Central Limit Theorem

Colby Community College

Definition

For all samples of the same size n with $n > 30$, the sampling distribution of $\bar{x}$ can be approximated by a normal distribution with mean μ and standard deviation $\dfrac{\sigma}{\sqrt{n}}$.

Definition

For all samples of the same size n with $n > 30$, the sampling distribution of $\bar{x}$ can be approximated by a normal distribution with mean μ and standard deviation $\dfrac{\sigma}{\sqrt{n}}$.

Note

The Central Limit Theorem holds for *any* population with *any* distribution.

Definition

For all samples of the same size n with $n > 30$, the sampling distribution of $\bar{x}$ can be approximated by a normal distribution with mean μ and standard deviation $\dfrac{\sigma}{\sqrt{n}}$.

Note

The Central Limit Theorem holds for *any* population with *any* distribution.

Note

If the population is normally distributed, then samples of any size will yield means that are normally distributed.

Definition

For all samples of the same size n with $n > 30$, the sampling distribution of $\bar{x}$ can be approximated by a normal distribution with mean μ and standard deviation $\dfrac{\sigma}{\sqrt{n}}$.

Note

The Central Limit Theorem holds for *any* population with *any* distribution.

Note

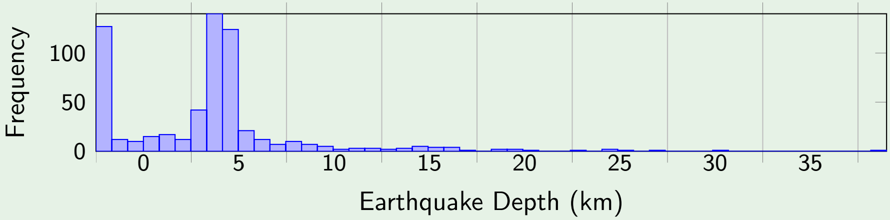
If the population is normally distributed, then samples of any size will yield means that are normally distributed.

Note

There are some rare cases where the requirement $n > 30$ isn't quite enough.

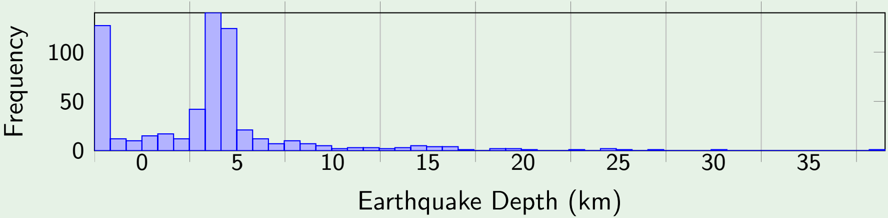
Example 1

If we look at depths of the 600 earthquakes in Data Set 21 we see that this distribution is not normal.

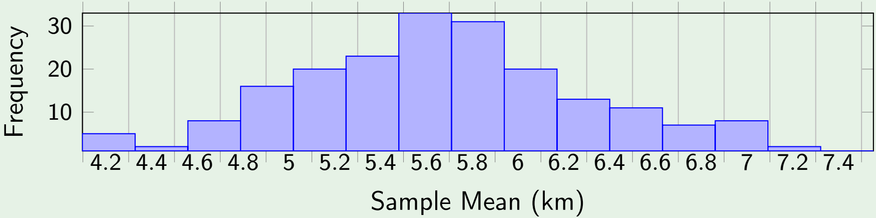


Example 1

If we look at depths of the 600 earthquakes in Data Set 21 we see that this distribution is not normal.



Looking at 200 sample means, each includes 50 randomly selected earthquake depths, we see that the distribution is approximately normal.



Universal Truth

The Central Limit Theorem describes a rule of nature that works throughout the universe.

If we could send a spaceship to a distant planet and collect samples of rocks and weight them, the sample means would have a distribution that is approximately normal.

Universal Truth

The Central Limit Theorem describes a rule of nature that works throughout the universe.

If we could send a spaceship to a distant planet and collect samples of rocks and weight them, the sample means would have a distribution that is approximately normal.

