Hypothesis Testing

MA 116

Summer 1 2025

Review: Sampling distribution of $\hat{p}$ rules

Fix a population of size N and a characteristic. Let p be the population portion of this characteristic in this population.

Let $\mu_{\hat{p}}$ denote the mean of the sample proportion, and let $\sigma_{\hat{p}}$ denote the standard deviation of the sample proportion. Note that $\sigma_{\hat{p}}$ depends on our choice of n, but this dependency is not reflected in its notation.

Theorem. Assume that any sampling is random.

$\mu_{\hat{p}} = p$. This holds regardless of n or N.

Theorem. Assume $n \leq 0.05N$ and that any sampling is random

1. As n increases, the shape of the distribution of the sample proportion becomes approximately normal. When $np(1-p) \geq 10$, we say the bell curve is a good approximation of the distribution of $\hat{p}$.
2. $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$.

Summary of CH.6

Fix a population and a sample size n. Assume all samplings are random.

Theorem		need $n < 0.05N$?
1	$\mu_{\overline{x}} = \mu$	NO
2	$\sigma_{\overline{x}} = \sigma/\sqrt{n}$	YES
3	x being approximately normal implies $\overline{x}$ is approximately normal	NO
4	CLT: $n \geq 30$ implies $\overline{x}$ is approximately normal	YES
5	$\mu_{\hat{p}} = p$	NO
6	$\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$	YES
7	$np(1-p) \geq 10$ implies $\hat{p}$ is approximately normal	YES

Example

Question. Note population is again not a quantitative data set.

Suppose the true fraction of all US citizens who trust the president is $p = 0.46$. Can you describe the sampling distribution of $\hat{p}$ with a sample size $n = 100$?

Step I. (i). 100 is surely less than $0.05\times$American population;
(ii) $np(1-p) = 100 \cdot 0.46 \cdot 0.54 = 24.84 \geq 10$. Then, we may say the distribution of $\hat{p}$ is approximately normal.
Step II. $\mu_{\hat{p}} = p = 0.46$ as always.
Step III. Since $n < 0.05N$ holds, we have that

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.46 \cdot 0.54}{100}} = 0.05.$$

Step IV. Conclusion: The sampling distribution of $\hat{p}$ with a sample size $n = 100$ is approximately normal with a mean at 0.46 and a standard deviation about 0.05.

Example. Calculating $z_{0.05}$.

Step I. Definition of $z_{0.05}$?
Step II. Draw a standard normal curve and label what 0.05 and $z_{0.05}$ mean in the diagram.
Step III. Look at the Standard Normal Distribution Table and find this $z_{0.05}$.

Example. Using the z_α lemma

The symmetry of the standard normal curve implies the following lemma.

