Lecture 8: Chapter 8

C C Moxley

UAB Mathematics

17 October 2016

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.

- 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 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 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.

Definition (Hypothesis)

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

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.

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.

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

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.

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:

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. It is the opposite of the null hypothesis as viewed through the lens of the claim.

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%.

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.

$\S 8.2 \; \mathsf{Example}$

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.

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?

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.

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 $H_1: p<0.65$,

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 rejecting the null hypothesis is equivalent to supporting the claim whereas failing to reject the null hypothesis is equivalent to not supporting the claim.

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

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

■ Testing using a critical value.

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.

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.

Definition (Test Statistic)

A **test statistic** is the result of converting a sample statistic into a value used to test the null hypothesis.

Definition (Test Statistic)

A **test statistic** is the result of converting a sample statistic into a value used to test the null hypothesis.

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

Definition (Test Statistic)

A **test statistic** is the result of converting a sample statistic into a value used to test the null hypothesis.

Proportion p	Mean μ with σ	• •	
$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$	$z = rac{ar{x} - \mu}{rac{\sigma}{\sqrt{n}}}$	$t=rac{ar{x}-\mu}{rac{s}{\sqrt{n}}}$	$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$

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

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

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.

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.

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.

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

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.

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%.

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.

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 the proportion was not greater than 50% when that was actually the case.

§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 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).

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.

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.

 $H_0: p = 0.95 \text{ and}$

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.

 $H_0: p = 0.95 \text{ and } H_1: p < 0.95.$

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.

$$H_0: p = 0.95 \text{ and } H_1: p < 0.95.$$

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

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.

$$H_0: p = 0.95 \text{ and } H_1: p < 0.95.$$

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

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

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.

 $H_0: p = 0.95 \text{ and } H_1: p < 0.95.$

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

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

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



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.

 $H_0: p = 0.95 \text{ and } H_1: p < 0.95.$

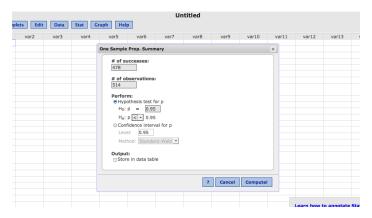
The test statistic is $z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.92996 - 0.95}{\sqrt{\frac{(0.95)(0.05)}{514}}} \approx -2.085$. So we test to see if

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

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!

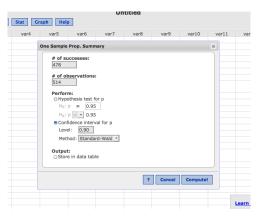


						, ,	Jntitled	i		
StatCru	nch App	lets Edit	Data	Stat Graph	Help					
Row 1 2 3	var1	var2	var3	Calculators Summary Stats Tables Z Stats	> >	var6	var7	var8	var9	var10
4				T Stats	>					
5				Proportion State		One Sar		With Data		
6 7 8				Variance Stats Regression ANOVA	> >	Two San	nple >	With Summar Power/Sample		
9				Nonparametrics	>					
11				Goodness-of-fit Control Charts	>					
13				Resample	>					
14										
15										
16										
17										
18 19										
20										



Data	Stat G	raph Help		U	ntitled				
var3	var4	var5	var6	var7	var8	var9	var10	var11	var12
		Options						22 26	
		H _A : p < 0.		Total Samp	le Prop.	Std. Err.	Z-Stat	P-value	
		p	478	514 0.9	2996109 0.0	0096131395	-2.0845334	0.0186	

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



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

			Untitled				
Stat	raph Help						
var4	var5	var6	var7 var8	3 var9	var10	var11	var12
	Options					20 8	
		ence interva n of successes ndard-Wald					
	Proportion	Count Tota	Sample Prop.	Std. Err.	L. Limit	U. Limit	
	р	478 514	0.92996109	0.011256944	0.91144506	0.94847712	

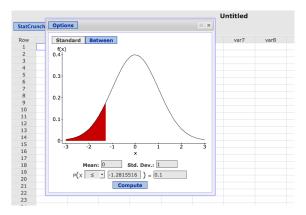
§8.3 Importance of Using an Appropriate Test

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!

§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!** We should have doubled out significance and created an 80% CI!

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

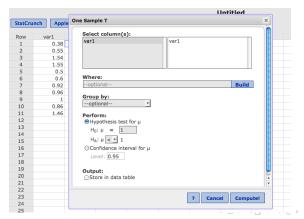


Use a 0.05 significance level to test the claim that the data below comes from a normally distributed population with mean less than 1.

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.

