

Hypothesis Testing

LAST WEEK

Confidence intervals use sample data to specify a range of values within which we are confident that the population parameter falls

Example: "With 95% certainty we conclude that the population proportion of people who own a car is between 0.401 and 0.420"

Example: "With 99% certainty we conclude that the mean income in the population is somewhere between \$31,200 and \$32,000"

Hypothesis Testing

THIS WEEK

Hypothesis testing (or significance testing) uses sample data to test particular claims about the value of a population parameter

Example: "Do our sample data support the assertion that more than 41% of people in the population own cars?"

Example: "Do our sample data support the assertion that the mean income in the population is greater than \$31,900?"

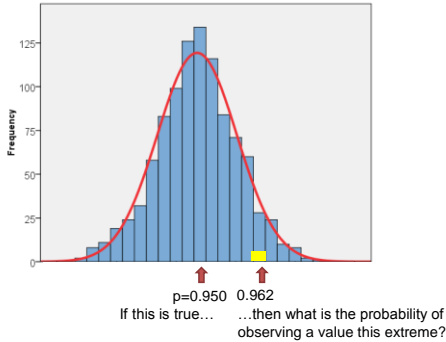
Hypothesis Testing

In 2012 the U.S. National Transportation Safety Board set a 5-year goal of having more than 95% of all American drivers use their seatbelts

To see whether they are on target for meeting that goal, they randomly sampled 1,000 American drivers in 2014

They found that 962—or 96.2%—of the 1,000 drivers they sampled use their seatbelts

Is this evidence sufficient to confidently conclude that more than 95% of *all* American drivers use their seatbelts?



Hypothesis Testing

We will consider hypothesis tests for...

- ...single population proportions (p)
- ...single population means (μ_v)
- ...differences in population proportions ($p_1 - p_2$)
- ...differences in population means ($\mu_1 - \mu_2$)

As with confidence intervals, the logic is the same for each type of hypothesis test

Hypothesis Testing

Hypothesis tests involve two competing (mutually exclusive and exhaustive) hypotheses; what these hypotheses are depends on the question at hand

Examples:

- Hyp. 1:* More than 95% of all U.S. drivers use their seat belts
- Hyp. 2:* 95% or fewer of all U.S. drivers use their seat belts

- Hyp. 1:* There are 12.1 penguins per square mile
- Hyp. 2:* The number of penguins per square mile is not 12.1

Hypothesis Testing

These two hypotheses are typically called...

...the "null" hypothesis (H_0) and

...the "alternative" hypothesis (H_1 or H_A ... H_1 in this class)

The **null hypothesis** often says that nothing is happening or that something is not true

The **alternative hypothesis** usually says that something is happening or that something is true

Hypothesis Testing

Null Hypothesis	Alternative Hypothesis
pH of Ocean Water is 7.8 $\mu = 7.8$	pH of Ocean Water is not 7.8 $\mu \neq 7.8$
The Average Skink Weighs an Ounce or Less $\mu \leq 1 \text{ oz}$	The Average Skink Weighs More than an Ounce $\mu > 1 \text{ oz}$
10% of People Carry Guns $p = 0.10$	The % Who Carry Guns is Not 10% $p \neq 0.10$
Coke and Pepsi Taste the Same $T_p = T_c$ (or, $T_p - T_c = 0$)	Coke and Pepsi Taste Different $T_p \neq T_c$ (or, $T_p - T_c \neq 0$)

Hypothesis Testing

Depending on the substantive situation, we might distinguish between **one-sided** and **two-sided** tests

If we are asking whether a population value is higher than a certain threshold (say, 0), then we use a one-sided test

Example: $H_0: \mu \leq 0$ $H_1: \mu > 0$

If we are asking whether a population value is lower than a certain threshold (say, 0), then we also use a one-sided test

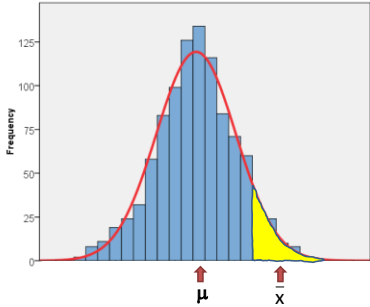
Example: $H_0: \mu \geq 0$ $H_1: \mu < 0$

