In chapter 7, we had access to the population mean $\bar{x}$ and we made probability statements about individual $x$ values taken from the population.

In Chapter 8, we began with a population having a known mean $\mu$ or proportion $p$;

Then we examined the sampling distribution of the corresponding sample statistic ($\bar{x}$ or $\hat{p}$) for samples of a given size, $n$.

In this chapter, we’ll be going in the opposite direction: based on sample data, we will be making estimates involving the (unknown) value of the population mean or proportion.
Definition

A **point estimate** of a population parameter is a single number that estimates the exact value of that parameter.

An **interval estimate** of a population parameter is an interval which includes a range of possible values that are likely to include the actual population parameter.

Example

Suppose the average GPA in a sample of 100 Fordham students is 3.3, what is the average GPA at Fordham?

**point estimate** \( \mu = 3.3 \)

**interval estimate**: \( \mu \in [2.8, 3.7] \)
An estimator is **unbiased** if the expected value of the sample statistic is the same as the actual value of the population parameter it is intended to estimate. For example $\bar{x}$, $\hat{p}$, and $s^2$ are unbiased estimators of $\mu$, $p$, and $\sigma^2$, respectively.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimator</th>
<th>Formula</th>
<th>Expected Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean: $\mu$</td>
<td>$\bar{x}$</td>
<td>$\frac{\sum_{i=1}^{n} x_i}{n}$</td>
<td>$\mathbb{E}[\bar{x}] = \mu$</td>
</tr>
<tr>
<td>Proportion: $p$</td>
<td>$\hat{p}$</td>
<td>$\hat{p} = \frac{x}{n}$</td>
<td>$\mathbb{E}[\hat{p}] = p$</td>
</tr>
<tr>
<td>Variance, $\sigma^2$</td>
<td>$s^2$</td>
<td>$\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{(n-1)}$</td>
<td>$\mathbb{E}[s^2] = \sigma^2$</td>
</tr>
</tbody>
</table>
Confidence Intervals: Definitions

- INTERVAL ESTIMATE: A range of values within which the actual value of the population parameter may fall.

- CONFIDENCE INTERVAL: An interval estimate for which there is a specified degree of certainty that the actual value of the population parameter will fall within the interval.

- CONFIDENCE LEVEL: This expresses the degree of certainty that an interval will include the actual value of the population parameter. It is usually stated as a percentage, commonly 90%, 95%, or 99%.

- LEVEL OF SIGNIFICANCE $\alpha$: This expresses the probability that an interval will NOT include the actual value of the population parameter. Note that $\alpha = 1 - \text{Confidence Level}$
Estimating Confidence Intervals of the Mean

Suppose we take a simple random sample of size $n$ from a population, and let $\bar{x}$ be the mean of this sample and $s^2$ its standard deviation. Estimating a confidence interval around the mean $\mu$ depends on whether or not the populations standard deviation $\sigma$ is known:

If $\sigma = \{$

- is known $\quad \mu \in [\bar{x} \pm z \ast \frac{\sigma}{\sqrt{n}}]$
- is unknown $\quad \mu \in [\bar{x} \pm t \ast \frac{s}{\sqrt{n}}]$, and $df = n - 1$

Where $z = \text{the zscore corresponding to the level of confidence desired. For example, } z = 1.96 \text{ corresponds to the 95% confidence level.}$

We will come back to $t$ latter.
Suppose we take a simple random sample of size $n$, with $\bar{x}$, $s^2$. Also, suppose $\sigma$ is known, then

$$\mu \in [\bar{x} \pm z * \frac{\sigma}{\sqrt{n}}]$$

Where $z =$ the $z$ score corresponding to the level of confidence desired.

