Bell Curves for Sample Means

Gene Quinn

Sample Means

Suppose a measure on a certain population follows a bell curve distribution with a mean of 70 and a standard deviation of 5.

If we take a sample of size 20 from this population and compute the sample mean or average, the question arises of what distribution should be used for the sample mean:

Does the sample mean $\overline{x}$ have a bell curve distribution?

Sample Means

Suppose a measure on a certain population follows a bell curve distribution with a mean of 70 and a standard deviation of 5.

If we take a sample of size 20 from this population and compute the sample mean or average, the question arises of what distribution should be used for the sample mean: Does the sample mean $\overline{x}$ have a bell curve distribution?

If so,

- What is the mean of the bell curve?
- What is the standard deviation?

The Normal Curve

If a measure in the *population* has a bell curve distribution with:

- mean u
- standard deviation s

The Normal Curve

If a measure in the *population* has a bell curve distribution with:

- mean u
- standard deviation s

Then the mean of a sample of size n also has a bell curve distribution with:

- mean u
- standard deviation $s/\sqrt{n}$

Sample Means

Even if the underlying population *does not* have a bell curve distribution, under very mild assumptions the sample mean *will* follow a bell curve, at least approximately.

Sample Means

Even if the underlying population *does not* have a bell curve distribution, under very mild assumptions the sample mean *will* follow a bell curve, at least approximately.

This result is called the *Central Limit Theorem*

Sample Means

Even if the underlying population *does not* have a bell curve distribution, under very mild assumptions the sample mean *will* follow a bell curve, at least approximately.

This result is called the *Central Limit Theorem*

The Central Limit Theorem is usually assumed to hold provided the individuals in the sample are independently selected and the sample size is 30 or more.

Sample Means

When the underlying population has a bell curve distribution, or the Central Limit Theorem applies, the *sample mean* for a sample of size n, given by

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

will have a bell curve distribution with the same mean as the population.

Sample Means

When the underlying population has a bell curve distribution, or the Central Limit Theorem applies, the *sample mean* for a sample of size n, given by

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

will have a bell curve distribution with the same mean as the population.

The only change is that the standard deviation of the sample is the population standard deviation, divided by $\sqrt{n}$

Sample Means

When the underlying population has a bell curve distribution, or the Central Limit Theorem applies, the *sample mean* for a sample of size n, given by

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

will have a bell curve distribution with the same mean as the population.

The only change is that the standard deviation of the sample is the population standard deviation, divided by $\sqrt{n}$

Once you adjust the standard deviation, you can treat the sample mean just like any other measurement with a bell curve distribution.

Example

Suppose a measure has a bell curve distribution in the population with mean 70 and standard deviation 5.

If a random sample of size 36 is taken and the sample mean or average is computed, what can be said about the distribution of the sample mean $\overline{x}$?

Example

Suppose a measure has a bell curve distribution in the population with mean 70 and standard deviation 5.

If a random sample of size 36 is taken and the sample mean or average is computed, what can be said about the distribution of the sample mean $\overline{x}$?

Answer: The sample mean should have a bell curve distribution with:

- a mean of 70
- a standard deviation of $5/\sqrt{36}$ or $5/6$

Example

Suppose a measure has a bell curve distribution in the population with mean 100 and standard deviation 15.

If a random sample of size 225 is taken and the sample mean or average is computed, what can be said about the distribution of the sample mean $\overline{x}$?

Example

Suppose a measure has a bell curve distribution in the population with mean 100 and standard deviation 15.

If a random sample of size 225 is taken and the sample mean or average is computed, what can be said about the distribution of the sample mean $\overline{x}$?

Answer: The sample mean should have a bell curve distribution with:

- a mean of 100
- a standard deviation of $15/\sqrt{225}$ or $15/15 = 1$

Example

Suppose a sample of 100 is taken population with mean of 60 and standard deviation 20.

If the sample mean or average is computed, what can be said about the distribution of the sample mean $\overline{x}$?

Example

Suppose a sample of 100 is taken population with mean of 60 and standard deviation 20.

