Lecture 8:
Chapter 8

C C Moxley

UAB Mathematics

1 March 16

§8.1 Hypotheses

In this section we formalize hypothesis testing. We test questions/claims like:

§8.1 Hypotheses

In this section we formalize hypothesis testing. We test questions/claims like:

- A state election board is not using a random process for selecting the ordering of candidates on a ballot because the Republican nominee has been listed second for 15 of the last 16 election cycles.

§8.1 Hypotheses

In this section we formalize hypothesis testing. We test questions/claims like:

- A state election board is not using a random process for selecting the ordering of candidates on a ballot because the Republican nominee has been listed second for 15 of the last 16 election cycles.
- The average weight of tennis balls manufactured by Wilson is less than 100 grams.

§8.1 Hypotheses

In this section we formalize hypothesis testing. We test questions/claims like:

- A state election board is not using a random process for selecting the ordering of candidates on a ballot because the Republican nominee has been listed second for 15 of the last 16 election cycles.
- The average weight of tennis balls manufactured by Wilson is less than 100 grams.
- A paper claims that most American consumers know that the Kindle is a e-book reader.

§8.1 Hypotheses

In this section we formalize hypothesis testing. We test questions/claims like:

- A state election board is not using a random process for selecting the ordering of candidates on a ballot because the Republican nominee has been listed second for 15 of the last 16 election cycles.
- The average weight of tennis balls manufactured by Wilson is less than 100 grams.
- A paper claims that most American consumers know that the Kindle is a e-book reader.
- A sample of 103 human body temperatures can be used to test whether or not the mean body temperature for humans is 98.6°F.

§8.2 Basic Hypothesis Testing

Definition (Hypothesis)

A statistical **hypothesis** is a claim or statement about a property of a population.

§8.2 Basic Hypothesis Testing

Definition (Hypothesis)

A statistical **hypothesis** is a claim or statement about a property of a population. And a **hypothesis test** is a procedure for testing a claim about a property of a population.

§8.2 Basic Hypothesis Testing

Definition (Hypothesis)

A statistical **hypothesis** is a claim or statement about a property of a population. And a **hypothesis test** is a procedure for testing a claim about a property of a population.

Example

From the previous, testing that the average weight of tennis balls manufactured by Wilson is less that 100 grams would be equivalent to testing the statement

$$\mu < 100,$$

where μ is the average weight of Wilson tennis balls.

§8.2 Basic Hypothesis Testing

To test a hypothesis, we use the rare event rule which we have discussed previously:

§8.2 Basic Hypothesis Testing

To test a hypothesis, we use the rare event rule which we have discussed previously:

If under the given assumption, the probability of a particular observed event is extremely small, we reject the assumption.

§8.2 Basic Hypothesis Testing

To test a hypothesis, we use the rare event rule which we have discussed previously:

If under the given assumption, the probability of a particular observed event is extremely small, we reject the assumption.

First we must form a hypothesis. Use these general rules:

§8.2 Basic Hypothesis Testing

To test a hypothesis, we use the rare event rule which we have discussed previously:

If under the given assumption, the probability of a particular observed event is extremely small, we reject the assumption.

First we must form a hypothesis. Use these general rules:

- The **null** hypothesis H_0 should always be that a population parameter is equal to some value.
- The **alternative** hypothesis H_1 should either be that the same parameter is not equal to, less than, or greater than the value above.

§8.2 Example

Translate into a null hypothesis and an alternative hypothesis the following claim:

The proportion of students at UAB who have taken a math class is at least 65%.

§8.2 Example

Translate into a null hypothesis and an alternative hypothesis the following claim:

The proportion of students at UAB who have taken a math class is at least 65%.

H_0 is easy.

§8.2 Example

Translate into a null hypothesis and an alternative hypothesis the following claim:

The proportion of students at UAB who have taken a math class is at least 65%.

H_0 is easy. $H_0 : p = 0.65$.

§8.2 Example

Translate into a null hypothesis and an alternative hypothesis the following claim:

The proportion of students at UAB who have taken a math class is at least 65%.

H_0 is easy. $H_0 : p = 0.65$. What should the other hypothesis be?

§8.2 Example

Translate into a null hypothesis and an alternative hypothesis the following claim:

The proportion of students at UAB who have taken a math class is at least 65%.