If we are asking whether a population value is simply different than a certain threshold (say, 0) then we use a two-sided test

Example: $H_0: \mu = 0$ $H_1: \mu \neq 0$

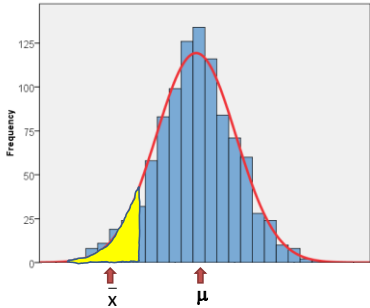
One-sided (upper tail) hypothesis test

Assuming that μ is true, what is the probability of observing one sample mean that far **above** μ ?



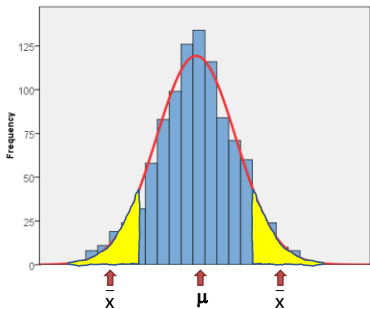
One-sided (lower tail) hypothesis test

Assuming that μ is true, what is the probability of observing one sample mean that far **below** μ ?



Two-sided hypothesis test

Assuming that μ is true, what is the probability of observing one sample mean that far **above or below** μ ?



Hypothesis Testing

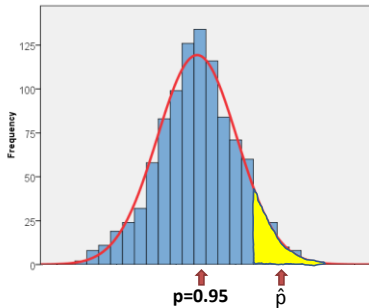
Do more than 95% of all American drivers use seat belts?

We will answer “yes” if we conclude that more than 95% use their seat belts. So, we employ a one-sided hypothesis test:

$$H_0: p \leq 0.95 \quad H_a: p > 0.95$$

One-sided (upper tail) hypothesis test

Assuming that $p=0.95$ or less, what is the probability of observing one sample mean that far **above** p ?



Hypothesis Testing

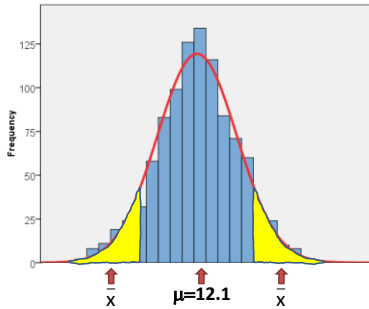
Are there 12.1 penguins per square mile in Antarctica?

We will answer “no” if we conclude that the mean number of penguins per square mile is greater than 12.1 or less than 12.1. So, we employ a two-sided hypothesis test:

$$H_0: \mu = 12.1 \quad H_a: \mu \neq 12.1$$

Two-sided hypothesis test

Assuming that μ is 12.1, what is the probability of observing one sample mean that far **above or below** 12.1?



Worksheet

Does the average dog weigh less than 20 pounds ?

H_0 : H_a :

Do Americans typically have 3 credit cards?

H_0 : H_a :

Are more than 10% of women in abusive relationships ?

H_0 : H_a :

Is it true that 5% of kids are in street gangs ?

H_0 : H_a :

Hypothesis Testing

1. State the null (H_0) and alternative (H_1) hypotheses
2. Check that the sample data conform to basic assumptions; if they do not, then do not go any further
3. Choose an α probability level ... that is, a probability associated with incorrectly rejecting the null hypothesis
4. Determine the "critical value" ... that is, how large the test statistic must be in order to reject the null hypothesis at the given α level
5. Calculate the test statistic
6. Compare the test statistic to the critical value
 - If the test statistic is as large as the critical value, then reject H_0 (with probability of α of doing so even though H_0 should not actually be rejected)
 - If the test statistic is less than or equal to the critical value, then do not reject H_0 (with probability of β of doing so even though H_0 should be rejected)

Example 1

In 2007 the U.S. National Transportation Safety Board set a 5-year goal of having more than 95% of all American drivers use their seatbelts

To see whether they are on target for meeting that goal, they randomly sampled 1,000 American drivers in 2012

They found that 962—or 96.2%—of the 1,000 drivers they sampled use their seatbelts

Is this evidence sufficient to confidently conclude that more than 95% of all American drivers use their seatbelts?

