$$

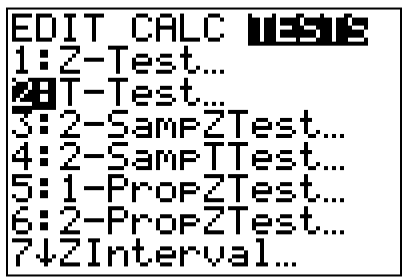
$$n = 50$$

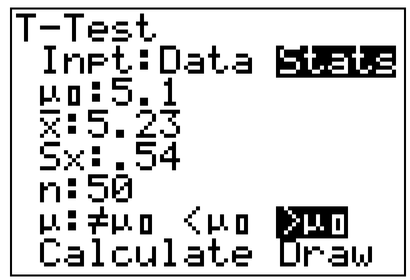
Example 1: A simple random sample of 50 adults is obtained, and each person's red blood cell count is measured. The sample mean is 5.23 and the sample standard deviation is 0.54. At the .01 level of significance, is our result significantly greater than 5.1?

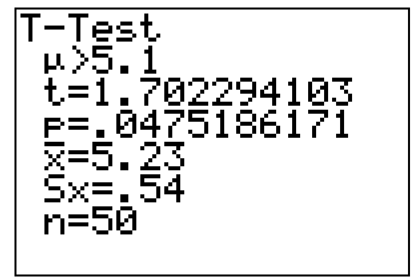
$\bar{x} = 5.23$

$s = 0.54$

$n = 50$







<u>Example 1:</u> A simple random sample of 50 adults is obtained, and each person's red blood cell count is measured. The sample mean is 5.23 and the sample standard deviation is 0.54. At the .01 level of significance, is our result significantly greater than 5.1?

$H_0 : \mu = 5.1 \quad (\leq)$

$H_1 : \mu > 5.1$

```
T-Test
 μ>5.1
 t=1.702294103
 p=.0475186171
 x̄=5.23
 Sx=.54
 n=50
```

Test Statistic Formula: $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

Degrees of Freedom: $df = 49$

Level of Significance: $\alpha = .01$

P-Value: $P = 0.0475$

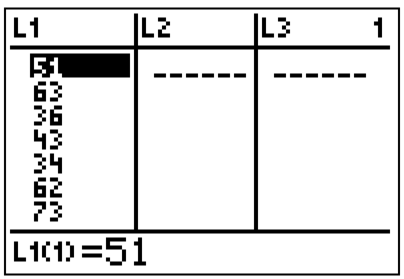
Decision: Fail to reject H_0

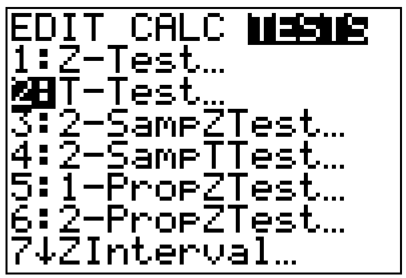
Example 2: The claim is that the *Merriam Webster Collegiate Dictionary* has, on average, 48 defined words per page. Below are the numbers of defined words found in a simple random sample of ten pages. Assuming that the overall distribution of words is normal, at the .05 level of significance determine if the data supports the above claim.

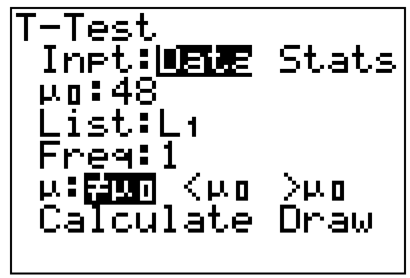
51, 63, 36, 43, 34, 62, 73, 39, 53, 79

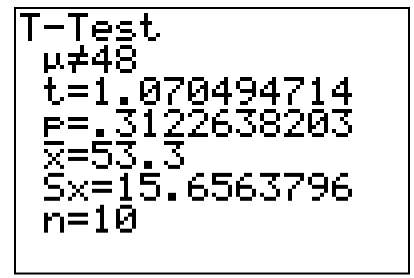
Example 2: The claim is that the *Merriam Webster Collegiate Dictionary* has, on average, 48 defined words per page. Below are the numbers of defined words found in a simple random sample of ten pages. Assuming that the overall distribution of words is normal, at the .05 level of significance determine if the data supports the above claim.