H_0 is easy. $H_0 : p = 0.65$. What should the other hypothesis be? Well, it can't be that $p \geq 0.65$ because that **includes** the null hypothesis.

§8.2 Example

Translate into a null hypothesis and an alternative hypothesis the following claim:

The proportion of students at UAB who have taken a math class is at least 65%.

H_0 is easy. $H_0 : p = 0.65$. What should the other hypothesis be? Well, it can't be that $p \geq 0.65$ because that **includes** the null hypothesis. The alternative hypothesis, then, is $H_1 : p < 0.65$,

§8.2 Example

Translate into a null hypothesis and an alternative hypothesis the following claim:

The proportion of students at UAB who have taken a math class is at least 65%.

H_0 is easy. $H_0 : p = 0.65$. What should the other hypothesis be? Well, it can't be that $p \geq 0.65$ because that **includes** the null hypothesis. The alternative hypothesis, then, is $H_1 : p < 0.65$, so failing to reject the null hypothesis is equivalent to supporting the claim whereas rejecting the null hypothesis is equivalent to not supporting the claim.

§8.2 Two Tests

Two types of hypothesis tests will be used in this chapter.

§8.2 Two Tests

Two types of hypothesis tests will be used in this chapter.

- Testing using a critical value.

§8.2 Two Tests

Two types of hypothesis tests will be used in this chapter.

- Testing using a critical value. This method is similar to constructing a confidence interval, except we may have one sided intervals now.

§8.2 Two Tests

Two types of hypothesis tests will be used in this chapter.

- Testing using a critical value. This method is similar to constructing a confidence interval, except we may have one sided intervals now.
- Testing using a P-value, which describes the area lying beyond a test statistic in a one- or two-sided manner.

§8.2 Test Statistic

Definition (Test Statistic)

A **test statistic** is the result of converting a sample statistic into a value used to test the null hypothesis. They are the "standardized" statistic because they generally take the form

$$\frac{\text{estimate} - \text{hypothesized value}}{\text{standard error}}.$$

§8.2 Test Statistic

Definition (Test Statistic)

A **test statistic** is the result of converting a sample statistic into a value used to test the null hypothesis. They are the "standardized" statistic because they generally take the form

$$\frac{\text{estimate} - \text{hypothesized value}}{\text{standard error}}.$$

Proportion p	Mean μ with σ	Mean μ w/o σ	Std. Dev. σ
$z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$	$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$	$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$	$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$

§8.2 Test Statistic

Definition (Test Statistic)

A **test statistic** is the result of converting a sample statistic into a value used to test the null hypothesis. They are the "standardized" statistic because they generally take the form

$$\frac{\text{estimate} - \text{hypothesized value}}{\text{standard error}}.$$

Proportion p	Mean μ with σ	Mean μ w/o σ	Std. Dev. σ
$z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$	$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$	$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$	$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$

The significance of a P-test is called α. We reject the null hypothesis if $p \leq \alpha$ and fail to reject it if $p > \alpha$.

§8.2 Test Statistic

Let ω be a test statistic. We calculate the P-value using the following rules:

§8.2 Test Statistic

Let ω be a test statistic. We calculate the P-value using the following rules:

- If the alternative hypothesis is a left-tailed ("less than") statement, then the P-value is the area to the left of the statistic ω using the appropriate distribution.

§8.2 Test Statistic

Let ω be a test statistic. We calculate the P-value using the following rules:

- If the alternative hypothesis is a left-tailed ("less than") statement, then the P-value is the area to the left of the statistic ω using the appropriate distribution.
- If the alternative hypothesis is a right-tailed ("greater than") statement, then the P-value is the area to the right of the statistic ω using the appropriate distribution.

§8.2 Test Statistic

Let ω be a test statistic. We calculate the P-value using the following rules:

- If the alternative hypothesis is a left-tailed ("less than") statement, then the P-value is the area to the left of the statistic ω using the appropriate distribution.
- If the alternative hypothesis is a right-tailed ("greater than") statement, then the P-value is the area to the right of the statistic ω using the appropriate distribution.
- If the alternative hypothesis is a two-tailed ("not equal to") statement, then the P-value is the area outside of the interval $(-\omega, \omega)$ (if ω is positive) or $(\omega, -\omega)$ (if ω is negative). In the case of a χ^2 test, we look at the small tail and double the area in that tail for the P-value.

