Confidence Intervals

Objectives

1. Determine confidence intervals for population mean
2. Determine confidence intervals for population proportion
3. Determine the necessary sample size
4. Determine a one-sided confidence interval

How Close Are We to the Population Mean?

In the last section, we looked at sampling distributions of the sample mean (and proportion too). Throughout the series, we've used computer simulations to examine statistical concepts.

How Close Are We to the Population Mean?

In the last section, we looked at sampling distributions of the sample mean (and proportion too). Throughout the series, we've used computer simulations to examine statistical concepts.

However, in real life, there are factors that can limit the number of studies and samples we can take: time, cost, etc.

How Close Are We to the Population Mean?

In the last section, we looked at sampling distributions of the sample mean (and proportion too). Throughout the series, we've used computer simulations to examine statistical concepts.

However, in real life, there are factors that can limit the number of studies and samples we can take: time, cost, etc.

So, for our samples, how confident are we that they contain the population mean?

How Close Are We to the Population Mean?

In the last section, we looked at sampling distributions of the sample mean (and proportion too). Throughout the series, we've used computer simulations to examine statistical concepts.

However, in real life, there are factors that can limit the number of studies and samples we can take: time, cost, etc.

So, for our samples, how confident are we that they contain the population mean?

That is where confidence intervals come into play.

How Confident Are We?

Confidence Interval

A **confidence interval** for a population parameter is an estimate of possible values for the parameter with a *given* certain level of confidence.

How Confident Are We?

Confidence Interval

A **confidence interval** for a population parameter is an estimate of possible values for the parameter with a *given* certain level of confidence.

Confidence Level

The **confidence level**, or **level of confidence**, is the percentage of the number of times our confidence intervals will contain the population parameter.

How Confident Are We?

Confidence Interval

A **confidence interval** for a population parameter is an estimate of possible values for the parameter with a *given* certain level of confidence.

Confidence Level

The **confidence level**, or **level of confidence**, is the percentage of the number of times our confidence intervals will contain the population parameter.

Typical confidence levels are 90%, 95%, 98%, and 99%.

Confidence Interval Setup

A confidence interval for a population parameter is in the form

$$\text{point estimate} \pm \text{margin of error}$$

Confidence Interval Setup

A confidence interval for a population parameter is in the form

$$\text{point estimate} \pm \text{margin of error}$$

Point Estimate

A **point estimate** is a value based on our sample data that represents a reasonable value of the population parameter.

Confidence Interval Setup

A confidence interval for a population parameter is in the form

$$\text{point estimate} \pm \text{margin of error}$$

Point Estimate

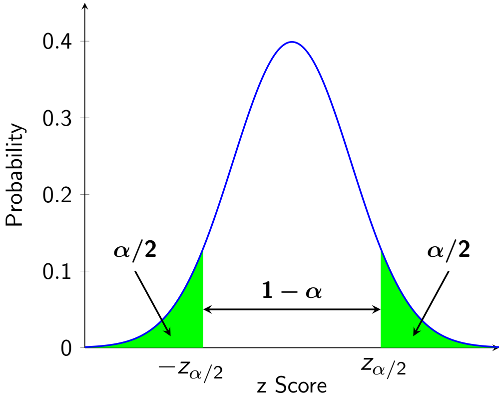
A **point estimate** is a value based on our sample data that represents a reasonable value of the population parameter.

The margin of error is in the form

$$\text{critical value} \times \text{standard error}$$

Critical Values

Critical values are typically in the form $z_{\alpha/2}$ where



Common Critical Values of $z_{\alpha/2}$

Confidence Level	Critical Value of $z_{\alpha/2}$

Common Critical Values of $z_{\alpha/2}$

Confidence Level	Critical Value of $z_{\alpha/2}$
90%	± 1.645

Common Critical Values of $z_{\alpha/2}$

Confidence Level	Critical Value of $z_{\alpha/2}$
90%	± 1.645
95%	± 1.96

Common Critical Values of $z_{\alpha/2}$

Confidence Level	Critical Value of $z_{\alpha/2}$
90%	± 1.645
95%	± 1.96
98%	± 2.326

Common Critical Values of $z_{\alpha/2}$

Confidence Level	Critical Value of $z_{\alpha/2}$
90%	± 1.645
95%	± 1.96
98%	± 2.326
99%	± 2.576

Common Critical Values of $z_{\alpha/2}$

Confidence Level	Critical Value of $z_{\alpha/2}$
90%	± 1.645
95%	± 1.96
98%	± 2.326
99%	± 2.576

Notice the higher the confidence level, the further away from $z = 0$ the critical value is.