51, 63, 36, 43, 34, 62, 73, 39, 53, 79









Example 2: The claim is that the *Merriam Webster Collegiate Dictionary* has, on average, 48 defined words per page. Below are the numbers of defined words found in a simple random sample of ten pages. Assuming that the overall distribution of words is normal, at the .05 level of significance determine if the data supports the above claim.

51, 63, 36, 43, 34, 62, 73, 39, 53, 79

$H_0 : \mu = 48$

$H_1 : \mu \neq 48$

Test Statistic Formula: $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

Degrees of Freedom: $df = 9$

Level of Significance: $\alpha = .05$

P-Value: $P = 0.3123$

Decision: Fail to reject H_0

PART 2
Testing a Claim about a Proportion

FACT: A binomial distribution with *n* trials and probability of success *p* and probability of failure *q = 1 – p* is approximately normally distributed if both *np* and *nq* are greater than or equal to 5.

$$np \geq 5$$

$$nq \geq 5$$

Example 3: In a Pew Research Center poll of 745 randomly selected adults, 589 said that it is morally wrong to not report all income on tax returns. Use a 0.05 significance level to test the claim that 75% of adults say that it is morally wrong to not report all income on tax returns.

Example 3: In a Pew Research Center poll of 745 randomly selected adults, 589 said that it is morally wrong to not report all income on tax returns. Use a 0.05 significance level to test the claim that 75% of adults say that it is morally wrong to not report all income on tax returns.

$$p = .75$$

$$x = 589$$

$$n = 745$$

$$\hat{p} = .7906$$

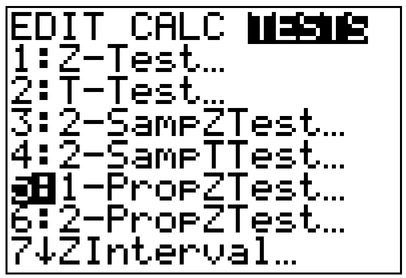
<u>Example 3:</u> In a Pew Research Center poll of 745 randomly selected adults, 589 said that it is morally wrong to not report all income on tax returns. Use a 0.05 significance level to test the claim that 75% of adults say that it is morally wrong to not report all income on tax returns.

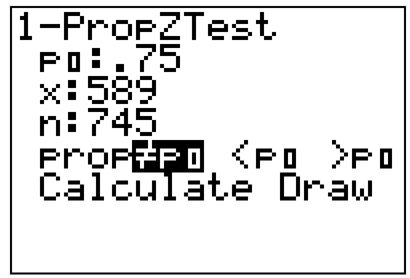
$p = .75$

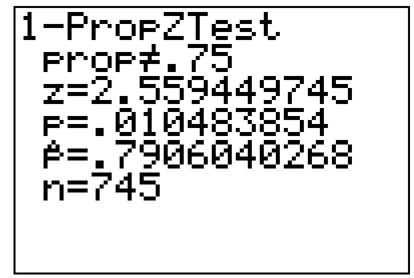
$x = 589$

$n = 745$

$\hat{p} = .7906$







Example 3: In a Pew Research Center poll of 745 randomly selected adults, 589 said that it is morally wrong to not report all income on tax returns. Use a 0.05 significance level to test the claim that 75% of adults say that it is morally wrong to not report all income on tax returns.

$H_0 : p = .75$

$H_1 : p \neq .75$

Test Statistic Formula: $z = \dfrac{\hat{p} - p}{\sqrt{(pq)/n}}$

Level of Significance: $\alpha = .05$

P-Value: $P = 0.0105$

Decision: Reject H_0

Example 4: In a presidential election, 308 out of 611 voters surveyed said that they voted for the candidate who won. If the actual percentage of votes for the winning candidate was 43%, use the .01 level of significance to test the claim that the percentage of voters who said they voted for the winner is not significantly different from 43%.

Example 4: In a presidential election, 308 out of 611 voters surveyed said that they voted for the candidate who won. If the actual percentage of votes for the winning candidate was 43%, use the .01 level of significance to test the claim that the percentage of voters who said they voted for the winner is not significantly different from 43%.

$$p = .43$$

$$x = 308$$

$$n = 611$$

$$\hat{p} = .5041$$

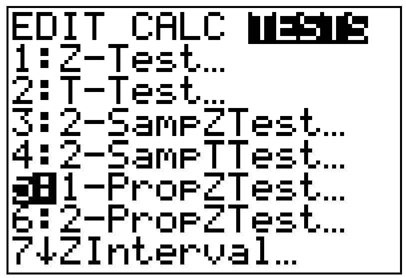
Example 4: In a presidential election, 308 out of 611 voters surveyed said that they voted for the candidate who won. If the actual percentage of votes for the winning candidate was 43%, use the .01 level of significance to test the claim that the percentage of voters who said they voted for the winner is not significantly different from 43%.