Assumptions: this assumes that either (1) the underlying population is normally distributed or (2) the sample size is $n > 30$. 
### $z$ Scores for Confidence Intervals

Commonly used confidence intervals and their corresponding $z$ values:

<table>
<thead>
<tr>
<th>Confidence</th>
<th>90%</th>
<th>95%</th>
<th>98%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z$</td>
<td>1.645</td>
<td>1.96</td>
<td>2.33</td>
<td>2.58</td>
</tr>
</tbody>
</table>

![Normal distribution diagram](image)
Find the $z$ score associated with 85% confidence level.

**Answer:**
Find the $z$ score associated with 85% confidence level.

**Answer:** We want $z$ such that $P[-z < Z < z] = 0.85$. Notice, it’s sufficient to find $z$ or $-z$. So:

1. To find $z$, note that $P[-z < Z < z] = 0.85$ implies that $P[Z > z] = \frac{1-0.85}{2}$, due to the symmetry of the normal distribution.
Find the $z$ score associated with 85% confidence level. 

**Answer:** We want $z$ such that $P[-z < Z < z] = 0.85$. Notice, it’s sufficient to find $z$ or $-z$. So:

1. To find $z$, note that $P[-z < Z < z] = 0.85$ implies that $P[Z > z] = \frac{1-0.85}{2}$, due to the symmetry of the normal distribution.
   This implies that $P[Z < z] = 1 - \frac{1-0.85}{2} = 0.925$. 

   Using this probability with the $Z$ table, we can find that $z = 1.44$. 

   Notice we can stop here because $-z = -1.44$. 

2. To find $-z$, note that $P[-z < Z < z] = 0.85$ implies,$ P[-z < Z] = 1 - \frac{1-0.85}{2} = 0.075$ due to the symmetry of the normal distribution.
   Using this probability with the $Z$ table, we can find that $-z = -1.44$. 

   Notice we can stop here because $z = 1.44$. 

Meshry (Fordham University) 

Chapter 9 

November 27, 2017 9 / 36
Find the $z$ score associated with 85% confidence level.

**Answer:** We want $z$ such that $P[-z < Z < z] = 0.85$. Notice, it’s sufficient to find $z$ or $-z$. So:

1. To find $z$, note that $P[-z < Z < z] = 0.85$ implies that $P[Z > z] = \frac{1-0.85}{2}$, due to the symmetry of the normal distribution.
   This implies that $P[Z < z] = 1 - \frac{1-0.85}{2} = 0.925$.
   Using this probability with the $Z$ table, we can find that $z = 1.44$. 
Find the $z$ score associated with 85% confidence level.

**Answer:** We want $z$ such that $P[-z < Z < z] = 0.85$. Notice, it’s sufficient to find $z$ or $-z$. So:

1. To find $z$, note that $P[-z < Z < z] = 0.85$ implies that $P[Z > z] = \frac{1-0.85}{2}$, due to the symmetry of the normal distribution.
   This implies that $P[Z < z] = 1 - \frac{1-0.85}{2} = 0.925$.
   Using this probability with the $Z$ table, we can find that $z = 1.44$.
   Notice we can stop here because $-z = -1.44$

2. To find $-z$, note that $P[-z < Z < z] = 0.85$ implies, $P[-z < Z] = \frac{1-0.85}{2} = 0.075$ due to the symmetry of the normal distribution.
Scores for Confidence Intervals: Example

Find the $z$ score associated with 85% confidence level.

**Answer:** We want $z$ such that $P[-z < Z < z] = 0.85$. Notice, it’s sufficient to find $z$ or $-z$. So:

1. To find $z$, note that $P[-z < Z < z] = 0.85$ implies that $P[Z > z] = \frac{1 - 0.85}{2}$, due to the symmetry of the normal distribution.
   This implies that $P[Z < z] = 1 - \frac{1 - 0.85}{2} = 0.925$.
   Using this probability with the $Z$ table, we can find that $z = 1.44$.
   Notice we can stop here because $-z = -1.44$.

2. To find $-z$, note that $P[-z < Z < z] = 0.85$ implies,
   $P[-z < Z] = \frac{1 - 0.85}{2} = 0.075$ due to the symmetry of the normal distribution.
   Using this probability with the $Z$ table, we can find that $-z = -1.44$. 