Example 1

1. State the null (H_0) and alternative (H_1) hypotheses

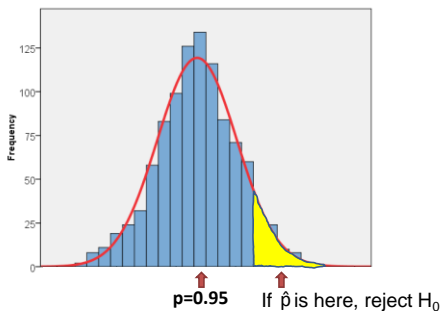
Because we are asking whether the proportion exceeds a particular value (instead of just being different from it), this is a one-sided hypothesis test

$$H_0: p \leq 0.95$$

$$H_1: p > 0.95$$

One-sided (upper tail) hypothesis test

Assuming that $p=0.95$ or less, what is the probability of observing one sample proportion that far **above** p ?



Example 1

2. Check that the sample data conform to basic assumptions; if they do not, then do not go any further

The sample must have been randomly selected from the population ... (True)

np_0 and $n(1-p_0)$ must both be at least 10 (where p_0 is the proportion we assume to be correct based on the null hypothesis)

$$np_0 = 1,000 \times 0.95 = 950 \text{ and } n(1-p_0) = 1,000 \times 0.05 = 50$$

(True)

Example 1

3. Choose an α probability level ... that is, a probability associated with incorrectly rejecting the null hypothesis

The precise α level we select is arbitrary, and varies by discipline (and sometimes by journal or even by author)

0.05 is most common; 0.01 is common, too

Let's go with $\alpha=0.05$ in this example

Example 1

4. Determine the "critical value" ... that is, how large the test statistic must be in order to reject the null hypothesis at the given α level

How large must the critical value Z^* be in order to have the area under the standard normal curve that exceeds Z^* be less than α ?

It can then be helpful to re-write the hypotheses in terms of the Z-score(s)

Example 1

5. Calculate the test statistic

All test statistics for hypothesis testing...

...subtract the population value that is assumed to be true under the null from the observed (sample) value

...divide that figure by the standard deviation of the sampling distribution of the statistic in question

For proportions:

$$Z = \frac{\text{sample estimate} - \text{null value}}{\text{null standard deviation}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Example 1

In our example:

$H_0: p \leq 0.95$

$H_1: p > 0.95$

$n = 1,000$

$$\hat{p} = \frac{962}{1,000} = 0.962$$

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{N}}} = \frac{0.962 - 0.950}{\sqrt{\frac{0.950(1-0.950)}{1,000}}} = 1.74$$

Example 1

6. Compare the test statistic to the critical value

- If the test statistic is larger than the critical value, then reject H_0 (with probability of α of doing so even though H_0 should not actually be rejected)
- If the test statistic is less than or equal to the critical value, then do not reject H_0 (with probability of β of doing so even though H_0 should be rejected)

In our example, our critical value Z^* is +1.645

We obtained a test statistic of 1.74

$H_0: p \leq 0.95$... fail to reject H_0 if test statistic $Z \leq 1.645$

$H_1: p > 0.95$... reject H_0 if test statistic $Z > 1.645$

Conclusion? **Reject H_0**

Example 2

A veterinarian claims that 6% of cats have FIDS (Feline Immune Deficiency Syndrome)

To evaluate this claim, researchers randomly sampled 320 cats
They found that 26—or 8.1%—of the 320 cats have FIDS

Is this evidence sufficient to confidently conclude that the population proportion of cats who have FIDS is different from 0.06?

Example 2

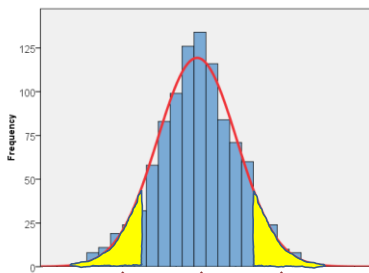
1. State the null (H_0) and alternative (H_1) hypotheses

Because we are asking whether the proportion is different from a particular value (regardless of whether it is higher or lower than that value) this is a two-sided hypothesis test

$$H_0: p = 0.06$$
$$H_1: p \neq 0.06$$

One-sided (upper tail) hypothesis test

Assuming that $p=0.95$ or less, what is the probability of observing one sample proportion that far **above** p ?