Notation

If all possible simple random samples of size n are selected from a population with mean μ and standard deviation σ, the mean of all sample means is denoted by:

Mean of all values of $\bar{x}$: $\mu_{\bar{x}} = \mu$

Standard deviation of all values of $\bar{x}$: $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

Universal Truth

The Central Limit Theorem describes a rule of nature that works throughout the universe.

If we could send a spaceship to a distant planet and collect samples of rocks and weight them, the sample means would have a distribution that is approximately normal.

Notation

If all possible simple random samples of size n are selected from a population with mean μ and standard deviation σ, the mean of all sample means is denoted by:

Mean of all values of $\bar{x}$:	$\mu_{\bar{x}} = \mu$
Standard deviation of all values of $\bar{x}$:	$\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

Note

$\sigma_{\bar{x}}$ is called the **standard error of the mean**.

Example 2

An elevator in a building has a sign stating the maximum capacity is "4000 lb—27 passengers." Because $4000/27 = 148$, this converts to a mean passenger weight of 148 lb when the elevator is full.

Example 2

An elevator in a building has a sign stating the maximum capacity is "4000 lb—27 passengers." Because $4000/27 = 148$, this converts to a mean passenger weight of 148 lb when the elevator is full.

We will assume a worst-case scenario in which the elevator is filled with 27 adult males. Based on Data Set 1, assume that adult males have weights that are normally distributed with a mean of 189 lb and a standard deviation of 39 lb.

Example 2

An elevator in a building has a sign stating the maximum capacity is "4000 lb—27 passengers." Because $4000/27 = 148$, this converts to a mean passenger weight of 148 lb when the elevator is full.

We will assume a worst-case scenario in which the elevator is filled with 27 adult males. Based on Data Set 1, assume that adult males have weights that are normally distributed with a mean of 189 lb and a standard deviation of 39 lb.

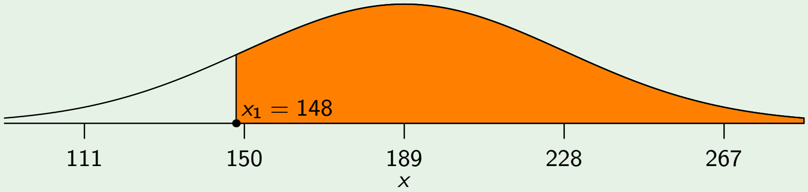
Let us find the probability that 1 randomly selected adult male has weight greater than 148 lb.

Example 2

An elevator in a building has a sign stating the maximum capacity is "4000 lb—27 passengers." Because $4000/27 = 148$, this converts to a mean passenger weight of 148 lb when the elevator is full.

We will assume a worst-case scenario in which the elevator is filled with 27 adult males. Based on Data Set 1, assume that adult males have weights that are normally distributed with a mean of 189 lb and a standard deviation of 39 lb.

Let us find the probability that 1 randomly selected adult male has weight greater than 148 lb.

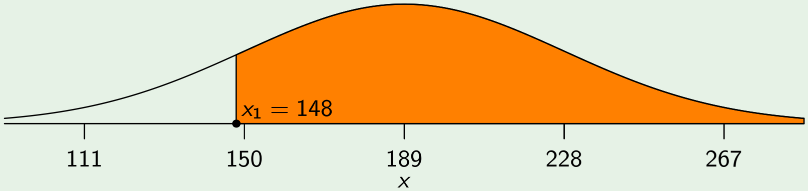


Example 2

An elevator in a building has a sign stating the maximum capacity is "4000 lb—27 passengers." Because $4000/27 = 148$, this converts to a mean passenger weight of 148 lb when the elevator is full.

We will assume a worst-case scenario in which the elevator is filled with 27 adult males. Based on Data Set 1, assume that adult males have weights that are normally distributed with a mean of 189 lb and a standard deviation of 39 lb.

Let us find the probability that 1 randomly selected adult male has weight greater than 148 lb.



Using technology we find that 0.8531 or 85.31% of males have weights greater than 148 lb.

Example 2

An elevator in a building has a sign stating the maximum capacity is "4000 lb—27 passengers." Because $4000/27 = 148$, this converts to a mean passenger weight of 148 lb when the elevator is full.

We will assume a worst-case scenario in which the elevator is filled with 27 adult males. Based on Data Set 1, assume that adult males have weights that are normally distributed with a mean of 189 lb and a standard deviation of 39 lb.