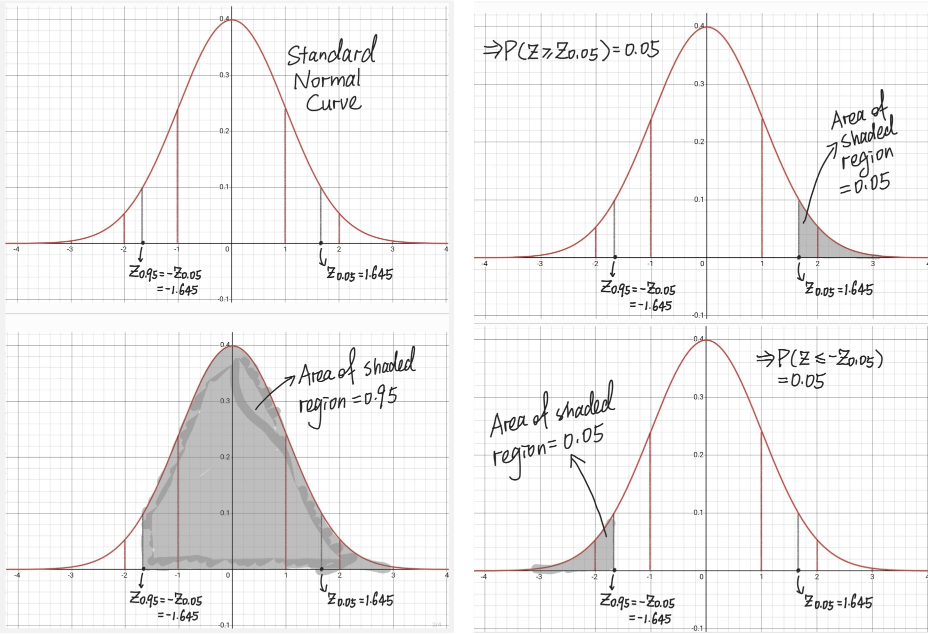
Lemma

For any $0 \leq \alpha \leq 1$ we have $-z_\alpha = z_{1-\alpha}$.

Question. Given the value $z_{0.05} = 1.645$ we just found, can you find $z_{0.95}$ and label $z_{0.95}$ on the diagram of a standard normal curve?
Step I. The lemma tells us that $z_{0.95} = z_{1-0.05} = -z_{0.05} = -1.645$.
Step II. Draw a standard normal curve and label what 0.05, 0.95, $z_{0.05}$, $z_{0.95}$ mean in the diagram.

Result

For any $0 \leq \alpha \leq 1$, we have that $P(z \leq -z_\alpha) = P(z \geq z_\alpha) = \alpha$.



Standard Normal Curve
$Z_{0.95} = -Z_{0.05} = -1.645$
$Z_{0.05} = 1.645$
$\Rightarrow P(Z \geq Z_{0.05}) = 0.05$
Area of shaded region $= 0.05$
$Z_{0.95} = -Z_{0.05} = -1.645$
$Z_{0.05} = 1.645$
Area of shaded region $= 0.95$
$Z_{0.95} = -Z_{0.05} = -1.645$
$Z_{0.05} = 1.645$
$\Rightarrow P(Z \leq -Z_{0.05}) = 0.05$
Area of shaded region $= 0.05$
$Z_{0.95} = -Z_{0.05} = -1.645$
$Z_{0.05} = 1.645$

	Population Parameter	Sample Statistic
Mean	μ	$\overline{x}$
Median	η	M
Variance	σ^2	s^2
Standard Deviation	σ	s
Proportion	p	$\hat{p}$

Plan of class

- 10.1 language of hypothesis testing
- 10.2 hypothesis test for a population proportion
- 10.3 hypothesis test for a population mean
- 11.1 inference about two population proportions
- 11.2 inference about two means
- 11.3 inference about two means
- Chapter 14: Regression

Language of hypothesis testing

Common settings when hypothesis test gets applied.

According to a Gallup poll conducted in 2022, 85% of Americans felt satisfied with the way things were going in their personal lives. A researcher wonders if the percentage of satisfied Americans is different today (a statement regarding a population proportion).

The packaging on a light bulb states that the bulb will last 500 hours under normal use. A consumer advocate would like to know if the mean lifetime of a bulb is less than 500 hours (a statement regarding the population mean).

Definition. (Hypothesis.)

A hypothesis is a statement regarding a **characteristic of a population**.

In the following example,
"80% of Americans felt satisfied with the way things were going in their personal lives" ($p = 0.8$) is a statement of population, so we can regard this statement as a hypothesis. While
"the percentage of satisfied Americans is different today" ($p \neq 0.8$) is also a statement of population, so we can regard this statement as a hypothesis too.

According to a Gallup poll conducted in 2008, 80% of Americans felt satisfied with the way things were going in their personal lives. A researcher wonders if the percentage of satisfied Americans is different today (a statement regarding a population proportion).

Definition. (Hypothesis testing.)

We **test these hypotheses using sample data** because it is usually impossible or impractical to gain access to the entire population. The procedure (or process) we use to test such statements is called **hypothesis testing**.

Because we use sample data to test hypotheses, we cannot state with 100% certainty that the statement is true. Instead, we can only determine whether the sample data **support** the statement, or not.