Common Critical Values of $z_{\alpha/2}$

Confidence Level	Critical Value of $z_{\alpha/2}$
90%	± 1.645
95%	± 1.96
98%	± 2.326
99%	± 2.576

Notice the higher the confidence level, the further away from $z = 0$ the critical value is.

The standard error is still $\frac{\sigma}{\sqrt{n}}$

Common Critical Values of $z_{\alpha/2}$

Confidence Level	Critical Value of $z_{\alpha/2}$
90%	± 1.645
95%	± 1.96
98%	± 2.326
99%	± 2.576

Notice the higher the confidence level, the further away from $z = 0$ the critical value is.

The standard error is still $\frac{\sigma}{\sqrt{n}}$

Note: If σ is unknown, you can use the sample standard deviation, s, when the sample size is large enough ($n \geq 30$).

Example 1

The waiting time of a sample of 100 patients at a hospital is 3.5 minutes with $s = 0.75$ minutes. Construct a 95% confidence interval for the mean waiting time of all patients at the hospital.

Example 1

The waiting time of a sample of 100 patients at a hospital is 3.5 minutes with $s = 0.75$ minutes. Construct a 95% confidence interval for the mean waiting time of all patients at the hospital.

$$\overline{x} \pm z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

Example 1

The waiting time of a sample of 100 patients at a hospital is 3.5 minutes with $s = 0.75$ minutes. Construct a 95% confidence interval for the mean waiting time of all patients at the hospital.

$$\overline{x} \pm z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

$$3.5 \pm 1.96\left(\frac{0.75}{\sqrt{100}}\right)$$

Example 1

The waiting time of a sample of 100 patients at a hospital is 3.5 minutes with $s = 0.75$ minutes. Construct a 95% confidence interval for the mean waiting time of all patients at the hospital.

$$\overline{x} \pm z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

$$3.5 \pm 1.96\left(\frac{0.75}{\sqrt{100}}\right)$$

$$3.5 \pm 0.147$$

Example 1

The waiting time of a sample of 100 patients at a hospital is 3.5 minutes with $s = 0.75$ minutes. Construct a 95% confidence interval for the mean waiting time of all patients at the hospital.

$$\overline{x} \pm z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

$$3.5 \pm 1.96\left(\frac{0.75}{\sqrt{100}}\right)$$

$$3.5 \pm 0.147$$

$$= 3.353 \text{ to } 3.647$$

Example 1

The waiting time of a sample of 100 patients at a hospital is 3.5 minutes with $s = 0.75$ minutes. Construct a 95% confidence interval for the mean waiting time of all patients at the hospital.

$$\overline{x} \pm z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

$$3.5 \pm 1.96\left(\frac{0.75}{\sqrt{100}}\right)$$

$$3.5 \pm 0.147$$

$$= 3.353 \text{ to } 3.647$$

A 95% confidence interval for the population mean waiting time is 3.353 to 3.647 minutes.

Error Bars

Often times, margins of error are graphed as *error bars* on a visual display.

Error Bars

Often times, margins of error are graphed as *error bars* on a visual display.

However, be advised that some graphs use standard deviation for their error bars.

Example 2

Suppose we take a sample of 45 gas prices and find the mean price per gallon is \$2.45 with a standard deviation of \$0.12.

(a) Create a 98% confidence interval for the population mean price per gallon.

Example 2

Suppose we take a sample of 45 gas prices and find the mean price per gallon is \$2.45 with a standard deviation of \$0.12.

(a) Create a 98% confidence interval for the population mean price per gallon.

$$2.45 \pm 2.326 \left(\frac{0.12}{\sqrt{45}} \right)$$

Example 2

Suppose we take a sample of 45 gas prices and find the mean price per gallon is \$2.45 with a standard deviation of \$0.12.