Let us find the probability that 27 randomly selected adult male has weight greater than 148 lb.

Example 2

An elevator in a building has a sign stating the maximum capacity is "4000 lb—27 passengers." Because $4000/27 = 148$, this converts to a mean passenger weight of 148 lb when the elevator is full.

We will assume a worst-case scenario in which the elevator is filled with 27 adult males. Based on Data Set 1, assume that adult males have weights that are normally distributed with a mean of 189 lb and a standard deviation of 39 lb.

Let us find the probability that 27 randomly selected adult male has weight greater than 148 lb.

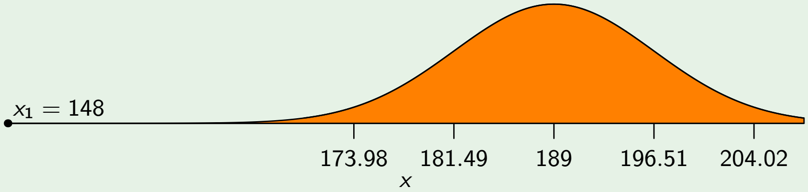
$$\mu_{\bar{x}} = \mu = 189$$
$$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 39/\sqrt{27} = 7.51$$

Example 2

An elevator in a building has a sign stating the maximum capacity is "4000 lb—27 passengers." Because $4000/27 = 148$, this converts to a mean passenger weight of 148 lb when the elevator is full.

We will assume a worst-case scenario in which the elevator is filled with 27 adult males. Based on Data Set 1, assume that adult males have weights that are normally distributed with a mean of 189 lb and a standard deviation of 39 lb.

Let us find the probability that 27 randomly selected adult male has weight greater than 148 lb.

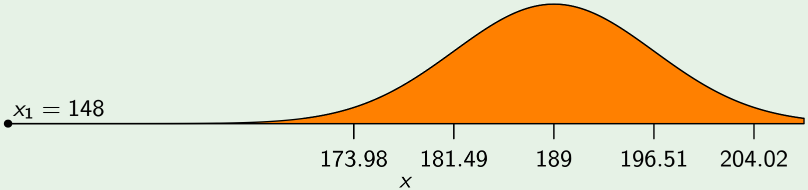


Example 2

An elevator in a building has a sign stating the maximum capacity is "4000 lb—27 passengers." Because $4000/27 = 148$, this converts to a mean passenger weight of 148 lb when the elevator is full.

We will assume a worst-case scenario in which the elevator is filled with 27 adult males. Based on Data Set 1, assume that adult males have weights that are normally distributed with a mean of 189 lb and a standard deviation of 39 lb.

Let us find the probability that 27 randomly selected adult male has weight greater than 148 lb.



Using technology we find that 0.9999 or 99.99% of the time 27 randomly selected males have a mean weight greater than 148 lb.

Rare Event Rule for Inferential Statistics

If, under a given assumption the probability of a particular observed event is very small and the observed event occurs *significantly less than* or *significantly greater than* what we typically expect with that assumption, we conclude that the assumption is probably not correct.

Example 3

Assume that the population of human body temperatures has a mean of 98.6°F, as is commonly believed. Also assume that the population standard deviation is 0.62°F.

Example 3

Assume that the population of human body temperatures has a mean of 98.6°F, as is commonly believed. Also assume that the population standard deviation is 0.62°F.

If a sample of size $n = 106$ is randomly selected, let's find the probability of getting a mean of 98.2°F or lower.

Example 3

Assume that the population of human body temperatures has a mean of 98.6°F, as is commonly believed. Also assume that the population standard deviation is 0.62°F.

If a sample of size $n = 106$ is randomly selected, let's find the probability of getting a mean of 98.2°F or lower.

$$\mu_{\bar{x}} = \mu = 98.6$$

$$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 0.62/\sqrt{106} = 0.0602197$$

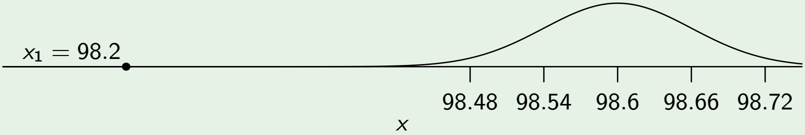
Example 3

Assume that the population of human body temperatures has a mean of 98.6°F, as is commonly believed. Also assume that the population standard deviation is 0.62°F.