If \hat{p} is here, reject H_0 $p=0.06$ If \hat{p} is here, reject H_0

Example 2

2. Check that the sample data conform to basic assumptions; if they do not, then do not go any further

The sample must have been randomly selected from the population ... (True)

np_0 and $n(1-p_0)$ must both be at least 10 (where p_0 is the proportion we assume to be correct based on the null hypothesis)

$$np_0 = 320 \times 0.06 = 19.2 \text{ and } n(1-p_0) = 320 \times 0.94 = 300.8$$

(True)

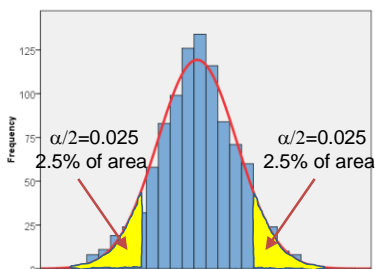
Example 2

3. Choose an α probability level ... that is, a probability associated with incorrectly rejecting the null hypothesis

Let's go with $\alpha=0.05$ in this example

One-sided (upper tail) hypothesis test

Assuming that $p=0.95$ or less, what is the probability of observing one sample proportion that far above p ?



If $Z < -Z^*$, reject H_0 $-Z^*$ $p=0.06$ Z^* If $Z > Z^*$, reject H_0

Example 2

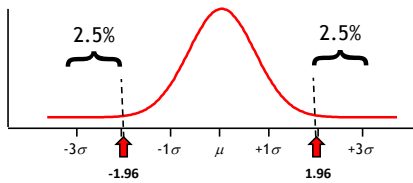
4. Determine the “critical value” ... that is, how large the test statistic must be in order to reject the null hypothesis at the given α level

What critical value Z^* must we choose so that $5\%/2=2.5\%$ of the area under the normal curve lies above $+Z^*$ and so that $5\%/2=2.5\%$ of the area under the normal curve lies below $-Z^*$?

1.96

Example 2

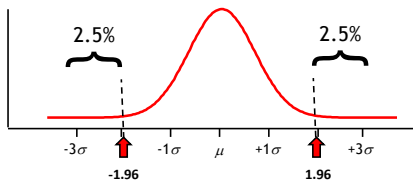
With $\alpha=0.05$ and a two-sided test ...



... a Z-score that is 1.96 in absolute value has 2.5% of the area under the curve beyond it in both directions

Example 2

With $\alpha=0.05$ and a two-sided test ...



... so our critical value Z^* is 1.96

$H_0: p = 0.06$... fail to reject H_0 if $|Z| \leq 1.96$

$H_1: p \neq 0.06$... reject H_0 if $|Z| > 1.96$

Example 2

In our example:

$$H_0: p = 0.06$$

$$H_1: p \neq 0.06$$

$$n = 320$$

$$\hat{p} = \frac{26}{320} = 0.08125$$

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.08125 - 0.06}{\sqrt{\frac{0.06(1-0.06)}{320}}} = 1.601$$

Example 2

6. Compare the test statistic to the critical value

In our example, our critical value Z^* is 1.96

We obtained a test statistic of 1.601

$$H_0: p = 0.06 \dots \text{fail to reject } H_0 \text{ if } |Z| \leq 1.96$$

$$H_1: p \neq 0.06 \dots \text{reject } H_0 \text{ if } |Z| > 1.96$$

Conclusion? **Fail to Reject H_0**

(But what if we had used $\alpha=0.10$?)

Errors in Hypothesis Testing

When we test a hypothesis, we can come to one of two conclusions

Reject the null hypothesis (H_0)

Fail to reject the null hypothesis (H_0)

Regardless of what we decide, we risk making an error

If we reject H_0 , it could be that H_0 is actually true ... such that we should not have rejected it

If we fail to reject H_0 , it could be that H_0 is actually false ... such we should have rejected it

Errors in Hypothesis Testing

A drug company has developed a new drug that they think will reduce rates of depression

The drug company will conclude that its new drug is effective if fewer than 10% of people who take it suffer from depression after taking the drug