(a) Create a 98% confidence interval for the population mean price per gallon.

$$2.45 \pm 2.326\left(\frac{0.12}{\sqrt{45}}\right)$$

$$= 2.45 \pm 0.042$$

Example 2

Suppose we take a sample of 45 gas prices and find the mean price per gallon is \$2.45 with a standard deviation of \$0.12.

(a) Create a 98% confidence interval for the population mean price per gallon.

$$2.45 \pm 2.326\left(\frac{0.12}{\sqrt{45}}\right)$$

$$= 2.45 \pm 0.042$$

$$= 2.408 \text{ to } 2.492$$

Example 2

Suppose we take a sample of 45 gas prices and find the mean price per gallon is \$2.45 with a standard deviation of \$0.12.

(a) Create a 98% confidence interval for the population mean price per gallon.

$$2.45 \pm 2.326\left(\frac{0.12}{\sqrt{45}}\right)$$

$$= 2.45 \pm 0.042$$

$$= 2.408 \text{ to } 2.492$$

A 98% confidence interval for the mean population price per gallon is \$2.41 to \$2.49

Example 2

(b) What would the confidence interval be if the sample size was 100?

Example 2

(b) What would the confidence interval be if the sample size was 100?

$$2.45 \pm 2.326 \left(\frac{0.12}{\sqrt{100}} \right)$$

Example 2

(b) What would the confidence interval be if the sample size was 100?

$$2.45 \pm 2.326\left(\frac{0.12}{\sqrt{100}}\right)$$

$$= 2.422 \text{ to } 2.478$$

Example 2

(b) What would the confidence interval be if the sample size was 100?

$$2.45 \pm 2.326 \left(\frac{0.12}{\sqrt{100}} \right)$$

$$= 2.422 \text{ to } 2.478$$

A 98% confidence interval for the mean population price per gallon is \$2.42 to \$2.48

Example 2

(c) Keeping the confidence interval the same, what happens as we increase the sample size?

Example 2

(c) Keeping the confidence interval the same, what happens as we increase the sample size?

Increasing the sample size decreases the standard error.

Example 2

(c) Keeping the confidence interval the same, what happens as we increase the sample size?

Increasing the sample size decreases the standard error.

By doing so, we decrease the width of our confidence interval, while still keeping the confidence level the same.

What the Confidence Interval Is Not

The following are incorrect interpretations of a confidence interval:

What the Confidence Interval Is Not

The following are incorrect interpretations of a confidence interval:

- There is a ____ % chance that the population mean is in this interval.

What the Confidence Interval Is Not

The following are incorrect interpretations of a confidence interval:

- There is a ____ % chance that the population mean is in this interval.
- There is a ____ % chance that the population mean is (*whatever the sample mean is*).

Objectives

1. Determine confidence intervals for population mean
2. Determine confidence intervals for population proportion
3. Determine the necessary sample size
4. Determine a one-sided confidence interval

Mean and Standard Error

Our point estimate for a population proportion is $\mu_{\hat{p}} = p$ and the standard error is $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

Mean and Standard Error

Our point estimate for a population proportion is $\mu_{\hat{p}} = p$ and the standard error is $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

Our confidence interval for the population proportion is

$$p \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

Example 3

In a recent sample of 578 voters, 301 of them favored a proposal to renovate a local park. Construct a 90% confidence interval for the population proportion of voters who favor the renovation.

Example 3

In a recent sample of 578 voters, 301 of them favored a proposal to renovate a local park. Construct a 90% confidence interval for the population proportion of voters who favor the renovation.

$\hat{p} = \frac{301}{578} \approx 0.521$

Example 3

In a recent sample of 578 voters, 301 of them favored a proposal to renovate a local park. Construct a 90% confidence interval for the population proportion of voters who favor the renovation.

$\hat{p} = \frac{301}{578} \approx 0.521$

$$0.521 \pm 1.645\sqrt{\frac{0.521(0.479)}{578}}$$

Example 3

In a recent sample of 578 voters, 301 of them favored a proposal to renovate a local park. Construct a 90% confidence interval for the population proportion of voters who favor the renovation.

$\hat{p} = \frac{301}{578} \approx 0.521$

$$0.521 \pm 1.645\sqrt{\frac{0.521(0.479)}{578}}$$

$$= 0.521 \pm 1.645(0.0208)$$

Example 3

In a recent sample of 578 voters, 301 of them favored a proposal to renovate a local park. Construct a 90% confidence interval for the population proportion of voters who favor the renovation.