Meshry (Fordham University)
From past experience, the population standard deviation of rod diameters produced by a machine has been found to be $\sigma = 0.053$ inches. For a simple random sample of $n = 30$ rods, the average diameter is found to be $\bar{x} = 1.4$ inches. What Is the 95% Confidence Interval for the Population Mean, $\mu$?

$$\bar{x} - z \frac{\sigma}{\sqrt{n}} \quad \bar{x} \quad \bar{x} + z \frac{\sigma}{\sqrt{n}}$$

$$1.400 - 1.96 \frac{0.053}{\sqrt{30}} \quad 1.400 \quad 1.400 + 1.96 \frac{0.053}{\sqrt{30}}$$

or

$$1.381 \quad 1.400 \quad 1.419$$
The following data values are a simple random sample from a population that is normally distributed, with $\sigma = 4$: {8, 10, 7, 8, 5, 13, 7, 10, 4, 6}. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n =$
The following data values are a simple random sample from a population that is normally distributed, with $\sigma = 4$: 
$\{8, 10, 7, 8, 5, 13, 7, 10, 4, 6\}$. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n = 10$, $\bar{x} =$
Estimating Confidence Intervals of the Mean: $\sigma$ is known, Example 2

The following data values are a simple random sample from a population that is normally distributed, with $\sigma = 4$: 
{8, 10, 7, 8, 5, 13, 7, 10, 4, 6}. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n = 10$, $\bar{x} = 7.8$, and $\sigma = 4$. So depending on the confidence level:
90% CI: $\mu \in \left[7.8 - 1.645 \frac{4}{\sqrt{10}}, 7.8 + 1.645 \frac{4}{\sqrt{10}}\right]$ 
95% CI: $\mu \in \left[7.8 - 1.96 \frac{4}{\sqrt{10}}, 7.8 + 1.96 \frac{4}{\sqrt{10}}\right]$ 
99% CI: $\mu \in \left[7.8 - 2.58 \frac{4}{\sqrt{10}}, 7.8 + 2.58 \frac{4}{\sqrt{10}}\right]$
The following data values are a simple random sample from a population that is normally distributed, with \( \sigma = 4 \): 
\{8, 10, 7, 8, 5, 13, 7, 10, 4, 6\}. Construct and interpret the 90\%, 95\%, and 99\% confidence intervals for the population mean, \( \mu \).

Answer: Notice that \( n = 10 \), \( \bar{x} = 7.8 \), and \( \sigma = 4 \). So depending on the confidence level:
90\% CI: \( \mu \in [7.8 - 1.645 \frac{4}{\sqrt{10}}, 7.8 + 1.645 \frac{4}{\sqrt{10}}] \Rightarrow \mu \in [5.7192, 9.8808] \)
95\% CI: \( \mu \in \)
Estimating Confidence Intervals of the Mean: $\sigma$ is known, Example 2

The following data values are a simple random sample from a population that is normally distributed, with $\sigma = 4$: \{8, 10, 7, 8, 5, 13, 7, 10, 4, 6\}. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n = 10$, $\bar{x} = 7.8$, and $\sigma = 4$. So depending on the confidence level:

- **90% CI:** $\mu \in [7.8 - 1.645 \frac{4}{\sqrt{10}}, 7.8 + 1.645 \frac{4}{\sqrt{10}}] \Rightarrow \mu \in [5.7192, 9.8808]
- **95% CI:** $\mu \in [7.8 - 1.96 \frac{4}{\sqrt{10}}, 7.8 + 1.96 \frac{4}{\sqrt{10}}] \Rightarrow \mu \in [5.3208, 10.2792]
- **99% CI:** $\mu \in [7.8 - 2.576 \frac{4}{\sqrt{10}}, 7.8 + 2.576 \frac{4}{\sqrt{10}}] \Rightarrow \mu \in [4.5365, 11.0635]$
Estimating Confidence Intervals of the Mean: $\sigma$ is known, Example 2