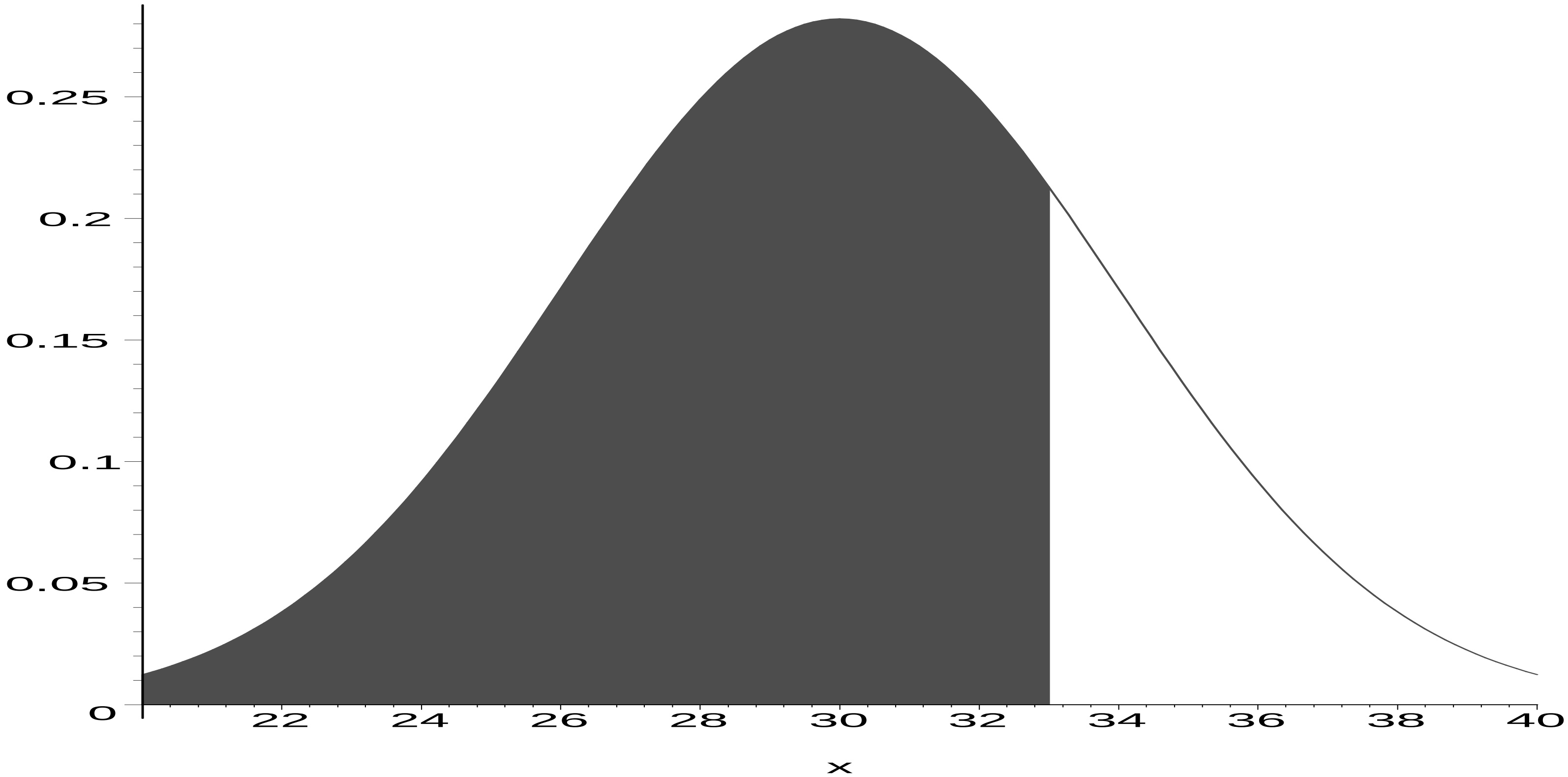
If the sample mean or average is computed, what can be said about the distribution of the sample mean $\overline{x}$?

Answer: The Central Limit Theorem applies, and the sample mean should have a bell curve distribution with:

- a mean of 60
- a standard deviation of $20/\sqrt{100}$ or $20/10 = 2$

Sample Means Below a Given Value

The proportion of *sample means* below a given value is:



=NORMDIST(value,u,s/SQRT(n),true)

Sample Means Below a Given Value

Example: Waist measurements of a certain population have bell curve distribution with mean 34.2 and standard deviation 2.1.

What population of samples of 100 have a mean waist measurement of 34.5 or less?

Sample Means Below a Given Value

Example: Waist measurements of a certain population have bell curve distribution with mean 34.2 and standard deviation 2.1.

What population of samples of 100 have a mean waist measurement of 34.5 or less?