§8.2 Error Types and Power of a Test

Let's make a claim that a certain proportion is more than 50%.

§8.2 Error Types and Power of a Test

Let's make a claim that a certain proportion is more than 50%.

Definition (Type I Error)

This is the error of rejecting a true null hypothesis.

§8.2 Error Types and Power of a Test

Let's make a claim that a certain proportion is more than 50%.

Definition (Type I Error)

This is the error of rejecting a true null hypothesis.

This would be like concluding that the proportion was not equal to 50% but was more than 50% when in reality the proportion was exactly 50%.

§8.2 Error Types and Power of a Test

Let's make a claim that a certain proportion is more than 50%.

Definition (Type I Error)

This is the error of rejecting a true null hypothesis.

This would be like concluding that the proportion was not equal to 50% but was more than 50% when in reality the proportion was exactly 50%.

Definition (Type II Error)

This is the error of failing to reject a false null hypothesis.

§8.2 Error Types and Power of a Test

Let's make a claim that a certain proportion is more than 50%.

Definition (Type I Error)

This is the error of rejecting a true null hypothesis.

This would be like concluding that the proportion was not equal to 50% but was more than 50% when in reality the proportion was exactly 50%.

Definition (Type II Error)

This is the error of failing to reject a false null hypothesis.

This would be like concluding that a population proportion was 50% when it actually was not.

§8.2 Error Types and Power of a Test

We use α to denote the the probability of making a Type I Error, i.e. it is the probability of rejecting a true null hypothesis.

§8.2 Error Types and Power of a Test

We use α to denote the the probability of making a Type I Error, i.e. it is the probability of rejecting a true null hypothesis.

We use β to denote the probability of making a Type II Error, i.e. it is the probability of failing to reject a false null hypothesis.

§8.2 Error Types and Power of a Test

We use α to denote the the probability of making a Type I Error, i.e. it is the probability of rejecting a true null hypothesis.

We use β to denote the probability of making a Type II Error, i.e. it is the probability of failing to reject a false null hypothesis.

The confidence of a test is $1-\alpha$ (the probability of failing to reject a true null hypothesis), and the power of a test is $1-\beta$ (the probability of rejecting a false null hypothesis).

§8.3 Example (Proportion)

A survey showed that, among the 514 human resource managers polled, 93% said that the appearance of a job applicant was important. Use a 10% significance level to test the claim that at least 95% of human resources managers think that a job applicant's appearance is important.

§8.3 Example (Proportion)

A survey showed that, among the 514 human resource managers polled, 93% said that the appearance of a job applicant was important. Use a 10% significance level to test the claim that at least 95% of human resources managers think that a job applicant's appearance is important.

$H_0 : p = 0.95$ and

§8.3 Example (Proportion)

A survey showed that, among the 514 human resource managers polled, 93% said that the appearance of a job applicant was important. Use a 10% significance level to test the claim that at least 95% of human resources managers think that a job applicant's appearance is important.

$H_0 : p = 0.95$ and $H_1 : p < 0.95$.

§8.3 Example (Proportion)

A survey showed that, among the 514 human resource managers polled, 93% said that the appearance of a job applicant was important. Use a 10% significance level to test the claim that at least 95% of human resources managers think that a job applicant's appearance is important.

$H_0 : p = 0.95$ and $H_1 : p < 0.95$.

The test statistic is $z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} = \frac{0.92996-0.95}{\sqrt{\frac{(0.95)(0.05)}{514}}} \approx -2.085$.

§8.3 Example (Proportion)

A survey showed that, among the 514 human resource managers polled, 93% said that the appearance of a job applicant was important. Use a 10% significance level to test the claim that at least 95% of human resources managers think that a job applicant's appearance is important.

$H_0 : p = 0.95$ and $H_1 : p < 0.95$.

The test statistic is $z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} = \frac{0.92996-0.95}{\sqrt{\frac{(0.95)(0.05)}{514}}} \approx -2.085$. So we test to see if

$$P(Z < -2.085)) \leq 0.1.$$

§8.3 Example (Proportion)

A survey showed that, among the 514 human resource managers polled, 93% said that the appearance of a job applicant was important. Use a 10% significance level to test the claim that at least 95% of human resources managers think that a job applicant's appearance is important.