The following data values are a simple random sample from a population that is normally distributed, with $\sigma = 4$: \{8, 10, 7, 8, 5, 13, 7, 10, 4, 6\}. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n = 10$, $\bar{x} = 7.8$, and $\sigma = 4$. So depending on the confidence level:

90% CI: $\mu \in [7.8 - 1.645 \frac{4}{\sqrt{10}}, 7.8 + 1.645 \frac{4}{\sqrt{10}}] \Rightarrow \mu \in [5.7192, 9.8808]$

95% CI: $\mu \in [7.8 - 1.96 \frac{4}{\sqrt{10}}, 7.8 + 1.96 \frac{4}{\sqrt{10}}] \Rightarrow \mu \in [5.3208, 10.2792]$

99% CI: $\mu \in [7.8 - 2.58 \frac{4}{\sqrt{10}}, 7.8 + 2.58 \frac{4}{\sqrt{10}}] \Rightarrow \mu \in [4.5365, 11.0635]$
A simple random sample of 30 has been collected from a population for which it is known that $\sigma = 10$. The sample mean has been calculated as $\bar{x} = 240$. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n =$
A simple random sample of 30 has been collected from a population for which it is known that \( \sigma = 10 \). The sample mean has been calculated as \( \bar{x} = 240 \). Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, \( \mu \).

Answer: Notice that \( n = 30, \bar{x} = \)
A simple random sample of 30 has been collected from a population for which it is known that \( \sigma = 10 \). The sample mean has been calculated as \( \bar{x} = 240 \). Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, \( \mu \).

Answer: Notice that \( n = 30 \), \( \bar{x} = 240 \) and \( \sigma = 10 \).
Estimating Confidence Intervals of the Mean: $\sigma$ is known, Example 3

A simple random sample of 30 has been collected from a population for which it is known that $\sigma = 10$. The sample mean has been calculated as $\bar{x} = 240$. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n = 30$, $\bar{x} = 240$ and $\sigma = 10$. So depending on the confidence level:
90% CI: $\mu \in$
Estimating Confidence Intervals of the Mean: $\sigma$ is known, Example 3

A simple random sample of 30 has been collected from a population for which it is known that $\sigma = 10$. The sample mean has been calculated as $\bar{x} = 240$. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n = 30$, $\bar{x} = 240$ and $\sigma = 10$. So depending on the confidence level:

90% CI: $\mu \in [240 - 1.645 \frac{10}{\sqrt{30}}, 240 + 1.645 \frac{10}{\sqrt{30}}] \Rightarrow \mu \in [236.997, 243.003]$  
95% CI: $\mu \in$
Estimating Confidence Intervals of the Mean: $\sigma$ is known, Example 3

A simple random sample of 30 has been collected from a population for which it is known that $\sigma = 10$. The sample mean has been calculated as $\bar{x} = 240$. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n = 30$, $\bar{x} = 240$ and $\sigma = 10$. So depending on the confidence level:

90% CI: $\mu \in [240 - 1.645 \frac{10}{\sqrt{30}}, 240 + 1.645 \frac{10}{\sqrt{30}}] \Rightarrow \mu \in [236.997, 243.003]$

95% CI: $\mu \in [240 - 1.96 \frac{10}{\sqrt{30}}, 240 + 1.96 \frac{10}{\sqrt{30}}] \Rightarrow \mu \in [236.4215, 243.5785]$

99% CI: $\mu \in$
Estimating Confidence Intervals of the Mean: $\sigma$ is known, Example 3