In this case the syntax for the formula is:

=NORMDIST(34.5,34.2,2.1/SQRT(100),true)

Sample Means Below a Given Value

Example: Waist measurements of a certain population have bell curve distribution with mean 34.2 and standard deviation 2.1.

What population of samples of 100 have a mean waist measurement of 34.5 or less?

In this case the syntax for the formula is:

=NORMDIST(34.5,34.2,2.1/SQRT(100),true)

In this case, the result is 0.92

Sample Means Below a Given Value

Example: Waist measurements of a certain population have bell curve distribution with mean 34.2 and standard deviation 2.1.

What population of samples of 100 have a mean waist measurement of 34.5 or less?

In this case the syntax for the formula is:

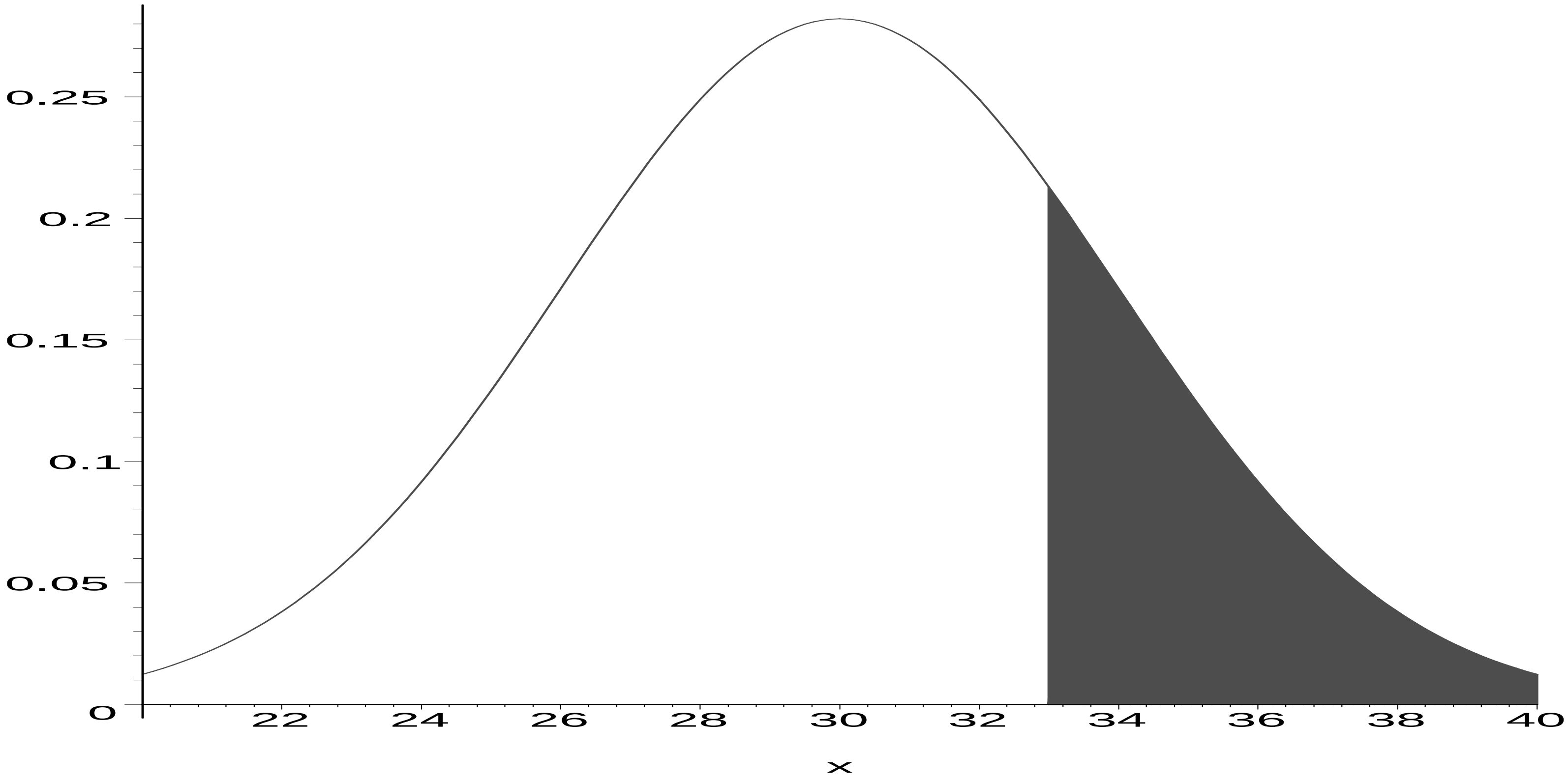
=NORMDIST(34.5,34.2,2.1/SQRT(100),true)

In this case, the result is 0.92

This tells us 92 percent of *samples* of 100 from a bell curve population with mean 34.2 and standard deviation 2.1 will have their average value less than or equal to 34.5.