$H_0 : p = 0.95$ and $H_1 : p < 0.95$.

The test statistic is $z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} = \frac{0.92996-0.95}{\sqrt{\frac{(0.95)(0.05)}{514}}} \approx -2.085$. So we test to see if

$$P(Z < -2.085)) \leq 0.1.$$

And because $P(Z < -2.085) = 0.0186$, we get that we must reject the null hypothesis.

§8.3 Example (Proportion)

A survey showed that, among the 514 human resource managers polled, 93% said that the appearance of a job applicant was important. Use a 10% significance level to test the claim that at least 95% of human resources managers think that a job applicant's appearance is important.

$H_0 : p = 0.95$ and $H_1 : p < 0.95$.

The test statistic is $z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} = \frac{0.92996-0.95}{\sqrt{\frac{(0.95)(0.05)}{514}}} \approx -2.085$. So we test to see if

$$P(Z < -2.085)) \leq 0.1.$$

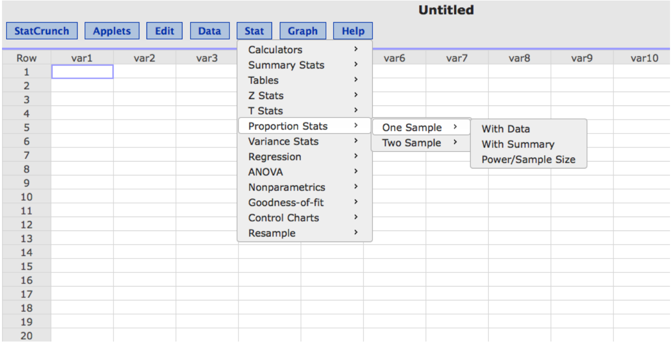
And because $P(Z < -2.085) = 0.0186$, we get that we must reject the null hypothesis. In this case, this means that the data does **not** support our original claim!

§8.3 Using StatCrunch for Hypothesis Testing

You can use StatCrunch to construct a hypothesis test for proportions! You **must** calculate the number of successes, though!

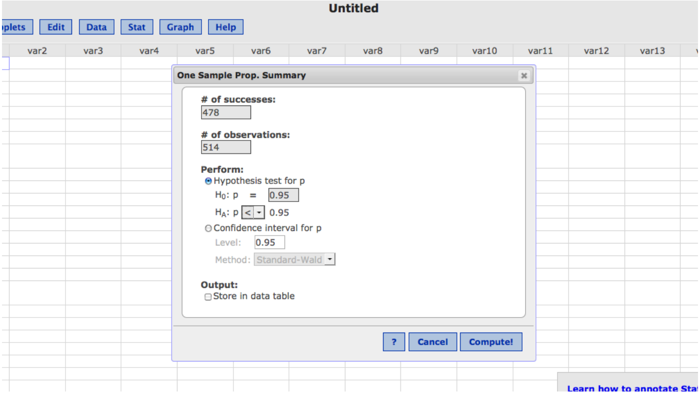
§8.3 Using StatCrunch for Hypothesis Testing

You can use StatCrunch to construct a hypothesis test for proportions! You **must** calculate the number of successes, though!



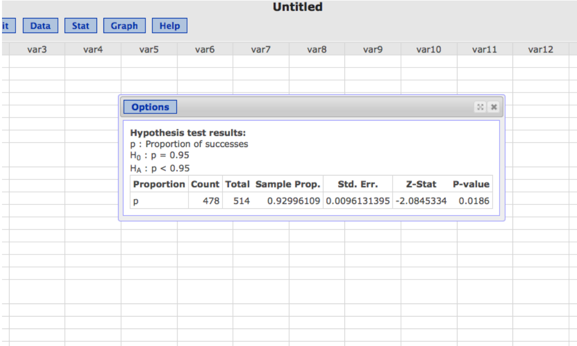
§8.3 Using StatCrunch for Hypothesis Testing

You can use StatCrunch to construct a hypothesis test for proportions! You **must** calculate the number of successes, though!