A simple random sample of 30 has been collected from a population for which it is known that $\sigma = 10$. The sample mean has been calculated as $\bar{x} = 240$. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n = 30$, $\bar{x} = 240$ and $\sigma = 10$. So depending on the confidence level:
- 90% CI: $\mu \in [240 - 1.645 \frac{10}{\sqrt{30}}, 240 + 1.645 \frac{10}{\sqrt{30}}] \Rightarrow \mu \in [236.997, 243.003]$
- 95% CI: $\mu \in [240 - 1.96 \frac{10}{\sqrt{30}}, 240 + 1.96 \frac{10}{\sqrt{30}}] \Rightarrow \mu \in [236.4215, 243.5785]$
- 99% CI: $\mu \in [240 - 2.58 \frac{10}{\sqrt{30}}, 240 + 2.58 \frac{10}{\sqrt{30}}] \Rightarrow \mu \in [235.2896, 244.7104]$
A simple random sample of 25 has been collected from a normally distributed population for which it is known that \( \sigma = 17 \). The sample mean has been calculated as 342.0, and the sample standard deviation is \( s = 14.9 \). Construct and interpret the 90\%, 95\%, and 99\% confidence intervals for the population mean, \( \mu \).

Answer: Notice that \( n = \)
A simple random sample of 25 has been collected from a normally distributed population for which it is known that \( \sigma = 17 \). The sample mean has been calculated as 342.0, and the sample standard deviation is \( s = 14.9 \). Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, \( \mu \).

Answer: Notice that \( n = 25, \bar{x} = \)
A simple random sample of 25 has been collected from a normally distributed population for which it is known that $\sigma = 17$. The sample mean has been calculated as 342.0, and the sample standard deviation is $s = 14.9$. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n = 25$, $\bar{x} = 342$, and $\sigma = 17$. So depending on the confidence level:

90% CI: $\mu \in \ldots$
A simple random sample of 25 has been collected from a normally distributed population for which it is known that $\sigma = 17$. The sample mean has been calculated as 342.0, and the sample standard deviation is $s = 14.9$. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n = 25$, $\bar{x} = 342$, and $\sigma = 17$. So depending on the confidence level:

90% CI: $\mu \in [342 - 1.645 \frac{17}{\sqrt{25}}, 342 + 1.645 \frac{17}{\sqrt{25}}] \Rightarrow \mu \in [336.407, 347.593]$

95% CI: $\mu \in$
A simple random sample of 25 has been collected from a normally distributed population for which it is known that $\sigma = 17$. The sample mean has been calculated as 342.0, and the sample standard deviation is $s = 14.9$. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n = 25$, $\bar{x} = 342$, and $\sigma = 17$. So depending on the confidence level:

90% CI: $\mu \in [342 - 1.645 \frac{17}{\sqrt{25}} , 342 + 1.645 \frac{17}{\sqrt{25}}] \Rightarrow \mu \in [336.407, 347.593]$

95% CI: $\mu \in [342 - 1.96 \frac{17}{\sqrt{25}} , 342 + 1.96 \frac{17}{\sqrt{25}}] \Rightarrow \mu \in [335.336, 348.664]$

99% CI: $\mu \in$
Estimating Confidence Intervals of the Mean: $\sigma$ is known, Example 4

A simple random sample of 25 has been collected from a normally distributed population for which it is known that $\sigma = 17$. The sample mean has been calculated as 342.0, and the sample standard deviation is $s = 14.9$. Construct and interpret the 90%, 95%, and 99% confidence intervals for the population mean, $\mu$.

Answer: Notice that $n = 25$, $\bar{x} = 342$, and $\sigma = 17$. So depending on the confidence level:

90% CI: $\mu \in [342 - 1.645 \frac{17}{\sqrt{25}}, 342 + 1.645 \frac{17}{\sqrt{25}}] \Rightarrow \mu \in [336.407, 347.593]$

95% CI: $\mu \in [342 - 1.96 \frac{17}{\sqrt{25}}, 342 + 1.96 \frac{17}{\sqrt{25}}] \Rightarrow \mu \in [335.336, 348.664]$

99% CI: $\mu \in [342 - 2.58 \frac{17}{\sqrt{25}}, 342 + 2.58 \frac{17}{\sqrt{25}}] \Rightarrow \mu \in [333.228, 350.772]$
You want to rent a one-bedroom apartment in NYC. You find the mean monthly rent for a random sample of 60 apartments advertised on Craig’s List is $1600. Assume a population standard deviation of $400. Construct 90%, 95%, and 99% confidence intervals for the average monthly one-bedroom rent in NYC.