To determine whether the new drug is effective the drug company might thus test the hypotheses:

$$H_0: p \geq 0.10$$

$$H_1: p < 0.10$$

Errors in Hypothesis Testing

The drug company risks making one of two errors:

First, the company could conclude that the drug works when in fact it does not ... this is called a Type I error

They would reject the null when in fact the null is true

Consequence: They mistakenly claim that they have an effective drug to combat depression

Second, the company could conclude that the drug does not work when in fact it does ... this is a Type II error

They would fail to reject H_0 when H_0 should be rejected

Consequence: They withhold an effective depression treatment from people who might benefit from it

Errors in Hypothesis Testing

		THE TRUTH IS...	
		H_0 is True	H_0 is False
WE CONCLUDE...	Fail to Reject H_0	✓	✗ Type II Error $P(\text{Type II Error})=\beta$
	Reject H_0	✗ Type I Error $P(\text{Type I Error})=\alpha$	✓

Worksheet

I think that 5% of people in Minneapolis own RVs

To see whether this is true, I randomly sampled 2,310 people in Minneapolis and asked them whether they own an RV

I found that 110—or 4.76%—of the 2,310 people own RVs

Is this evidence sufficient to confidently (use $\alpha=0.01$) conclude that the population proportion of people in Minneapolis who own RVs is different from 0.05?

Worksheet

I think the poverty rate is more than 20%

To see whether this is true, I randomly sampled 1,100 people and determined whether they were below the poverty line

I found that 242—or 22.0%—of the 1,100 people were poor

Is this evidence sufficient to confidently (use $\alpha=0.05$) conclude that the population proportion of people who are poor is more than 0.20?

BREAK

Hypothesis Tests

We will consider four different sorts of hypothesis tests, all of which follow the same logic

- Hypothesis Tests for Proportions**
Use \hat{p} to infer p , the population proportion
- Hypothesis Tests for Means**
Use \bar{Y} to infer μ_y , the population mean of Y
- Hypothesis Tests for Differences in Proportions**
Use $\hat{p}_1 - \hat{p}_2$ to infer the difference between two population proportions, p_1 and p_2
- Hypothesis Tests for Differences in Means**
Use $\bar{Y}_1 - \bar{Y}_2$ to infer the difference between two population means, μ_{Y1} and μ_{Y2}

How to Test a Hypothesis

- State the null (H_0) and alternative (H_1) hypotheses
- Check that the sample data conform to basic assumptions; if they do not, then do not go any further
- Choose an α probability level ... that is, a probability associated with incorrectly rejecting the null hypothesis
- Determine the "critical value" ... that is, how large the test statistic must be in order to reject the null hypothesis at the given α level
- Calculate the test statistic
- Compare the test statistic to the critical value
 - If the test statistic is larger than the critical value, then reject H_0 (with probability of α of doing so even though H_0 should not actually be rejected)
 - If the test statistic is less than or equal to the critical value, then do not reject H_0 (with probability of β of doing so even though H_0 should be rejected)

Hypotheses about a Mean: Example #1

A 2001 census of penguins found that there were 12.1 penguins per square mile in Antarctica

Now researchers are interested in seeing how the size of the population of penguins in Antarctica has changed

They randomly sampled 200 square mile sections in Antarctica, and observed the number of penguins that lived on each section

In their sample they observed a mean of 11.7 penguins per square mile, with $s=2.3$ penguins

Is this evidence sufficient to confidently conclude that the size of the penguin population has changed?

Hypotheses about a Mean: Example #1

1. State the null (H_0) and alternative (H_1) hypotheses

Because we are asking whether the mean is *different* from a particular value (as opposed to being either *higher* or *lower* than a particular value), this is a two-sided hypothesis test

$$H_0: \mu = 12.1$$

$$H_1: \mu \neq 12.1$$

Hypotheses about a Mean: Example #1

2. Check that the sample data conform to basic assumptions; if they do not, then do not go any further

The random sample is small (<30) and symmetrically distributed with no outliers, and the population of measurements is bell-shaped (**Not True**)

--- or ---

The size of the random sample is large (≥ 30), regardless of the shape of the distribution of the measurements in the population (**True**)

Hypotheses about a Mean: Example #1

3. Choose an α probability level ... that is, a probability associated with incorrectly rejecting the null hypothesis

Again, the precise α level we select is arbitrary, and varies by discipline (and sometimes by journal or author)