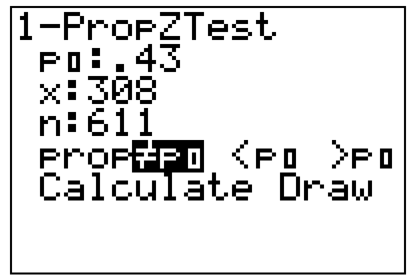
$$p = .43$$

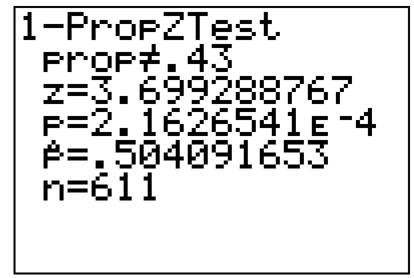
$$x = 308$$

$$n = 611$$

$$\hat{p} = .5041$$







Example 4: In a presidential election, 308 out of 611 voters surveyed said that they voted for the candidate who won. If the actual percentage of votes for the winning candidate was 43%, use the .01 level of significance to test the claim that the percentage of voters who said they voted for the winner is not significantly different from 43%.

$H_0: p = .43$

$H_1: p \neq .43$

Test Statistic Formula: $z = \dfrac{\hat{p} - p}{\sqrt{(pq)/n}}$

Level of Significance: $\alpha = .01$

P-Value: $P = 0.0002$

Decision: Reject H_0

PART 3
Testing a Claim about Two Means using Independent Samples

CONDITION: The samples are either large, n_1 & n_2 both greater than 30, or they come from populations that are normally distributed. Additionally, we do not assume that the variances of the samples are equal, but we do assume that the samples are independent.

TEST STATISTIC:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

TEST STATISTIC:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

$$\text{degrees of freedom} = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

Example 5: Below is information from two independent samples on the average number of words in a day spoken by men and women. Test at the .05 level of significance the claim that men and women speak, on average, the same number of words each day.

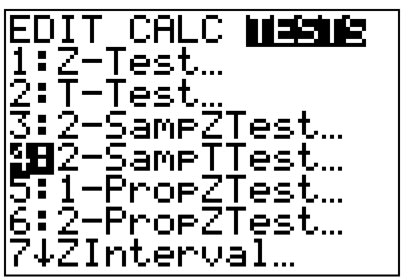
Men: $\bar{x}_1 = 15668.5 \quad s_1 = 8632.5 \quad n_1 = 186$

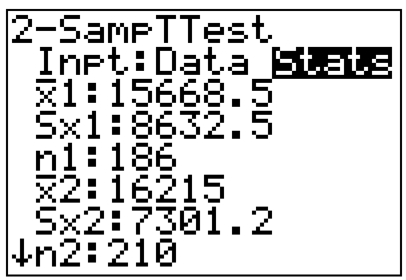
Women: $\bar{x}_2 = 16215.0 \quad s_2 = 7301.2 \quad n_2 = 210$

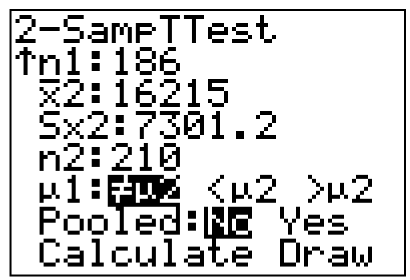
Example 5: Below is information from two independent samples on the average number of words in a day spoken by men and women. Test at the .05 level of significance the claim that men and women speak, on average, the same number of words each day.

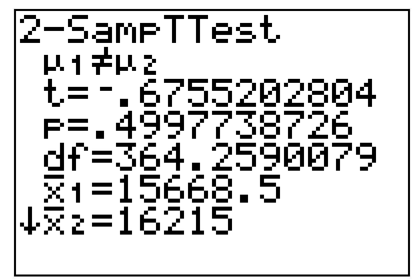
Men: $\bar{x}_1 = 15668.5 \quad s_1 = 8632.5 \quad n_1 = 186$

Women: $\bar{x}_2 = 16215.0 \quad s_2 = 7301.2 \quad n_2 = 210$









Example 5: Below is information from two independent samples on the average number of words in a day spoken by men and women. Test at the .05 level of significance the claim that men and women speak, on average, the same number of words each day.