To conduct a hypothesis test, we say a hypothesis test is based on two types of hypotheses.

1. The **null hypothesis**, denoted H_0, is a statement to be tested. H_0 is a hypothesis that's assumed to be true until evidence indicates otherwise.
2. The **alternative hypothesis**, denoted H_1, is a statement that we are trying to find evidence to support.

Example.

According to a Gallup poll conducted in 2008, 80% of Americans felt satisfied with the way things were going in their personal lives. A researcher wonders if the percentage of satisfied Americans is different today.

$H_0: \ p = 0.8$
$H_1: \ p \neq 0.8$

Steps in Hypothesis Testing

1. Make a statement regarding the nature of the population.
2. Collect evidence (sample data) to test the statement.
3. Analyze the data to assess the plausibility of the statement.

In practice: Steps in Hypothesis Testing

1. Make a statement regarding the nature of the population, i.e. determine the population parameter (μ, p?) we are using, determine H_0 and H_1.
2. Collect evidence (sample data) to test the statement. I.e. Obtain a sample to use. Assume H_0 to be true all the time.
3. Analyze the data to assess the plausibility of the statement. Assume H_0 to be true all the time, i.e. if $H_0: \ p = 0.8$ then we actually take $\mu_{\hat{p}} = p = 0.8$ to construct a distribution of $\hat{p}$, and see where our particular sample's proportion lies on this distribution.

Null hypothesis and alternative hypothesis

Null hypothesis

The null hypothesis is a statement of **status quo** or **no difference** and **always contains a statement of equality**. The null hypothesis is assumed to be true until we have evidence to the contrary. We seek evidence that supports the statement in the alternative hypothesis.

1. **Two-tailed test.** H_0 : parameter$= c$; H_1 : parameter$\neq c$, for c a specific value of the parameter we are testing.
2. **Left-tailed test.** H_0 : parameter$= c$; H_1 : parameter$< c$, for c a specific value of the parameter we are testing.
3. **Right-tailed test.** H_0 : parameter$= c$; H_1 : parameter$> c$, for c a specific value of the parameter we are testing.

Left- and right-tailed tests are referred to as **one-tailed tests**.

Examples.

Two-tailed test.

According to a Gallup poll conducted in 2008, 80% of Americans felt satisfied with the way things were going in their personal lives. A researcher wonders if the percentage of satisfied Americans is different today.

$H_0 : \ p = 0.8$

$H_1 : \ p \neq 0.8$ (because we are trying to show the proportion is different today)

Left-tailed Test.

The packaging on a light bulb states that the bulb will last 500 hours under normal use. A consumer advocate would like to know if the mean lifetime of a bulb is less than 500 hours.

$H_0 : \ \mu = 500$

$H_1 : \ \mu < 500$ (because we are trying to show the mean life time is less that 500 hours)

The statement we are trying to gather evidence for, which is dictated by the researcher before any data are collected, determines the structure of the alternative hypothesis (two-tailed, left-tailed, or right-tailed).

Example.

The label on a can of soda states that the can contains 12 ounces of liquid. A consumer advocate would be concerned only if the mean contents are less than 12 ounces, so the alternative hypothesis is $H_1 : \mu < 12$ ounces.

However, a quality-control engineer for the soda manufacturer would be concerned if there is too little or too much soda in the can, so the alternative hypothesis would be $H_1 : \mu \neq 12$ ounces. In both cases, however, the null hypothesis is a statement of no difference between the manufacturer's assertion on the label and the actual mean contents of the can, so the null hypothesis is $H_0 : \mu = 12$ ounces.

Example.

Determine the null and alternative hypotheses. State whether the test is two-tailed, left-tailed, or right-tailed.

a. The Medco pharmaceutical company has just developed a new antibiotic for children. Of children taking competing antibiotics, 2% experience headaches as a side effect. A researcher for the Food and Drug Administration (FDA) wishes to know if the percentage of children taking the new antibiotic who experience headaches as a side effect is different from $2\% = 0.02$.

b. A placement exam is structured so the mean time to complete the exam is 60 minutes. A community college administrator is concerned that the exam takes longer than 60 minutes.