Answer: Notice that $n =$
You want to rent a one-bedroom apartment in NYC. You find the mean monthly rent for a random sample of 60 apartments advertised on Craig’s List is $1600. Assume a population standard deviation of $400. Construct 90%, 95%, and 99% confidence intervals for the average monthly one-bedroom rent in NYC.

Answer: Notice that $n = 60$, $\bar{x} =$
Estimating Confidence Intervals of the Mean: \( \sigma \) is known, Example 5

You want to rent a one-bedroom apartment in NYC. You find the mean monthly rent for a random sample of 60 apartments advertised on Craig’s List is $1600. Assume a population standard deviation of $400. Construct 90%, 95%, and 99% confidence intervals for the average monthly one-bedroom rent in NYC.

Answer: Notice that \( n = 60, \bar{x} = 342, \) and \( \sigma = 17. \) So depending on the confidence level:
90% CI: \( \mu \in \)
You want to rent a one-bedroom apartment in NYC. You find the mean monthly rent for a random sample of 60 apartments advertised on Craig’s List is $1600. Assume a population standard deviation of $400. Construct 90%, 95%, and 99% confidence intervals for the average monthly one-bedroom rent in NYC.

Answer: Notice that $n = 60$, $\bar{x} = 342$, and $\sigma = 17$. So depending on the confidence level:

90% CI: $\mu \in [1600 – 1.645 \frac{400}{\sqrt{60}} , 1600 + 1.645 \frac{400}{\sqrt{60}}] \Rightarrow \mu \in [1515.1, 1684.9]

95% CI: $\mu \in$
You want to rent a one-bedroom apartment in NYC. You find the mean monthly rent for a random sample of 60 apartments advertised on Craig’s List is $1600. Assume a population standard deviation of $400. Construct 90%, 95%, and 99% confidence intervals for the average monthly one-bedroom rent in NYC.

Answer: Notice that $n = 60$, $\bar{x} = 342$, and $\sigma = 17$. So depending on the confidence level:

90% CI: $\mu \in [1600 - 1.645 \frac{400}{\sqrt{60}}, 1600 + 1.645 \frac{400}{\sqrt{60}}]$ $\Rightarrow \mu \in [1515.1, 1684.9]$

95% CI: $\mu \in [1600 - 1.96 \frac{400}{\sqrt{60}}, 1600 + 1.96 \frac{400}{\sqrt{60}}]$ $\Rightarrow \mu \in [1498.8, 1701.2]$

99% CI: $\mu \in$
You want to rent a one-bedroom apartment in NYC. You find the mean monthly rent for a random sample of 60 apartments advertised on Craig’s List is $1600. Assume a population standard deviation of $400. Construct 90%, 95%, and 99% confidence intervals for the average monthly one-bedroom rent in NYC.

Answer: Notice that \( n = 60, \bar{x} = 342, \) and \( \sigma = 17. \) So depending on the confidence level:

90% CI: \( \mu \in [1600 - 1.645 \frac{400}{\sqrt{60}}, 1600 + 1.645 \frac{400}{\sqrt{60}}] \Rightarrow \mu \in [1515.1, 1684.9] \)

95% CI: \( \mu \in [1600 - 1.96 \frac{400}{\sqrt{60}}, 1600 + 1.96 \frac{400}{\sqrt{60}}] \Rightarrow \mu \in [1498.8, 1701.2] \)