$\hat{p} = \frac{301}{578} \approx 0.521$

$$0.521 \pm 1.645\sqrt{\frac{0.521(0.479)}{578}}$$

$$= 0.521 \pm 1.645(0.0208)$$

$$= 0.521 \pm 0.034$$

Example 3

In a recent sample of 578 voters, 301 of them favored a proposal to renovate a local park. Construct a 90% confidence interval for the population proportion of voters who favor the renovation.

$\hat{p} = \frac{301}{578} \approx 0.521$

$$\begin{aligned} & 0.521 \pm 1.645\sqrt{\frac{0.521(0.479)}{578}} \\ &= 0.521 \pm 1.645(0.0208) \\ &= 0.521 \pm 0.034 \\ &= (0.487,\ 0.555) \end{aligned}$$

Example 3

In a recent sample of 578 voters, 301 of them favored a proposal to renovate a local park. Construct a 90% confidence interval for the population proportion of voters who favor the renovation.

$\hat{p} = \frac{301}{578} \approx 0.521$

$$
\begin{aligned}
&0.521 \pm 1.645\sqrt{\frac{0.521(0.479)}{578}} \\
&= 0.521 \pm 1.645(0.0208) \\
&= 0.521 \pm 0.034 \\
&= (0.487, 0.555)
\end{aligned}
$$

A 90% confidence interval for the population proportion of voters who favor the renovation is 48.7 to 55.5%

Example 4

The results of a sample of 40 students who took a pass/fail statistics exam are shown below. Construct a 98% confidence interval for the population proportion of students who passed the exam.

Pass	Pass	Fail	Pass	Pass	Pass	Fail	Pass
Fail	Pass	Pass	Pass	Fail	Pass	Pass	Pass
Pass	Fail	Fail	Pass	Pass	Fail	Pass	Pass
Pass	Pass	Fail	Fail	Pass	Pass	Pass	Fail

Example 4

$$p = 0.75$$

Example 4

$$p = 0.75$$

$$0.75 \pm 2.326\sqrt{\frac{0.75(0.25)}{40}}$$

Example 4

$$p = 0.75$$

$$0.75 \pm 2.326\sqrt{\frac{0.75(0.25)}{40}}$$
$$= 0.75 \pm 2.326(0.07)$$

Example 4

$$p = 0.75$$

$$\begin{aligned} & 0.75 \pm 2.326\sqrt{\frac{0.75(0.25)}{40}} \\ & = 0.75 \pm 2.326(0.07) \\ & = 0.75 \pm 0.159 \end{aligned}$$

Example 4

$$p = 0.75$$

$$
\begin{aligned}
& 0.75 \pm 2.326\sqrt{\frac{0.75(0.25)}{40}} \\
& = 0.75 \pm 2.326(0.07) \\
& = 0.75 \pm 0.159 \\
& = (0.591,\ 0.909)
\end{aligned}
$$

Example 4

$$p = 0.75$$

$$\begin{aligned} & 0.75 \pm 2.326\sqrt{\frac{0.75(0.25)}{40}} \\ &= 0.75 \pm 2.326(0.07) \\ &= 0.75 \pm 0.159 \\ &= (0.591,\, 0.909) \end{aligned}$$

A 98% confidence interval for the population proportion of students who passed the exam is 59 to 91%.

Objectives

1. Determine confidence intervals for population mean
2. Determine confidence intervals for population proportion
3. Determine the necessary sample size
4. Determine a one-sided confidence interval

Margin of Error for Population Mean

Our margin of error, E, for a population mean has been given as

$$E = z_{\alpha/2} \left(\frac{\sigma}{\sqrt{n}} \right)$$

Margin of Error for Population Mean

Our margin of error, E, for a population mean has been given as

$$E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Solving for n:

Margin of Error for Population Mean

Our margin of error, E, for a population mean has been given as

$$E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Solving for n:

$$\frac{E}{z_{\alpha/2}} = \frac{\sigma}{\sqrt{n}}$$

Margin of Error for Population Mean

Our margin of error, E, for a population mean has been given as

$$E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Solving for n:

$$\frac{E}{z_{\alpha/2}} = \frac{\sigma}{\sqrt{n}}$$

$$E\sqrt{n} = \sigma \cdot z_{\alpha/2}$$

Margin of Error for Population Mean

Our margin of error, E, for a population mean has been given as

$$E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Solving for n:

$$\frac{E}{z_{\alpha/2}} = \frac{\sigma}{\sqrt{n}}$$

$$E\sqrt{n} = \sigma \cdot z_{\alpha/2}$$

$$\sqrt{n} = \frac{\sigma \cdot z_{\alpha/2}}{E}$$

Margin of Error for Population Mean

Our margin of error, E, for a population mean has been given as

$$E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Solving for n:

$$\frac{E}{z_{\alpha/2}} = \frac{\sigma}{\sqrt{n}}$$

$$E\sqrt{n} = \sigma \cdot z_{\alpha/2}$$

$$\sqrt{n} = \frac{\sigma \cdot z_{\alpha/2}}{E}$$

$$n = \left(\frac{\sigma \cdot z_{\alpha/2}}{E}\right)^2$$

Example 5

How large of a sample size should we take if we want to create a 95% confidence interval that is within \$0.02 per gallon of the mean cost of gasoline? Assume the standard deviation is \$0.12.

Example 5

How large of a sample size should we take if we want to create a 95% confidence interval that is within \$0.02 per gallon of the mean cost of gasoline? Assume the standard deviation is \$0.12.

$$n = \left(\frac{\sigma \cdot z_{\alpha/2}}{E}\right)^2$$

Example 5

How large of a sample size should we take if we want to create a 95% confidence interval that is within \$0.02 per gallon of the mean cost of gasoline? Assume the standard deviation is \$0.12.

$$n = \left(\frac{\sigma \cdot z_{\alpha/2}}{E}\right)^2$$

$$n = \left(\frac{0.12 \cdot 1.96}{0.02}\right)^2$$

Example 5

How large of a sample size should we take if we want to create a 95% confidence interval that is within \$0.02 per gallon of the mean cost of gasoline? Assume the standard deviation is \$0.12.

$$n = \left(\frac{\sigma \cdot z_{\alpha/2}}{E}\right)^2$$

$$n = \left(\frac{0.12 \cdot 1.96}{0.02}\right)^2$$

$$n = 138.2976$$

Example 5

How large of a sample size should we take if we want to create a 95% confidence interval that is within \$0.02 per gallon of the mean cost of gasoline? Assume the standard deviation is \$0.12.

$$n = \left(\frac{\sigma \cdot z_{\alpha/2}}{E}\right)^2$$

$$n = \left(\frac{0.12 \cdot 1.96}{0.02}\right)^2$$

$$n = 138.2976 \to 139$$

Example 5

How large of a sample size should we take if we want to create a 95% confidence interval that is within \$0.02 per gallon of the mean cost of gasoline? Assume the standard deviation is \$0.12.

$$n = \left(\frac{\sigma \cdot z_{\alpha/2}}{E}\right)^2$$

$$n = \left(\frac{0.12 \cdot 1.96}{0.02}\right)^2$$

$$n = 138.2976 \to 139$$

We would need a sample of at least 139.

Sample Size for Population Proportion

If we solve

$$E = z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

for n, we get

$$n = p(1-p)\left(\frac{z_{\alpha/2}}{E}\right)^2$$

Example 6

A recent poll shows that 301 out of 578 voters favor renovating a local park. How large of a sample must we take to create a 95% confidence interval that is within 4% of the population proportion?

Example 6

A recent poll shows that 301 out of 578 voters favor renovating a local park. How large of a sample must we take to create a 95% confidence interval that is within 4% of the population proportion?

$p \approx 0.521$

Example 6

A recent poll shows that 301 out of 578 voters favor renovating a local park. How large of a sample must we take to create a 95% confidence interval that is within 4% of the population proportion?

$p \approx 0.521$

$$n = 0.521(0.479)\left(\frac{1.96}{0.04}\right)^2$$

Example 6

A recent poll shows that 301 out of 578 voters favor renovating a local park. How large of a sample must we take to create a 95% confidence interval that is within 4% of the population proportion?

$p \approx 0.521$

$$n = 0.521(0.479)\left(\frac{1.96}{0.04}\right)^2$$

$$n = 599.191159$$

Example 6

A recent poll shows that 301 out of 578 voters favor renovating a local park. How large of a sample must we take to create a 95% confidence interval that is within 4% of the population proportion?