Step I. Determine the population parameter. μ? p?
Step II. Determine $H_0 : \mu =$? $p =$?

Answer

Determine the null and alternative hypotheses. State whether the test is two-tailed, left-tailed, or right-tailed.

a. The Medco pharmaceutical company has just developed a new antibiotic for children. Of children taking competing antibiotics, 2% experience headaches as a side effect. A researcher for the Food and Drug Administration (FDA) wishes to know if the percentage of children taking the new antibiotic who experience headaches as a side effect is different from $2\% = 0.02$.

b. A placement exam is structured so the mean time to complete the exam is 60 minutes. A community college administrator is concerned that the exam takes longer than 60 minutes.

a. $H_0: p = 0.02$; $H_1: p \neq 0.02$. The test is two-tailed.

Answer

Determine the null and alternative hypotheses. State whether the test is two-tailed, left-tailed, or right-tailed.

a. The Medco pharmaceutical company has just developed a new antibiotic for children. Of children taking competing antibiotics, 2% experience headaches as a side effect. A researcher for the Food and Drug Administration (FDA) wishes to know if the percentage of children taking the new antibiotic who experience headaches as a side effect is different from $2\% = 0.02$.

b. A placement exam is structured so the mean time to complete the exam is 60 minutes. A community college administrator is concerned that the exam takes longer than 60 minutes.

b. $H_0 : \mu = 60$; $H_1 : \mu > 60$. The test is right-tailed.

Type I and Type II Errors

Sample data are used to decide **whether or not to reject** the statement in the null hypothesis. Because this decision is based on incomplete (sample vs. population) information, there is the possibility of making an incorrect decision. In fact, there are four possible outcomes from hypothesis testing.

Four Outcomes from Hypothesis Testing.

1. **Reject the null hypothesis** when the alternative hypothesis is true. This decision would be correct.
2. **Do not reject the null hypothesis** when the null hypothesis is true. This decision would be correct.
3. **Reject the null hypothesis** when the null hypothesis is true. This decision would be incorrect. This is called a **Type I error**.
4. **Do not reject the null hypothesis** when the alternative hypothesis is true. This decision would be incorrect. This is called a **Type II error**.

RMK.

When saying "**Reject the null hypothesis** when the alternative hypothesis is true," the [reject the null hypothesis] is an action we make based on analysis of sample data, while the [the alternative hypothesis is true] is a true fact about our population that we may have no knowledge of.

		Reality	
		H_0 Is True	H_1 Is True
Conclusion	Do Not Reject H_0	Correct Conclusion	Type II Error
	Reject H_0	Type I Error	Correct Conclusion

Example.

The Medco pharmaceutical company has just developed a new antibiotic. Two percent of children taking competing antibiotics experience headaches as a side effect. A researcher for the Food and Drug Administration wishes to know if the percentage of children taking the new antibiotic who experience a headache as a side effect is different from The researcher conducts a hypothesis test with $H_0: p = 0.02$ and $H_1: p \neq 0.02$.

What does it mean to make a Type I error?

A Type I error occurs if the sample evidence lead us to reject the null hypothesis when the null hypothesis is true.

Example.

The Medco pharmaceutical company has just developed a new antibiotic. Two percent of children taking competing antibiotics experience headaches as a side effect. A researcher for the Food and Drug Administration wishes to know if the percentage of children taking the new antibiotic who experience a headache as a side effect is different from The researcher conducts a hypothesis test with $H_0 : p = 0.02$ and $H_1 : p \neq 0.02$.

What does it mean to make a Type I error?

A Type I error is made if the sample leads the researcher to believe that $p \neq 0.02$ (that is, we reject the null hypothesis) when, in fact, the proportion of children who experience a headache is not different from 0.02.

Example.

The Medco pharmaceutical company has just developed a new antibiotic. Two percent of children taking competing antibiotics experience headaches as a side effect. A researcher for the Food and Drug Administration wishes to know if the percentage of children taking the new antibiotic who experience a headache as a side effect is different from The researcher conducts a hypothesis test with $H_0 : p = 0.02$ and $H_1 : p \neq 0.02$.