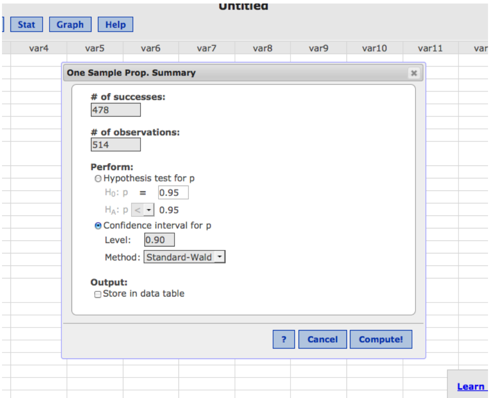
§8.3 Using StatCrunch for Hypothesis Testing

You can use StatCrunch to construct a hypothesis test for proportions! You **must** calculate the number of successes, though!



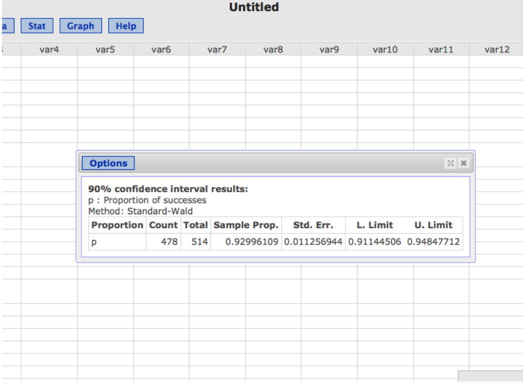
§8.3 Using StatCrunch for Hypothesis Testing

You can also use a confidence interval or critical value to test a hypothesis.



§8.3 Using StatCrunch for Hypothesis Testing

You can also use a confidence interval or critical value to test a hypothesis.



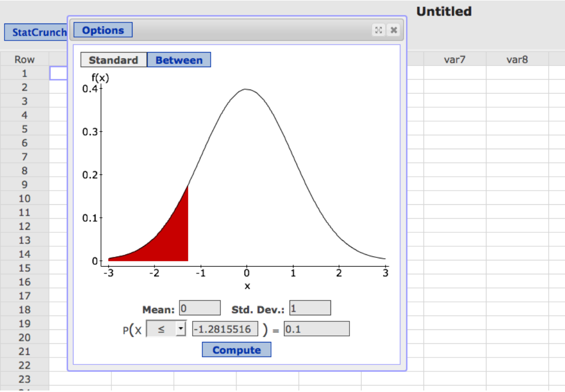
§8.3 Importance of Using an Appropriate Test

Notice the Confidence Interval Method resulting in failing to reject the null hypothesis because it is inherently **two-tailed**. **You must change the significance/confidence level when testing and one-tailed claim using a confidence interval!**

Notice the Confidence Interval Method resulting in failing to reject the null hypothesis because it is inherently **two-tailed**. **You must change the significance/confidence level when testing and one-tailed claim using a confidence interval!** We should have doubled our significance and created an 80% CI!

§8.3 Using StatCrunch for Hypothesis Testing

You can also use a confidence interval or critical value to test a hypothesis.



§8.4 Example (Mean)

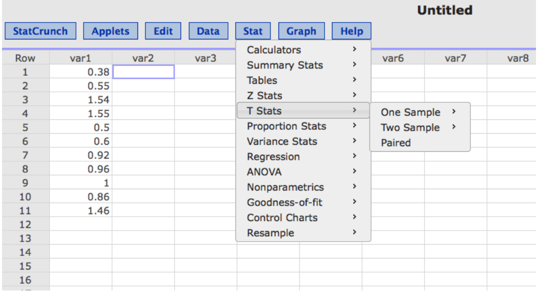
Use a 0.05 significance level to test the claim that the data below comes from a normally distributed population with mean less than 1.

0.38, 0.55, 1.54, 1.55, 0.50, 0.60, 0.92, 0.96, 1.00, 0.86, 1.46.

§8.4 Example (Mean)

Use a 0.05 significance level to test the claim that the data below comes from a normally distributed population with mean less than 1.

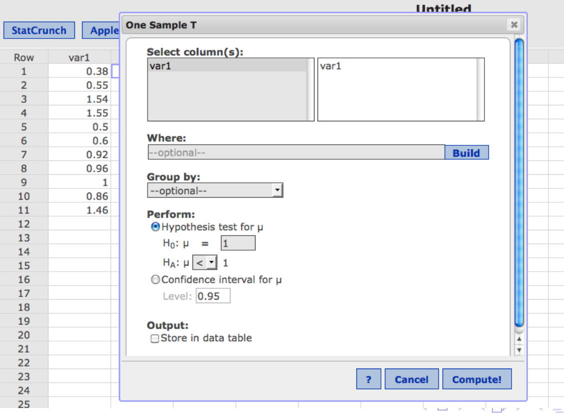
0.38, 0.55, 1.54, 1.55, 0.50, 0.60, 0.92, 0.96, 1.00, 0.86, 1.46.



