Confidence Intervals for Means

So far we have been mostly concerned with confidence intervals for proportions.

Confidence Intervals for Means

So far we have been mostly concerned with confidence intervals for proportions.

Often we wish to determine the average of some measure in a population other than a proportion.

Confidence Intervals for Means

So far we have been mostly concerned with confidence intervals for proportions.

Often we wish to determine the average of some measure in a population other than a proportion.

The same problems we mentioned in connection with sampling for proportions are present.

Confidence Intervals for Means

So far we have been mostly concerned with confidence intervals for proportions.

Often we wish to determine the average of some measure in a population other than a proportion.

The same problems we mentioned in connection with sampling for proportions are present.

As a result, sampling is often the only viable option.

Confidence Intervals for Means

So far we have been mostly concerned with confidence intervals for proportions.

Often we wish to determine the average of some measure in a population other than a proportion.

The same problems we mentioned in connection with sampling for proportions are present.

As a result, sampling is often the only viable option.

As before, it is important to determine the accuracy of the results.

Confidence Intervals for Means

As usual, we assume a simple random sample is drawn.

Confidence Intervals for Means

As usual, we assume a simple random sample is drawn.

We assume the sampling is done without replacement.

Confidence Intervals for Means

As usual, we assume a simple random sample is drawn.

We assume the sampling is done without replacement.

We also assume that the population is large enough that this makes little difference.

Confidence Intervals for Means

As usual, we assume a simple random sample is drawn.

We assume the sampling is done without replacement.

We also assume that the population is large enough that this makes little difference.

If you sample from a very large population, sampling with and without replacement are pretty much the same.

Confidence Intervals for Means

As usual, we assume a simple random sample is drawn.

We assume the sampling is done without replacement.

We also assume that the population is large enough that this makes little difference.

If you sample from a very large population, sampling with and without replacement are pretty much the same.

If you sample from a small population, it can make a difference.

Confidence Intervals for Means

As usual, we assume a simple random sample is drawn.

We assume the sampling is done without replacement.

We also assume that the population is large enough that this makes little difference.

If you sample from a very large population, sampling with and without replacement are pretty much the same.

If you sample from a small population, it can make a difference.

There are ways of adjusting the results to allow for this.

Confidence Intervals for Means

Suppose the sample consists of n individuals,

$$(x_1, x_2, x_3, \ldots, x_n)$$

Confidence Intervals for Means

Suppose the sample consists of n individuals,

$$(x_1, x_2, x_3, \ldots, x_n)$$

We calculate the sample mean, denoted by $\overline{x}$, in the usual way:

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Confidence Intervals for Means

Suppose the sample consists of n individuals,

$$(x_1, x_2, x_3, \ldots, x_n)$$

We calculate the sample mean, denoted by $\overline{x}$, in the usual way:

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

The only difference from sampling proportions is that the x values are usually not all zeroes and ones.

Confidence Intervals for Means

As before, we want some idea of the accuracy of our estimated mean.

Confidence Intervals for Means

As before, we want some idea of the accuracy of our estimated mean.

In order to calculate this, we need additional information.

Unlike the situation with proportions, knowing the mean does not tell us anything about the standard deviation of the $x's$.

Confidence Intervals for Means

As before, we want some idea of the accuracy of our estimated mean.

In order to calculate this, we need additional information.

Unlike the situation with proportions, knowing the mean does not tell us anything about the standard deviation of the $x's$.

The only difference from sampling proportions is that the x values are usually not all zeroes and ones.

Confidence Intervals for Means

There are two common ways of dealing with this problem.

Confidence Intervals for Means

There are two common ways of dealing with this problem.

The first way is to treat the standard deviation of the $x's$ as a known quantity.

Confidence Intervals for Means

There are two common ways of dealing with this problem.

The first way is to treat the standard deviation of the $x's$ as a known quantity.

This may be realistic if you have earlier results from a very large sample.

Confidence Intervals for Means

There are two common ways of dealing with this problem.

The first way is to treat the standard deviation of the $x's$ as a known quantity.