What does it mean to make a Type II error?

A Type II error occurs if the sample evidence leads us to not reject the null hypothesis when the alternative hypothesis is true.

Example.

The Medco pharmaceutical company has just developed a new antibiotic. Two percent of children taking competing antibiotics experience headaches as a side effect. A researcher for the Food and Drug Administration wishes to know if the percentage of children taking the new antibiotic who experience a headache as a side effect is different from The researcher conducts a hypothesis test with $H_0: p = 0.02$ and $H_1: p \neq 0.02$.

What does it mean to make a Type II error?

A Type II error is made if the sample leads the researcher to not reject the null hypothesis that the proportion of children experiencing a headache is equal to 0.02 when, in fact, the proportion of children who experience a headache is different from 0.02.

The Probability of Making a Type I or Type II Error

Let's regard **making Type I error** and **making Type II error** as two events. Then we may talk about the probabilities that these two events happen. Let

$$\alpha = P(\text{making Type I error})$$

$$\beta = P(\text{making Type II error})$$

level of significance

The probability of making a Type I error, α, is chosen by the researcher before the sample data are collected. This probability is referred to as the level of significance.

Level of significance

The choice of the level of significance depends on the consequences of making a Type I error. If the consequences are severe, the level of significance should be small (say, $\alpha = 0.01$). However, if the consequences are not severe, a higher level of significance can be chosen (say $\alpha = 0.05$ or $\alpha = 0.1$).

Why is the level of significance not always set at $\alpha = 0.01$? Answer: **Reducing the probability of making a Type I error increases the probability of making a Type II error.**

As the probability of a Type I error increases, the probability of a Type II error decreases, and vice versa.

State Conclusions to Hypothesis Tests

This topic is important and will be tested.

Because any hypothesis test decision is based on incomplete (sample vs. population) information, we never say that we **accept** the null hypothesis. without having access to the entire population, we don't know the exact value of the parameter stated in the null hypothesis. Rather, we say that we **do not reject** the null hypothesis if our sample indicates that the null hypothesis H_0 could be true.

The conclusion to a hypothesis test is ALWAYS as follows: There (is/is not) sufficient evidence to conclude that [insert $\mathbf{H}_1$ statement].

Example.

The Medco pharmaceutical company has just developed a new antibiotic. Two percent of children taking competing antibiotics experience headaches as a side effect. A researcher for the Food and Drug Administration wishes to know if the percentage of children taking the new antibiotic who experience a headache as a side effect is different from The researcher conducts a hypothesis test with $H_0: p = 0.02$ and $H_1: p \neq 0.02$.

Suppose that the sample evidence indicates that the null hypothesis is rejected. State the conclusion.

Example.

The Medco pharmaceutical company has just developed a new antibiotic. Two percent of children taking competing antibiotics experience headaches as a side effect. A researcher for the Food and Drug Administration wishes to know if the percentage of children taking the new antibiotic who experience a headache as a side effect is different from The researcher conducts a hypothesis test with $H_0: p = 0.02$ and $H_1: p \neq 0.02$.

Suppose that the sample indicates that the null hypothesis is rejected. State the conclusion.

Answer. Because the null hypothesis ($p = 0.02$) is rejected, there is sufficient evidence to conclude that the proportion of children who experience a headache as a side effect is different from 0.02.

Example.

The Medco pharmaceutical company has just developed a new antibiotic. Two percent of children taking competing antibiotics experience headaches as a side effect. A researcher for the Food and Drug Administration wishes to know if the percentage of children taking the new antibiotic who experience a headache as a side effect is different from The researcher conducts a hypothesis test with $H_0 : p = 0.02$ and $H_1 : p \neq 0.02$.

Suppose that the sample evidence indicates that the null hypothesis is not rejected. State the conclusion.

Example.