							١	Untitled	
StatCru	nch Appl	ets Edit	Data	Stat	Graph	Help			
Row	var1	var2	var3	Calcu	ators nary Stats	>	var6	var7	var8
2	0.38			Tables		>			
3	1.54			Z Sta	s	>			
4	1.55			T Stat	S	->	One Sar	nple >	
5	0.5				rtion Stats	>	Two Sar	nple >	
6	0.6			Varia	nce Stats	>	Paired		
7	0.92			Regre	ssion				
8	0.96			ANOV	Α	>			
9	1			Nonpa	arametrics	>			
10	0.86			Good	ness-of-fit	>			
11	1.46			Contr	ol Charts	>			
12				Resar	nple	>			
14									
15									
16									

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.



$\S 8.4$ 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.

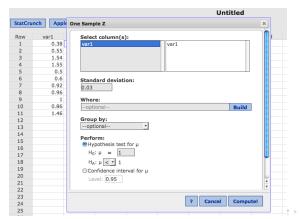
StatCru	nch Appl	ets Edit	Data Sta	t Graph		Help	Untitled			
Row	var1	var2	var3	var4	ar5	var6	var7			
1	0.38		1				50 96			
2	0.55	Options					20 36			
3	1.54									
4	1.55		Hypothesis test results:							
5	0.5	μ : Mean C								
6	0.6	H _Δ : μ < 1								
7	0.92									
8	0.96	Variable	Sample Mean	Std. Err.	DF	T-Stat	P-value			
9	1	var1	0.93818182	0.12749915	10	-0.48485172	0.3191			
10	0.86									
11	1.46									
12										
13										

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.

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.

StatCrui	nch Apple	ts Edit	Data	Stat Graph	Help	U	ntitle	d	
Row 1 2	var1 0.38 0.55	var3	Calculators Summary Stats Tables	> >	var6	var7	var8	var9	
3	1.54			Z Stats	>	One Sam	ple >	With Data	
4	1.55			T Stats	>	Two Sam	ple >	With Summa	ry
5	0.5			Proportion Stats	> '			Power/Sample	e Size
6	0.6			Variance Stats	>				_
7	0.92			Regression	>				
8	0.96			ANOVA	>				
9	1			Nonparametrics	>				
10	0.86			Goodness-of-fit	,				
11	1.46			Control Charts	,				
12					,				
13				Resample	,				
14									
15									
16									

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.



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.

StatCru	ınch App	ets Edit		Data	Stat	Graph He	_	Untitled		
Row	var1	var2	٧	ar3	var4	var5	var6	var7		
1	0.38									
2	0.55	Options						00 16		
3	1.54									
4	1.55		Hypothesis test results:							
5	0.5	μ : Mean o	f va	iriable						
6	0.6	$H_0: \mu = 1$								
7	0.92	H _A : μ < 1								
8	0.96	Standard of	levi	ation =	0.03					
9	1	Variable	n	Samp	le Mean	Std. Err.	Z-Stat	P-value		
10	0.86	var1	11	0.9	3818182	0.0090453403	-6.8342571	<0.0001		
11	1.46									
12										
13										
14										
15										
**										

$(\S 8.4 \; \mathsf{Example} \; (\mathsf{Standard} \; \mathsf{Deviation}))$

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.

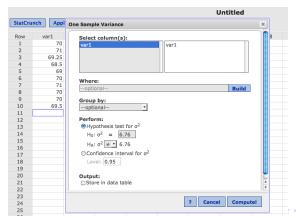
§8.4 Example (Standard Deviation)

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.

				1	Untitle	d			
StatCrun	Apple	ts Edit	Data	Stat Graph	Help	J			
Row	var1 70	var2	var3	Calculators Summary Stats	>	var6	var7	var8	var
2	71			Tables	>				
3	69.25			Z Stats	>				
4	68.5			T Stats	>				
5	69			Proportion Stats	>				
6	70			Variance Stats	>	One Sar	nple >	With Data	
7	71			Regression	>		nple >	With Summary	
8	70			ANOVA	>	Homoge		Power/Sample	
9	70			Nonparametrics	>	Homoge	incity	rowel/ Sample	Size
10	69.5			Goodness-of-fit	,				
11				Control Charts	,				
12					,				
13				Resample					
14									
15									
16									

§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.



$(\S 8.4 \; \mathsf{Example} \; (\mathsf{Standard} \; \mathsf{Deviation}))$

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.

StatCru	nch Appl	ets Edit	Data	at	Graph Help		Untitled			
Row	var1	var2	var3	var4	var5	var6	var7			
1	70									
2	71	Options					×			
3	69.25									
4	68.5		Hypothesis test results:							
5	69		ce of variable							
6	70	$H_0: \sigma^2 = 0$	5.76							
7	71	H _A : σ ² ≠ i	5.76							
8	70	Variable	Sample Var.	DF	Chi-Square Stat	P-value				
9	70									
10	69.5	var1	0.63958333	9	0.85151627	0.0006				
11										
12										
13										

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?

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?

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

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.