Let's go with $\alpha=0.01$ in this example

Hypotheses about a Mean: Example #1

4. Determine the "critical value" ... that is, how large the test statistic must be in order to reject the null hypothesis at the given α level

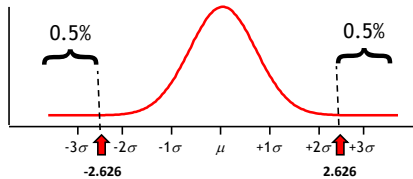
Because we are dealing with sample means, our test statistic will be a t-score (not a Z-score)

What critical value t^* corresponds to having $1\%/2=0.5\%$ of the area under the normal curve above $+t^*$ and 0.5% of the area under the normal curve below $-t^*$?

2.626

Hypotheses about a Mean: Example #1

With $\alpha=0.01$, $N=200$, and a two-sided test ...



... so our critical value t^* is 2.626 (at least using our class's t table)

$H_0: \mu = 12.1$... fail to reject H_0 if $|t| \leq 2.626$

$H_1: \mu \neq 12.1$... reject H_0 if $|t| > 2.626$

Hypotheses about a Mean: Example #1

5. Calculate the test statistic

All test statistics for hypothesis testing...

...subtract the population value that is assumed to be true under the null from the observed (sample) value

...divide that figure by the standard deviation of the sampling distribution of the statistic in question

For single means:

$$t = \frac{\bar{Y} - \mu_y}{s_y / \sqrt{n}}$$

Hypotheses about a Mean: Example #1

In our example:

$$H_0: \mu = 12.1$$

$$H_1: \mu \neq 12.1$$

$$\bar{Y} = 11.7$$

$$s_Y = 2.3$$

$$n = 200$$

$$t = \frac{\bar{Y} - \mu_Y}{s_Y / \sqrt{N}} = \frac{11.7 - 12.1}{2.3 / \sqrt{200}} = -2.459$$

Hypotheses about a Mean: Example #1

6. Compare the test statistic to the critical value

- If the test statistic is larger than the critical value, then reject H_0
(with probability of α of doing so even though H_0 should not actually be rejected)
- If the test statistic is less than or equal to the critical value, then do not reject H_0 (with probability of β of doing so even though H_0 should be rejected)

In our example, our critical value t^* is 2.626

We obtained a test statistic of -2.459

$$H_0: \mu = 12.1 \dots \text{fail to reject } H_0 \text{ if } |t| \leq 2.626$$

$$H_1: \mu \neq 12.1 \dots \text{reject } H_0 \text{ if } |t| > 2.626$$

Conclusion? **Fail to Reject H_0**

(But what if α had been 0.05?)

Worksheet

Does the average undergraduate student bring more than \$40 to class?

Using data from a survey conducted on the first day of my undergraduate research methods class, I found that the mean amount of money that the 75 students had on them was \$43 with a standard deviation of \$78 (Assume for now that students in my class represent a random sample of all undergraduate students)

Use $\alpha = 0.05$

Hypotheses about Differences in Proportions: Example #1

Are men and women equally likely to smoke?

To find out, I will use data from the 1975-1994 General Social Surveys (GSS)

Among men: 2,839 out of 7,014 smoke ($p=0.405$, or 40.5%)

Among women: 2,886 out of 9,359 smoke ($p=0.308$, or 30.8%)

Is this sufficient evidence to confidently conclude that men and women are not equally likely to smoke?

Hypotheses about Differences in Proportions: Example #1

1. State the null (H_0) and alternative (H_1) hypotheses

Because we are asking whether or not the difference in proportions is different from a particular value (zero), this is a two-sided hypothesis test

$$H_0: p_M = p_F \quad (\text{or, } p_M - p_F = 0)$$

$$H_1: p_M \neq p_F \quad (\text{or, } p_M - p_F \neq 0)$$

Hypotheses about Differences in Proportions: Example #1

2. Check that the sample data conform to basic assumptions; if they do not, then do not go any further

The two samples must be independent and representative of the two populations from which they were drawn (True)

As with confidence intervals, $N_1 \hat{p}_1$, $N_1(1-\hat{p}_1)$, $N_2 \hat{p}_2$, and $N_2(1-\hat{p}_2)$ must all exceed 5 (and preferably 10) (True)