The Medco pharmaceutical company has just developed a new antibiotic. Two percent of children taking competing antibiotics experience headaches as a side effect. A researcher for the Food and Drug Administration wishes to know if the percentage of children taking the new antibiotic who experience a headache as a side effect is different from The researcher conducts a hypothesis test with $H_0 : p = 0.02$ and $H_1 : p \neq 0.02$.

Suppose that the sample evidence indicates that the null hypothesis is not rejected. State the conclusion.

Answer. Because the null hypothesis is not rejected, **there is not sufficient evidence to say that the proportion of children who experience a headache as a side effect is different from 0.02.**

It is not accurate to say "Because the null hypothesis is not rejected, the proportion of children who experience a headache as a side effect is 0.02" or "is approximately 0.02."

10.2 Hypothesis testing for a population proportion

Let's recall theorems that may help us describe a sampling distribution of $\hat{p}$. Suppose all samplings are random.

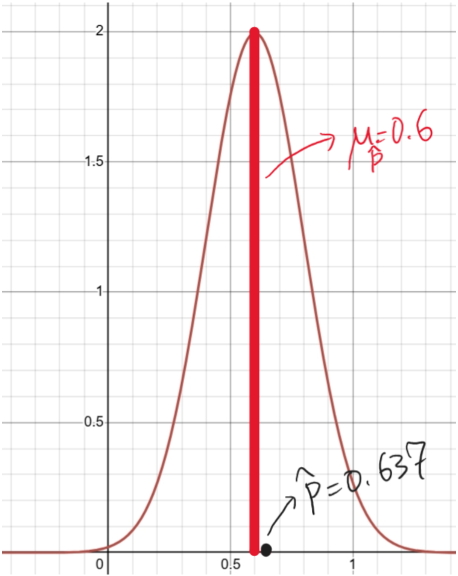
$\mu_{\hat{p}} = p$

When $n < 0.05N$, the formula

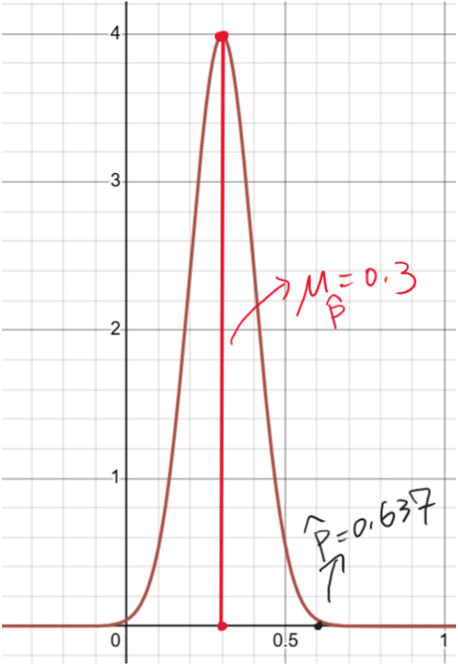
$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

holds.

When $n < 0.05N$ and $np(1-p) \geq 10$, the distribution of $\hat{p}$ is approximately normal.



If we know the sampling distribution of $\hat{p}$ can be approximated by this curve, then a particular sample proportion $\hat{p} = 0.637$ we get seems to be a good estimation of p.



If we know the sampling distribution of $\hat{p}$ can be approximated by this curve, then a particular sample proportion $\hat{p} = 0.637$ we get seems to be a bad estimation of p.

Definition. (Statistically significant.)

Suppose we perform a hypothesis test with a sample. If the sample statistic (e.g. $\hat{p}$ of this particular sample) is unlikely under the assumption that H_0 is true, then we say the result (e.g. $\hat{p}$ of this particular sample) is **statistically significant**. In this case we reject the null hypothesis.

Example. Suppose a politician wants to know if a majority (more than 50%) of their constituents are in favor of a certain policy.

We have $H_0 : p = 0.5$, $H_1 : p > 0.5$.

Suppose the politician hires a polling firm to obtain a random sample of 1000 registered voters in their district and finds that 534 are in favor of the policy, so $\hat{p} = 0.534$ for this particular sample. We'd like to determine whether a sample proportion $\hat{p} = 0.534$ is statistically significant.