Sample Means Above a Given Value

For samples of size n, the proportion of *sample means* above a given value is:



=1-NORMDIST(value,u,s/SQRT(n),true)

Sample Means Above a Given Value

Example: Waist measurements of a certain population have bell curve distribution with mean 34.2 and standard deviation 2.1. A sample of size 144 is taken and the average is computed.

What proportion of the time will this average be greater than 34?

Sample Means Above a Given Value

Example: Waist measurements of a certain population have bell curve distribution with mean 34.2 and standard deviation 2.1. A sample of size 144 is taken and the average is computed.

What proportion of the time will this average be greater than 34?

In this case the syntax for the formula is:

=1-NORMDIST(34,34.2,2.1/SQRT(144),true)

Sample Means Above a Given Value

Example: Waist measurements of a certain population have bell curve distribution with mean 34.2 and standard deviation 2.1. A sample of size 144 is taken and the average is computed.

What proportion of the time will this average be greater than 34?

In this case the syntax for the formula is:

=1-NORMDIST(34,34.2,2.1/SQRT(144),true)

In this case, the result is 0.87

Sample Means Above a Given Value

Example: Waist measurements of a certain population have bell curve distribution with mean 34.2 and standard deviation 2.1. A sample of size 144 is taken and the average is computed.

What proportion of the time will this average be greater than 34?

In this case the syntax for the formula is:

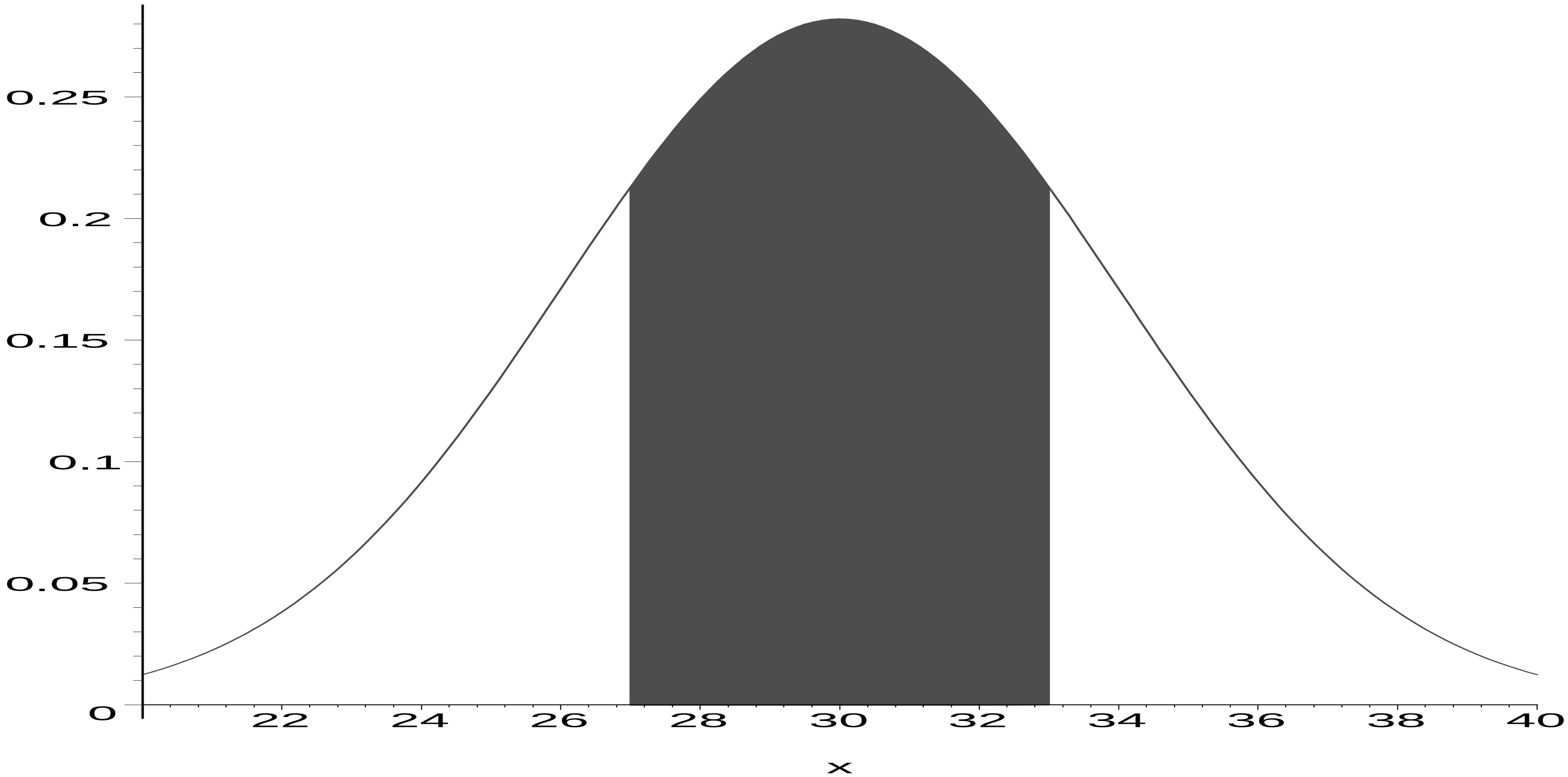
=1-NORMDIST(34,34.2,2.1/SQRT(144),true)