Hypotheses about Differences in Proportions: Example #1

3. Choose an α probability level ... that is, a probability associated with incorrectly rejecting the null hypothesis

Let's go with $\alpha=0.05$ in this example

Hypotheses about Differences in Proportions: Example #1

4. Determine the "critical value" ... that is, how large the test statistic must be in order to reject the null hypothesis at the given α level

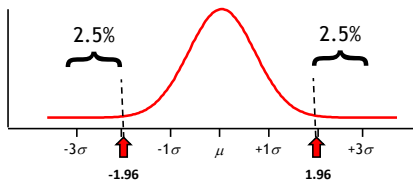
Because we are dealing with sample proportions, our test statistic will be a Z-score

What critical value Z^* corresponds to having 5%/2=2.5% of the area under the normal curve above $+Z^*$ and 2.5% of the area under the normal curve below $-Z^*$?

1.96

Hypotheses about Differences in Proportions: Example #1

With $\alpha=0.05$ and a two-sided test ...



... so our critical value Z^* is 1.96

$H_0: p_M = p_F$... fail to reject H_0 if $|Z| \leq 1.96$

$H_1: p_M \neq p_F$... reject H_0 if $|Z| > 1.96$

Hypotheses about Differences in Proportions: Example #1

5. Calculate the test statistic

All test statistics for hypothesis testing...

...subtract the population value that is assumed to be true under the null from the observed (sample) value

...divide that figure by the standard deviation of the sampling distribution of the statistic in question

For differences in proportions:

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}} \quad \text{where} \quad \hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$$

Hypotheses about Differences in Proportions: Example #1

In our example: $\hat{p}_M = 0.405$ $n_M = 7,014$
 $\hat{p}_F = 0.308$ $n_F = 9,359$

$$\hat{p} = \frac{n_M\hat{p}_M + n_F\hat{p}_F}{n_M + n_F} = \frac{7,014(0.405) + 9,359(0.308)}{7,014 + 9,359} = 0.350$$

$$Z = \frac{\hat{p}_M - \hat{p}_F - 0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_M} + \frac{\hat{p}(1-\hat{p})}{n_F}}} = \frac{0.405 - 0.308 - 0}{\sqrt{\frac{0.350(1-0.350)}{7,014} + \frac{0.350(1-0.350)}{9,359}}} = 12.9$$

Hypotheses about Differences in Proportions: Example #1

6. Compare the test statistic to the critical value

- If the test statistic is larger than the critical value, then reject H_0 (with probability of α of doing so even though H_0 should not actually be rejected)
- If the test statistic is less than or equal to the critical value, then do not reject H_0 (with probability of β of doing so even though H_0 should be rejected)

In our example, our critical value Z^* is 1.96

We obtained a test statistic of 12.9

$H_0: p_M = p_F$... fail to reject H_0 if $|Z| \leq 1.96$

$H_1: p_M \neq p_F$... reject H_0 if $|Z| > 1.96$

Conclusion? **Reject H_0**

Worksheet

Did high school age boys' condom use during last intercourse increase between 2009 and 2014?

2009 Youth Risk Behavior Survey

69% of high school age boys used a condom
n = 2,640

2014 National Survey of Sexual Health and Behavior

79% of high school age boys used a condom
n = 57

Use $\alpha=0.05$

Hypotheses about Differences in Means: Example #1

Someone claims that women have better vocabularies than men

To test this claim we use data from the General Social Survey, which contains a short vocabulary test

In these data:

$\bar{Y}_M=5.93$ $s_M=2.200$ $n_M=8,709$
 $\bar{Y}_F=6.01$ $s_F=2.153$ $n_F=11,530$

Is this evidence sufficient to confidently conclude that (in the population) women have better vocabularies than men?