A politician wants to know if a majority (more than 50%) of their constituents are in favor of a certain policy. A particular sample of size 100 obtained has $\hat{p} = 0.534$.

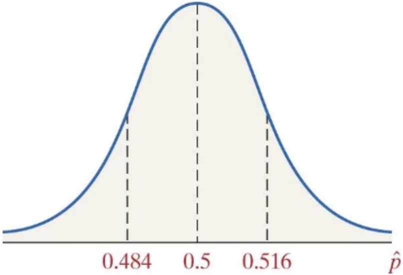
Step I. Assume $\mathbf{p = 0.5}$ and fix a level of confidence α. (Assume the null hypothesis to be true until the sample indicates otherwise.) In this example, let's use $\alpha = 0.05$.

Step II. (i) $n = 1000 < 0.05N$ because the population size N is assumed to be large. (ii) Calculate $np(1-p) = 250 \geq 10$. We may claim that the distribution of $\hat{p}$ is approximately normal.

Step III. $\mu_{\hat{p}} = p = 0.5$. Since the population size N is assumed to be large, we have $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}} = 0.016$.

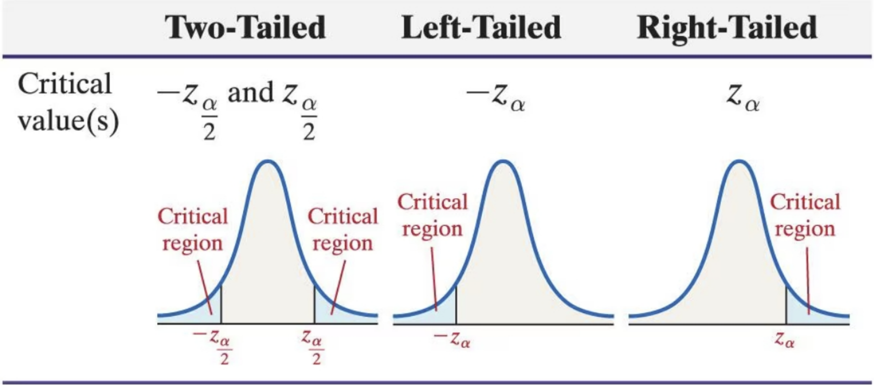
A politician wants to know if a majority (more than 50%) of their constituents are in favor of a certain policy. A particular sample of size 100 obtained has $\hat{p} = 0.534$.

Step IV. Draw a diagram of the normal distribution of $\hat{p}$, still assume that $p = 0.5$.



A politician wants to know if a majority (more than 50%) of their constituents are in favor of a certain policy. A particular sample of size 100 obtained has $\hat{p} = 0.534$.

Step V. Determine if this hypothesis test is two-tailed, left-tailed, or right-tailed. Determine the critical value(s) and the critical region(s).



A politician wants to know if a majority (more than 50%) of their constituents are in favor of a certain policy. A particular sample of size 100 obtained has $\hat{p} = 0.534$.

Step VI. Calculate the **text statistics** via the forluma

$$z_0 = \frac{\hat{p} - \mu_{\hat{p}}}{\sigma_{\hat{p}}}.$$

In this example, we have

$$z_0 = \frac{0.534 - 0.5}{0.016} = 2.15.$$

Step VII. Compare the **text statistics** z_0 to the critical value(s).

Two-Tailed	Left-Tailed	Right-Tailed
If $z_0 < -z_{\frac{\alpha}{2}}$ or $z_0 > z_{\frac{\alpha}{2}}$, reject the null hypothesis.	If $z_0 < -z_{\alpha}$, reject the null hypothesis.	If $z_0 > z_{\alpha}$, reject the null hypothesis.

A politician wants to know if a majority (more than 50%) of their constituents are in favor of a certain policy. A particular sample of size 100 obtained has $\hat{p} = 0.534$.