§8.4 Example (Mean)

Use a 0.05 significance level to test the claim that the data below comes from a normally distributed population with mean less than 1.

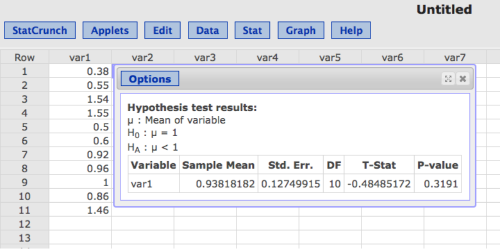
0.38, 0.55, 1.54, 1.55, 0.50, 0.60, 0.92, 0.96, 1.00, 0.86, 1.46.



§8.4 Example (Mean)

Use a 0.05 significance level to test the claim that the data below comes from a normally distributed population with mean less than 1.

0.38, 0.55, 1.54, 1.55, 0.50, 0.60, 0.92, 0.96, 1.00, 0.86, 1.46.



§8.4 Example (Mean)

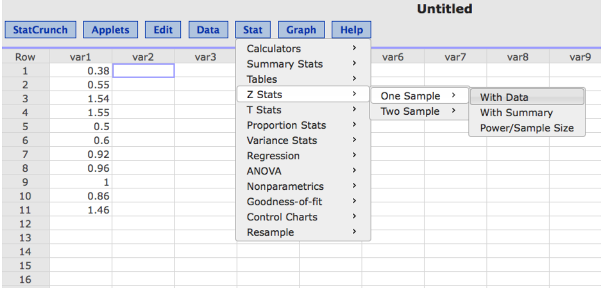
Use a 0.05 significance level to test the claim that the data below comes from a normally distributed population with mean less than 1 and standard deviation 0.03.

0.38, 0.55, 1.54, 1.55, 0.50, 0.60, 0.92, 0.96, 1.00, 0.86, 1.46.

§8.4 Example (Mean)

Use a 0.05 significance level to test the claim that the data below comes from a normally distributed population with mean less than 1 and standard deviation 0.03.

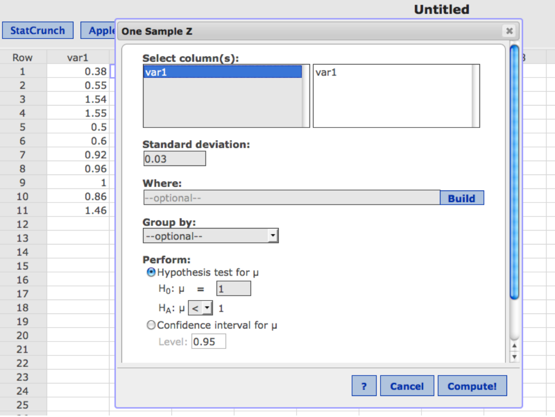
0.38, 0.55, 1.54, 1.55, 0.50, 0.60, 0.92, 0.96, 1.00, 0.86, 1.46.



§8.4 Example (Mean)

Use a 0.05 significance level to test the claim that the data below comes from a normally distributed population with mean less than 1 and standard deviation 0.03.

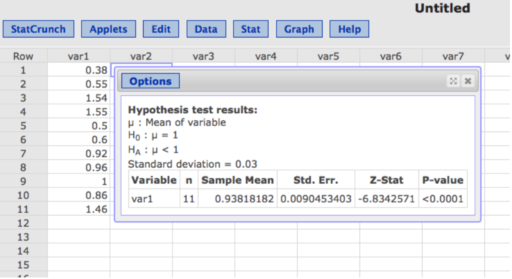
0.38, 0.55, 1.54, 1.55, 0.50, 0.60, 0.92, 0.96, 1.00, 0.86, 1.46.



§8.4 Example (Mean)

Use a 0.05 significance level to test the claim that the data below comes from a normally distributed population with mean less than 1 and standard deviation 0.03.