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

Test Statistic Formula: $t = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

Level of Significance: $\alpha = .05$

P-Value: $P = 0.4997738726 \approx 0.5000$

Decision: Fail to reject H_0

Example 6: A group of supermodels has an average height of 70 inches with a standard deviation of 1.5 inches. If a sample of heights of 40 ordinary women has a mean of 63.2 inches and a standard deviation of 2.7 inches, use a 0.01 significance level to test the claim that the mean height of the supermodels is greater than that of ordinary women. Assume all samples come from normal distributions.

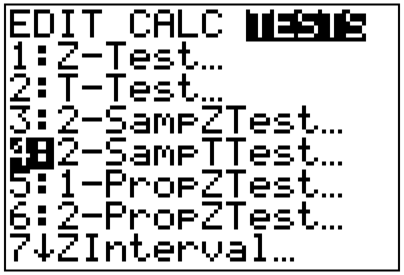
Super: $\bar{x}_1 = 70$ $s_1 = 1.5$ $n_1 = 9$

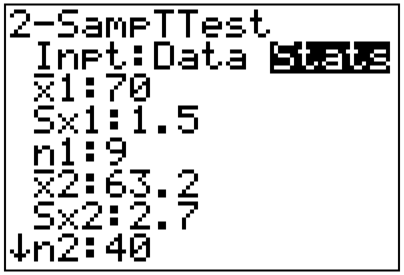
Ordinary: $\bar{x}_2 = 63.2$ $s_2 = 2.7$ $n_2 = 40$

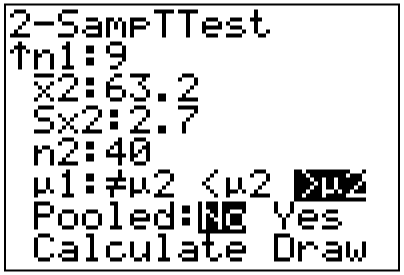
Example 6: A group of supermodels has an average height of 70 inches with a standard deviation of 1.5 inches. If a sample of heights of 40 ordinary women has a mean of 63.2 inches and a standard deviation of 2.7 inches, use a 0.01 significance level to test the claim that the mean height of the supermodels is greater than that of ordinary women. Assume all samples come from normal distributions.

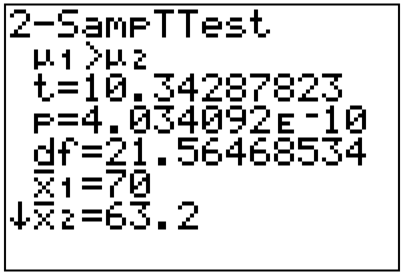
Super: $\bar{x}_1 = 70 \quad s_1 = 1.5 \quad n_1 = 9$

Ordinary: $\bar{x}_2 = 63.2 \quad s_2 = 2.7 \quad n_2 = 40$









Example 6: A group of supermodels has an average height of 70 inches with a standard deviation of 1.5 inches. If a sample of heights of 40 ordinary women has a mean of 63.2 inches and a standard deviation of 2.7 inches, use a 0.01 significance level to test the claim that the mean height of the supermodels is greater than that of ordinary women. Assume all samples come from normal distributions.

$H_0 : \mu_1 = \mu_2 \quad (\leq)$

$H_1 : \mu_1 > \mu_2$

Test Statistic Formula: $t = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

Level of Significance: $\alpha = .01$

P-Value: $P = 4.034092 \times 10^{-10} \approx 0^+$

Decision: Reject H_0

Example 6: A group of supermodels has an average height of 70 inches with a standard deviation of 1.5 inches. If a sample of heights of 40 ordinary women has a mean of 63.2 inches and a standard deviation of 2.7 inches, use a 0.01 significance level to test the claim that the mean height of the supermodels is greater than that of ordinary women. Assume all samples come from normal distributions.

$H_0 : \mu_1 = \mu_2 \quad (\leq)$

$H_1 : \mu_1 > \mu_2$

Test Statistic Formula: $t = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

Level of Significance: $\alpha = .01$

P-Value: $P = 4.034092 \times 10^{-10} \approx 0^+$

If the P is low, the null must go!

Decision: Reject H_0