This may be realistic if you have earlier results from a very large sample.

Another possibility is that the population was standardized in some way and the standard deviation is known.

Confidence Intervals for Means

There are two common ways of dealing with this problem.

The first way is to treat the standard deviation of the $x's$ as a known quantity.

This may be realistic if you have earlier results from a very large sample.

Another possibility is that the population was standardized in some way and the standard deviation is known.

Common IQ tests are standardized to have a standard deviation of 15.

Standard Deviation is Known

Either way, we treat the standard deviation as a known quantity SD_x.

Standard Deviation is Known

Either way, we treat the standard deviation as a known quantity SD_x.

In this situation, SD_x is assumed to have a constant numerical value.

Standard Deviation is Known

Either way, we treat the standard deviation as a known quantity SD_x.

In this situation, SD_x is assumed to have a constant numerical value.

If the population follows a bell curve, SD_x is the same as the standard deviation of the bell curve.

Standard Deviation is Known

Either way, we treat the standard deviation as a known quantity SD_x.

In this situation, SD_x is assumed to have a constant numerical value.

If the population follows a bell curve, SD_x is the same as the standard deviation of the bell curve.

The standard deviation of the mean of a sample of size n is then given by:

Standard Deviation is Known

Either way, we treat the standard deviation as a known quantity SD_x.

In this situation, SD_x is assumed to have a constant numerical value.

If the population follows a bell curve, SD_x is the same as the standard deviation of the bell curve.

The standard deviation of the mean of a sample of size n is then given by:

$$SD_{\overline{x}} = \frac{SD_x}{\sqrt{n}}$$

Standard Deviation is Unknown

The other common situation is that we *do not* know the standard deviation of x.

Standard Deviation is Unknown

The other common situation is that we *do not* know the standard deviation of x.

Many times, there is no reliable estimate of SD_x available.

Standard Deviation is Unknown

The other common situation is that we *do not* know the standard deviation of x.

Many times, there is no reliable estimate of SD_x available.

In this case, the standard deviation of x must be estimated from the sample.

Standard Deviation is Unknown

The other common situation is that we *do not* know the standard deviation of x.

Many times, there is no reliable estimate of SD_x available.

In this case, the standard deviation of x must be estimated from the sample.

This means that the estimates of the mean and standard deviation both come from the sample.

Standard Deviation is Unknown

The other common situation is that we *do not* know the standard deviation of x.

Many times, there is no reliable estimate of SD_x available.

In this case, the standard deviation of x must be estimated from the sample.

This means that the estimates of the mean and standard deviation both come from the sample.

Standard Deviation is Known

When we have estimated both the sample mean and standard deviation from the sample, the estimates are slightly less accurate than if the variance was known.

Standard Deviation is Known

When we have estimated both the sample mean and standard deviation from the sample, the estimates are slightly less accurate than if the variance was known.

In this case, the sample mean has a distribution similar to a bell curve, but slightly wider.

Standard Deviation is Known

When we have estimated both the sample mean and standard deviation from the sample, the estimates are slightly less accurate than if the variance was known.

In this case, the sample mean has a distribution similar to a bell curve, but slightly wider.

This distribution is called the t distribution, or sometimes *student's t* distribution

Standard Deviation is Known

When we have estimated both the sample mean and standard deviation from the sample, the estimates are slightly less accurate than if the variance was known.

In this case, the sample mean has a distribution similar to a bell curve, but slightly wider.

This distribution is called the t distribution, or sometimes *student's t* distribution

Most statistics books will have a table for the t distribution in the appendix.

Standard Deviation is Known

When we have estimated both the sample mean and standard deviation from the sample, the estimates are slightly less accurate than if the variance was known.

In this case, the sample mean has a distribution similar to a bell curve, but slightly wider.

This distribution is called the t distribution, or sometimes *student's t* distribution

Most statistics books will have a table for the t distribution in the appendix.

It is a fact that the t distribution becomes indistinguishable from the bell curve as the sample size grows.