Since $z_0 = 2.15 > z_{\alpha/2} = z_{0.025} = 1.96$, we reject the null hypothesis.

Step VIII. Draw conclusion to our hypothesis test.

Because the null hypothesis is rejected, there is sufficient evidence to say that a majority (more than 50%) of constituents are in favor of this policy.

Hypothesis testing of population proportion: another method

Confidence interval of a point estimator

Fix a population parameter p. We obtain a particular sample of size n and calculate its sample proportion $\hat{p} = \frac{b}{n}$. Suppose $n < 0.05N$ and $n\hat{p}(1-\hat{p}) \geq 10$. Fix some α that lies between 0 and 1. Consider the interval

$$\left[\hat{p} - z_{\alpha/2}\sigma_{\hat{p}},\, \hat{p} + z_{\alpha/2}\sigma_{\hat{p}}\right].$$

This is called a $(1-\alpha)100\%$ confidence interval of p, or a confidence interval of p with a level of confidence $(1-\alpha)100\%$.

By our conditions we know that $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$. From mathematical experience, this formula implies that $\sigma_{\hat{p}}$ is **insensitive** to changes of p. Thus, in practice, we use

$$\sigma_{\hat{p}} \cong \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

We say a $(1-\alpha)100\%$ confidence interval of p (about $\hat{p}$) is

$$\left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

Example. Suppose we would like to get a 95% confidence interval for p in the setting of

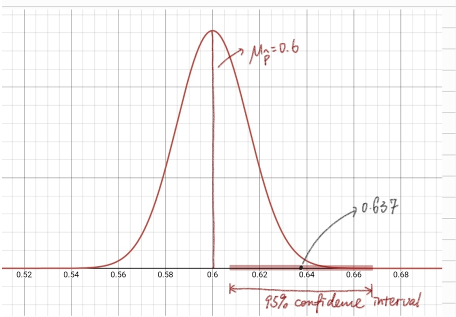
> 1000 people are randomly chosed from all US citizens and 637 answer that they trust the president.

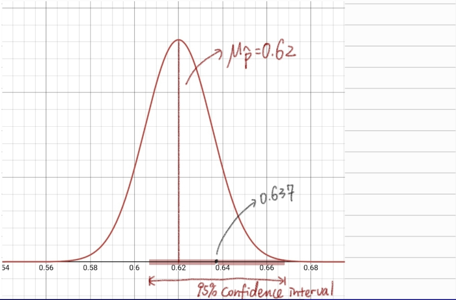
Let's first figure out what our α is. $(1-\alpha)100\% = 95\%$ implies that $\alpha = 0.05$, so $\alpha/2 = 0.025$. SND Table tells us that $z_\alpha = z_{0.025} = 1.96$. Then $\sigma_{\hat{p}} \cong \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \cong 0.0152$, so our 95% confidence interval of p is

$$[0.637 - 1.96 \cdot 0.0152,\ 0.637 + 1.96 \cdot 0.0152]$$

which is about [0.607, 0.667].

Interpret this interval





A 95% confidence interval indicates that 95% of all random samples of size n from the population, whose parameter p we want to estimate, will results in an interval $\left[\hat{p} - z_{0.05/2}\sigma_{\hat{p}},\ \hat{p} + z_{0.05/2}\sigma_{\hat{p}}\right]$ that contains the parameter p.

Interpret this interval

Strictly speaking, a 95% confidence interval

$$\left[\hat{p} - z_{0.05/2}\sigma_{\hat{p}},\ \hat{p} + z_{0.05/2}\sigma_{\hat{p}}\right]$$

obtained from a particular sample does not imply that there is a 95% probability that p lies in this interval. This is because p is assumed to be a fixed value (intrinsic to our population and does not depend on specific sample chosen), not a random value. So saying "there is a 95% probability that p lies in this interval" makes no sense.

From the illustrations, we see that given a particular sample with some $\hat{p}$, we do not know if p lies in the 95% confidence interval about this $\hat{p}$.

We do not know if the random sample we obtained is one of the 95% samples whose interval contain p, or not.