Hypotheses about Differences in Means: Example #1

1. State the null (H_0) and alternative (H_1) hypotheses

Because we are asking whether the difference in means is less than a particular value (zero), this is a one-sided hypothesis test

$H_0: \mu_W \leq \mu_M$ (or $\mu_M - \mu_F \geq 0$)
 $H_1: \mu_W > \mu_M$ (or $\mu_M - \mu_F < 0$)

Hypotheses about Differences in Means: Example #1

- 2. Check that the sample data conform to basic assumptions; if they do not, then do not go any further**

Each of the two random samples are small (<30) and symmetrically distributed with no outliers, and the populations of measurements is bell-shaped (**Not True**)

--- or ---

The size of the random samples are each large (≥ 30), regardless of the shape of the distributions of the measurements in the populations (**True**)

Hypotheses about Differences in Means: Example #1

- 3. Choose an α probability level ... that is, a probability associated with incorrectly rejecting the null hypothesis**

Let's go with $\alpha=0.05$ in this example

Hypotheses about Differences in Means: Example #1

- 4. Determine the "critical value" ... that is, how large the test statistic must be in order to reject the null hypothesis at the given α level**

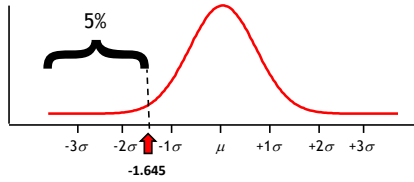
Because we are dealing with sample means, our test statistic will be a t-score (not a Z-score)

What critical value t^* corresponds to having 5% of the area under the normal curve below $-t^*$?

-1.645

Hypotheses about Differences in Means: Example #1

With $\alpha=0.05$, $n_M=8,709$, $n_F=11,530$, & a one-sided test:



... so our critical value t^* is -1.645

$H_0: \mu_W \leq \mu_M$... fail to reject H_0 if $t \geq -1.645$

$H_1: \mu_W > \mu_M$... reject H_0 if $t < -1.645$

Hypotheses about Differences in Means: Example #1

5. Calculate the test statistic

All test statistics for hypothesis testing...

...subtract the population value that is assumed to be true under the null from the observed (sample) value

...divide that figure by the standard deviation of the sampling distribution of the statistic in question

For differences in means:

$$t = \frac{\bar{Y}_M - \bar{Y}_F - 0}{\sqrt{\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}}}$$

Hypotheses about Differences in Means: Example #1

In our example:

$\bar{Y}_M=5.93$ $s_M=2.200$ $n_M=8,709$

$\bar{Y}_F=6.01$ $s_F=2.153$ $n_F=11,530$

$$t = \frac{\bar{Y}_M - \bar{Y}_F - 0}{\sqrt{\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}}} = \frac{5.93 - 6.01 - 0}{\sqrt{\frac{2.200^2}{8,709} + \frac{2.153^2}{11,530}}} = -2.58$$

Hypotheses about Differences in Means: Example #1

6. Compare the test statistic to the critical value

- If the test statistic is larger than the critical value, then reject H_0 (with probability of α of doing so even though H_0 should not actually be rejected)
- If the test statistic is less than or equal to the critical value, then do not reject H_0 (with probability of β of doing so even though H_0 should be rejected)

In our example, our critical value t^* is -1.645

We obtained a test statistic of -2.58

$H_0: \mu_W \leq \mu_M$... fail to reject H_0 if $t \geq -1.645$

$H_1: \mu_W > \mu_M$... reject H_0 if $t < -1.645$

Conclusion? **Reject H_0**

Worksheet

Do men and women tend to bring the same amount of money to class?

Using data from a survey conducted on the first day of my undergraduate research methods class, which included **21** men and **54** women, I found that men brought an average of **\$38** with a standard deviation of **\$59**, while women brought an average of **\$45** with a standard deviation of **\$85**

(Assume for now that students in my class represent a random sample of all undergraduate students)

Use $\alpha=0.05$

Worksheet

Were Minnesota kids more likely to be in families that received food stamps in 2010 than in 2007?

Data from 2007 and 2010 American Community Surveys:

2007: 7.1% of 13,511 sampled MN kids in such families

2010: 11.7% of 12,832 sampled MN kids in such families

1. Construct a 99% confidence interval for the difference in proportions between these two years
2. Test the hypothesis that there was no difference between the two years in the rate of food stamp receipt. Use an α of 0.01.

Want More?

David Lane's Books

http://onlinestatbook.com/2/logic_of_hypothesis_testing/logic_hypothesis.html

http://davidmlane.com/hyperstat/logic_hypothesis.html

Lowry's Book (Chapter 7)

<http://vassarstats.net/textbook/>

Dallal's Book

<http://www.jerrydallal.com/LHSP/sigtest.htm>

Stat Trek's Discussion

<http://stattrek.com/hypothesis-test/hypothesis-testing.aspx>