In this case, the result is 0.87

This tells us that the mean of a sample of size 144 from a population with mean 34.2 and standard deviation 2.1 will be 34 or greater 87 percent of the time.

Sample Means Between Two Values

The proportion of means of a sample of size n between two given values **a** and **b** is:



NORMDIST(b,u,s/SQRT(n),true)-

NORMDIST(a,m,s/SQRT(n),true)

Sample Means Between Two Values

Example: Samples of 81 waist measurements are taken from a population with mean 34.2 and standard deviation 2.1.

What proportion of the sample means will between 34 and 35?

Sample Means Between Two Values

Example: Samples of 81 waist measurements are taken from a population with mean 34.2 and standard deviation 2.1.

What proportion of the sample means will between 34 and 35?

In this case the syntax for the formula is:

=NORMDIST(35,34.2,2.1/SQRT(81),true)-NORMDIST(34,34.2,2.1/SQRT(81),true)

Sample Means Between Two Values

Example: Samples of 81 waist measurements are taken from a population with mean 34.2 and standard deviation 2.1.

What proportion of the sample means will between 34 and 35?

In this case the syntax for the formula is:

=NORMDIST(35,34.2,2.1/SQRT(81),true)-NORMDIST(34,34.2,2.1/SQRT(81),true)

In this case, the result is 0.80

Sample Means Between Two Values

Example: Samples of 81 waist measurements are taken from a population with mean 34.2 and standard deviation 2.1.

What proportion of the sample means will between 34 and 35?

In this case the syntax for the formula is:

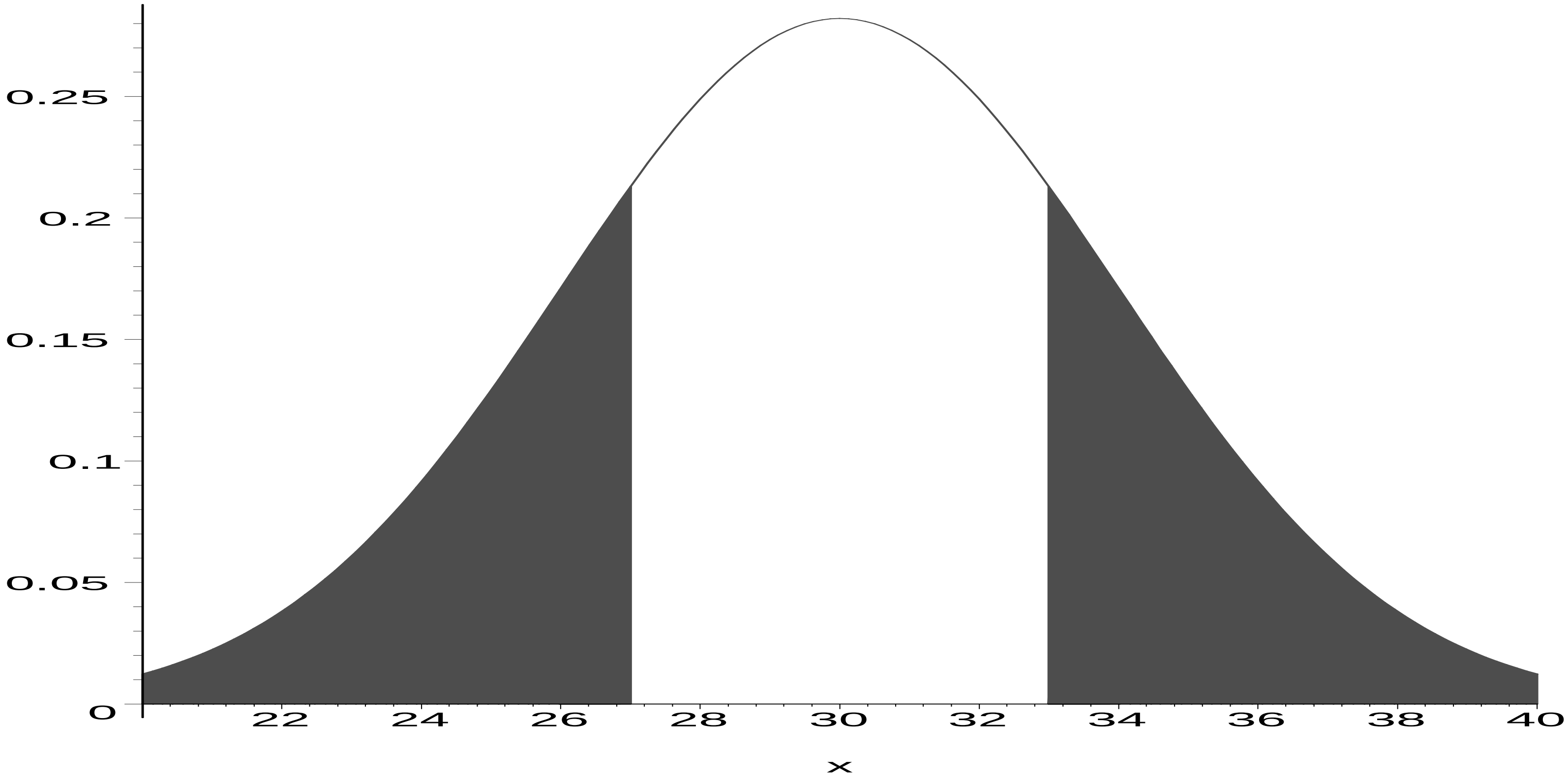
=NORMDIST(35,34.2,2.1/SQRT(81),true)-NORMDIST(34,34.2,2.1/SQRT(81),true)

In this case, the result is 0.80

This tells us 80 percent of the time, a sample of 81 from a population with mean 34.2 and standard deviation 2.1 will fall between 34 and 35.

Sample Mean Outside an Interval

The proportion of sample means below a given value **a** or above a second value **b** is:



1-(**NORMDIST(b,u,s/SQRT(n),true)-**

NORMDIST(a,u,s/SQRT(n),true))

Sample Mean Outside an Interval

Example: A sample of 80 waist measurements from a population with mean 34.2 and standard deviation 2.1 is taken.

What percentage of the time will the sample mean be either below 34 or above 34.6?

Sample Mean Outside an Interval

Example: A sample of 80 waist measurements from a population with mean 34.2 and standard deviation 2.1 is taken.

What percentage of the time will the sample mean be either below 34 or above 34.6?

In this case the syntax for the formula is:

=1-(NORMDIST(34.6,34.2,2.1/SQRT(80),true)-NORMDIST(34,34.2,2.1/SQRT(80),true))

Sample Mean Outside an Interval

Example: A sample of 80 waist measurements from a population with mean 34.2 and standard deviation 2.1 is taken.

What percentage of the time will the sample mean be either below 34 or above 34.6?

In this case the syntax for the formula is:

=1-(NORMDIST(34.6,34.2,2.1/SQRT(80),true)-NORMDIST(34,34.2,2.1/SQRT(80),true))

In this case, the result is 0.24

Sample Mean Outside an Interval

Example: A sample of 80 waist measurements from a population with mean 34.2 and standard deviation 2.1 is taken.

What percentage of the time will the sample mean be either below 34 or above 34.6?

In this case the syntax for the formula is:

=1-(NORMDIST(34.6,34.2,2.1/SQRT(80),true)-NORMDIST(34,34.2,2.1/SQRT(80),true))

In this case, the result is 0.24

This tells us that if a sample of size 80 is taken from a population with mean 34.2 and standard deviation 2.1, then 24 percent of the time the mean will fall below 34 or above 34.6.