If a sample of size $n = 106$ is randomly selected, let's find the probability of getting a mean of 98.2°F or lower.

$$\mu_{\bar{x}} = \mu = 98.6$$

$$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 0.62/\sqrt{106} = 0.0602197$$



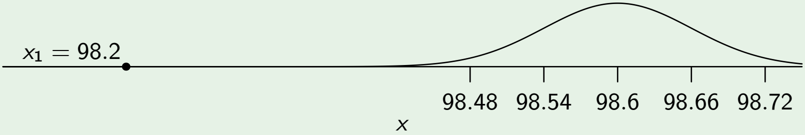
Example 3

Assume that the population of human body temperatures has a mean of 98.6°F, as is commonly believed. Also assume that the population standard deviation is 0.62°F.

If a sample of size $n = 106$ is randomly selected, let's find the probability of getting a mean of 98.2°F or lower.

$$\mu_{\bar{x}} = \mu = 98.6$$

$$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 0.62/\sqrt{106} = 0.0602197$$



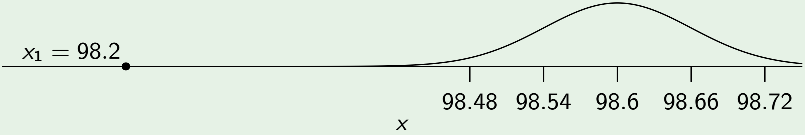
Using technology we see that the probability is 0.000000000155.

Example 3

Assume that the population of human body temperatures has a mean of 98.6°F, as is commonly believed. Also assume that the population standard deviation is 0.62°F.

If a sample of size $n = 106$ is randomly selected, let's find the probability of getting a mean of 98.2°F or lower.

$$\mu_{\bar{x}} = \mu = 98.6$$
$$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 0.62/\sqrt{106} = 0.0602197$$



Using technology we see that the probability is 0.000000000155.

If 98.6°F is the mean of our body temperature, there is a very small probability of getting a sample mean of 98.2°F.

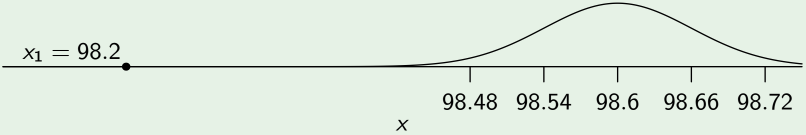
Example 3

Assume that the population of human body temperatures has a mean of 98.6°F, as is commonly believed. Also assume that the population standard deviation is 0.62°F.

If a sample of size $n = 106$ is randomly selected, let's find the probability of getting a mean of 98.2°F or lower.

$$\mu_{\bar{x}} = \mu = 98.6$$

$$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 0.62/\sqrt{106} = 0.0602197$$



Using technology we see that the probability is 0.0000000000155.

If 98.6°F is the mean of our body temperature, there is a very small probability of getting a sample mean of 98.2°F.

It is then reasonable to conclude that the population mean is lower than 98.6°F. (It's really closer to 98.2°F.)

Note

In applying the Central Limit Theorem, we assume that the population has infinitely many members. When we sample with replacement, this is effectively true.

Note

In applying the Central Limit Theorem, we assume that the population has infinitely many members. When we sample with replacement, this is effectively true.

Correction for a Finite Population

When sampling without replacement and the sample size n is greater than 5% of the finite population size N (that is $n > 0.05N$), adjust the standard deviation of sample means $\sigma_{\bar{x}}$ by multiplying it by this **finite population correction factor**:

$$\sqrt{\frac{N-n}{N-1}}$$

Note

In applying the Central Limit Theorem, we assume that the population has infinitely many members. When we sample with replacement, this is effectively true.

Correction for a Finite Population

When sampling without replacement and the sample size n is greater than 5% of the finite population size N (that is $n > 0.05N$), adjust the standard deviation of sample means $\sigma_{\bar{x}}$ by multiplying it by this **finite population correction factor**:

$$\sqrt{\frac{N-n}{N-1}}$$

Note

You *do not* use the finite population correction factor when:

- You sample with replacement.
- The population if infinite.
- Sample size does not exceed 5% of the population size.