0.38, 0.55, 1.54, 1.55, 0.50, 0.60, 0.92, 0.96, 1.00, 0.86, 1.46.



§8.4 Example (Standard Deviation)

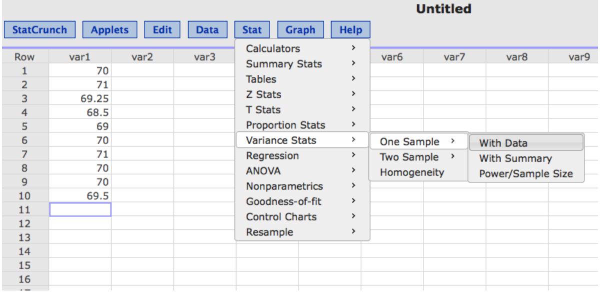
Use a 0.01 significance level to test the claim that the data below comes from a normally distributed population with standard deviation equal to 2.6.

70, 71, 69.25, 68.5, 69, 70, 71, 70, 70, 69.5

§8.4 Example (Standard Deviation)

Use a 0.01 significance level to test the claim that the data below comes from a normally distributed population with standard deviation equal to 2.6.

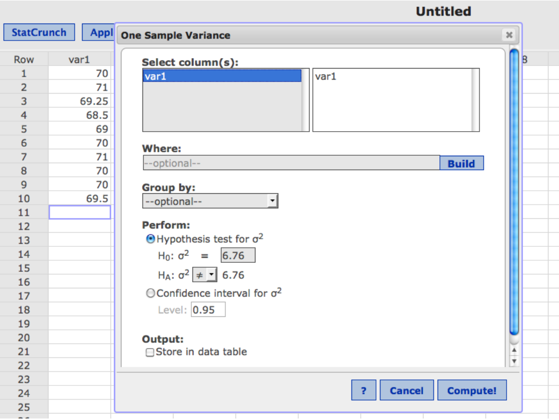
70, 71, 69.25, 68.5, 69, 70, 71, 70, 70, 69.5



§8.4 Example (Standard Deviation)

Use a 0.01 significance level to test the claim that the data below comes from a normally distributed population with standard deviation equal to 2.6.

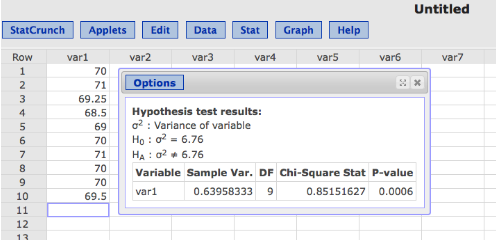
70, 71, 69.25, 68.5, 69, 70, 71, 70, 70, 69.5



§8.4 Example (Standard Deviation)

Use a 0.01 significance level to test the claim that the data below comes from a normally distributed population with standard deviation equal to 2.6.

70, 71, 69.25, 68.5, 69, 70, 71, 70, 70, 69.5



§8 Mastery

What would be a Type I error if the claim was that the proportion of people who write with their left hand is equal to 0.1? What would be a Type 2 error?

§8 Mastery

The claim is that women have heights with $\sigma = 5$cm. After the test is conducted, it's found that the P-value was 0.0055. If we're conducting a test with 99% confidence, do we support this claim or not?

We want to test the claim that at least 98% of Cheez-Its have at least 1.5mg of salt on them. In a sample of 120 Cheez-It crackers, we find that 118 have at least 1.5mg of salt on them. Do we support our claim or not?

§8 Mastery

The brain volumes of cows are given below. We want to test the claim that the population of cow brain volumes has mean equal to 1100 square centimeters. Assume brain volumes of cows are normally distributed.

$$963, 1027, 1272, 1079, 1070, 1173, 1067, 1347, 1100, 1204$$

The claim is that for nicotine amounts in a certain brand of cigarettes, $\mu > 20.0$mg. A test of 200 cigarettes showed the mean to be 20.1mg. The standard deviation for nicotine in this brand of cigarettes is 0.9mg. Can we support our claim with a 10% significance?

§8 Mastery

A simple random sample of 40 men results in a standard deviation of 10.3 heartbeats per minute. Men's heartbeats are normally distributed. We wish to test the claim that men's heartbeats have a standard deviation of 10 heartbeats per minute with a significance of 0.05.