99% CI: \( \mu \in [1600 - 2.58 \frac{400}{\sqrt{60}}, 1600 + 2.58 \frac{400}{\sqrt{60}}] \Rightarrow \mu \in [1466.8, 1733.2] \)
Confidence Interval Estimate for the Population Proportion

Suppose a proportion $p$ of individuals in a population have a certain trait, and assume we don’t know the value of $p$. Take a sample of size $n$, and let $\hat{p}$ be the proportion of individuals in the sample with that trait. We can construct a confidence interval for the population proportion $p$ as follows:

$$p \in \left[ \hat{p} \pm z \star \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \right] = \left[ \hat{p} \pm z \star \sqrt{\frac{\hat{p}\hat{q}}{n}} \right]$$

This assumes the normal distribution as an approximation to the binomial distribution, which holds whenever $n\hat{p} > 5$ and $n\hat{q} > 5$. The approximation becomes better for large values of $n$ and whenever $\hat{p}$ is closer to 0.5.
Out of a sample of 1008 adults, 22% responded YES to a survey question. What is the 95% confidence interval for the population proportion who would have answered “YES” to the same question?

Answer: \( \hat{p} = 0.22 \), and \( z = \pm 1.96 \), so

\[
\hat{p} \pm z \sqrt{\frac{\hat{p} \hat{q}}{n}} = 0.22 \pm 1.96 \sqrt{\frac{0.22 \times 0.78}{1008}} = [0.194, 0.246]
\]
A pharmaceutical company found that 46% of 1000 U.S. adults surveyed knew neither their blood pressure nor their cholesterol level. Assuming the persons surveyed to be a simple random sample of U.S. adults, construct 99% confidence interval for \( p \), the population proportion of U.S. adults who would have given the same answer if a census had been taken instead of a survey.

Answer: \( \hat{p} = 0.46 \), and \( z = \pm 2.58 \), so 99% CI:

\[
p \in \left[ \hat{p} \pm z \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right] = \left[ 0.46 \pm 2.58 \right] = [0.42, 0.50]\]
Confidence Interval Estimate for the Population Proportion: Example 2

A pharmaceutical company found that 46% of 1000 U.S. adults surveyed knew neither their blood pressure nor their cholesterol level. Assuming the persons surveyed to be a simple random sample of U.S. adults, construct 99% confidence interval for $p$, the population proportion of U.S. adults who would have given the same answer if a census had been taken instead of a survey.

Answer: $\hat{p} = 0.46$, and $z = \pm 2.58$, so 99% CI:

$p \in \left[ \hat{p} \pm z * \sqrt{\frac{\hat{p}\hat{q}}{n}} \right] = \left[ 0.46 \pm 2.58 * \sqrt{\frac{0.46*0.54}{1000}} \right] = [0.42, 0.50]$
Confidence Interval Estimate for the Population Proportion: Example 3

An airline has surveyed a simple random sample of air travelers to find out whether they would be interested in paying a higher fare in order to have access to e-mail during their flight. Of the 400 travelers surveyed, 80 said email access would be worth a slight extra cost. Construct 90% confidence interval for the population proportion of air travelers who are in favor of the airline’s email idea.

Answer: \( \hat{p} = \frac{80}{400} = 0.2 \), and \( z = \pm 1.645 \), so 99% CI:

\[ p \in \]
An airline has surveyed a simple random sample of air travelers to find out whether they would be interested in paying a higher fare in order to have access to e-mail during their flight. Of the 400 travelers surveyed, 80 said email access would be worth a slight extra cost. Construct 90% confidence interval for the population proportion of air travelers who are in favor of the airline’s email idea.

Answer: \( \hat{p} = \frac{80}{400} = 0.2 \), and \( z = \pm 1.645 \), so 99% CI:

\[
p \in \left[ \hat{p} \pm z \star \sqrt{\frac{\hat{p}\hat{q}}{n}} \right] = \left[ 0.2 \pm 1.645 \star \sqrt{\frac{0.2\times0.8}{400}} \right] = [0.18, 0.22]
\]