$p \approx 0.521$

$$n = 0.521(0.479)\left(\frac{1.96}{0.04}\right)^2$$

$$n = 599.191159 \to 600$$

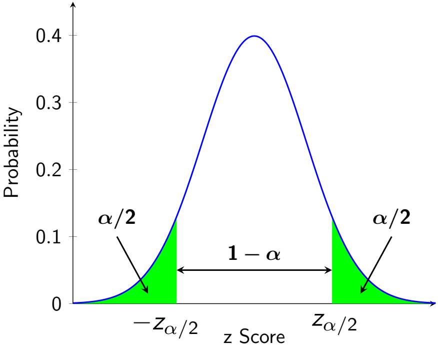
We would need a sample size of at least 600.

Objectives

1. Determine confidence intervals for population mean
2. Determine confidence intervals for population proportion
3. Determine the necessary sample size
4. Determine a one-sided confidence interval

Two-Sided vs. One-Sided

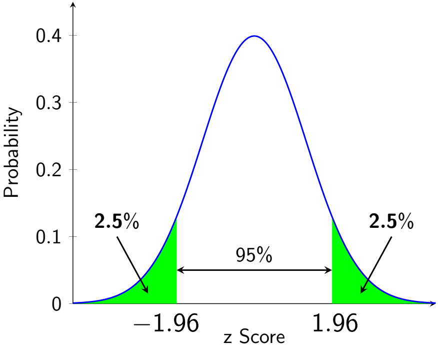
With a two-sided confidence interval (what we've discussed so far), the total area outside of our interval is *alpha*:



Notice the endpoints of our interval are $\pm z_{\alpha/2}$.

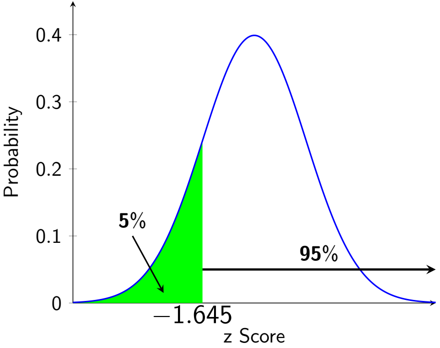
Two-Sided vs. One-Sided

For example, constructing a 95% confidence interval will mean that a total of 5% will be outside of our interval:



Two-Sided vs. One-Sided

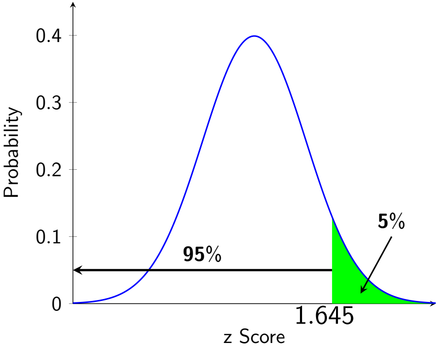
For a one-sided interval, in order to have the same area outside of our interval, our critical values will need to change:



The value $z = -1.645$ is a **lower bound** on the one-sided 95% confidence interval.

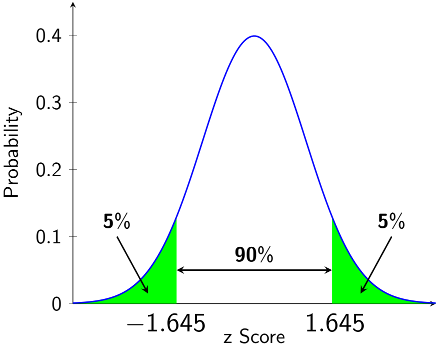
Two-Sided vs. One-Sided

The value $z = 1.645$ is an **upper bound** on the one-sided 95% confidence interval going the other way (to the left).



Two-Sided vs. One-Sided

We can also get those values of $z = -1.645$ and $z = 1.645$ by constructing a 90% confidence interval:



Two-Sided vs. One-Sided

For left- or right-tailed hypothesis testing, some statistical software will report the one-sided interval, others will only report the usual two-sided interval with an adjusted value of *alpha*:

Critical Values

Two-Tailed	One-Tailed
$z_{\alpha/2}$	z_{α}