

CENTRAL LIMIT THEOREM: CENTRAL LIMIT THEOREM FOR SAMPLE MEANS*

Susan Dean
Barbara Illowsky, Ph.D.

This work is produced by OpenStax-CNX and licensed under the
Creative Commons Attribution License 3.0[†]

Suppose X is a random variable with a distribution that may be known or unknown (it can be any distribution). Using a subscript that matches the random variable, suppose:

- a. μ_X = the mean of X
- b. σ_X = the standard deviation of X

If you draw random samples of size n , then as n increases, the random variable \bar{X} which consists of sample means, tends to be **normally distributed** and

$$\bar{X} \sim N\left(\mu_X, \frac{\sigma_X}{\sqrt{n}}\right)$$

The Central Limit Theorem for Sample Means says that if you keep drawing larger and larger samples (like rolling 1, 2, 5, and, finally, 10 dice) and **calculating their means** the sample means form their own **normal distribution** (the sampling distribution). The normal distribution has the same mean as the original distribution and a variance that equals the original variance divided by n , the sample size. n is the number of values that are averaged together not the number of times the experiment is done.

To put it more formally, if you draw random samples of size n , the distribution of the random variable \bar{X} , which consists of sample means, is called the **sampling distribution of the mean**. The sampling distribution of the mean approaches a normal distribution as n , the sample size, increases.

The random variable \bar{X} has a different z-score associated with it than the random variable X . \bar{x} is the value of \bar{X} in one sample.

$$z = \frac{\bar{x} - \mu_X}{\left(\frac{\sigma_X}{\sqrt{n}}\right)} \quad (1)$$

μ_X is both the average of X and of \bar{X} .

$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}}$ = standard deviation of \bar{X} and is called the **standard error of the mean**.

Example 1

An unknown distribution has a mean of 90 and a standard deviation of 15. Samples of size $n = 25$ are drawn randomly from the population.

Problem 1

Find the probability that the **sample mean** is between 85 and 92.

*Version 1.23: Jun 18, 2012 9:49 am -0500

[†]<http://creativecommons.org/licenses/by/3.0/>

Solution

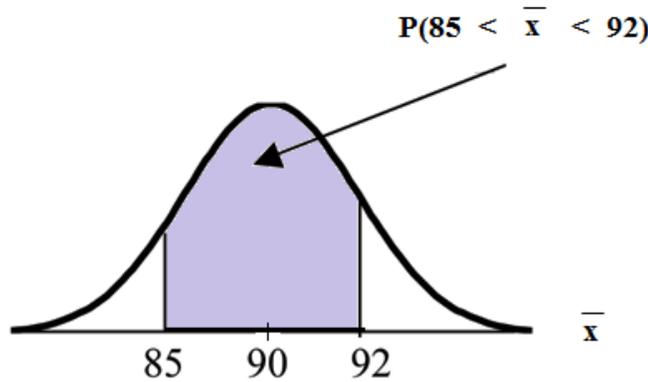
Let X = one value from the original unknown population. The probability question asks you to find a probability for the **sample mean**.

Let \bar{X} = the mean of a sample of size 25. Since $\mu_X = 90$, $\sigma_X = 15$, and $n = 25$;
then $\bar{X} \sim N\left(90, \frac{15}{\sqrt{25}}\right)$

Find $P(85 < \bar{x} < 92)$ Draw a graph.

$$P(85 < \bar{x} < 92) = 0.6997$$

The probability that the sample mean is between 85 and 92 is 0.6997.



TI-83 or 84: normalcdf(lower value, upper value, mean, standard error of the mean)

The parameter list is abbreviated (lower value, upper value, μ , $\frac{\sigma}{\sqrt{n}}$)

$$\text{normalcdf}\left(85, 92, 90, \frac{15}{\sqrt{25}}\right) = 0.6997$$

Problem 2

Find the value that is 2 standard deviations above the expected value (it is 90) of the sample mean.

Solution

To find the value that is 2 standard deviations above the expected value 90, use the formula

$$\text{value} = \mu_X + (\# \text{of STDEVs}) \left(\frac{\sigma_X}{\sqrt{n}}\right)$$

$$\text{value} = 90 + 2 \cdot \frac{15}{\sqrt{25}} = 96$$

So, the value that is 2 standard deviations above the expected value is 96.

Example 2

The length of time, in hours, it takes an "over 40" group of people to play one soccer match is normally distributed with a **mean of 2 hours** and a **standard deviation of 0.5 hours**. A **sample of size $n = 50$** is drawn randomly from the population.

Problem

Find the probability that the **sample mean** is between 1.8 hours and 2.3 hours.

Solution

Let X = the time, in hours, it takes to play one soccer match.

The probability question asks you to find a probability for the **sample mean time, in hours**, it takes to play one soccer match.

Let \bar{X} = the **mean** time, in hours, it takes to play one soccer match.

If $\mu_X = \underline{\hspace{2cm}}$, $\sigma_X = \underline{\hspace{2cm}}$, and $n = \underline{\hspace{2cm}}$, then $\bar{X} \sim N(\underline{\hspace{2cm}}, \underline{\hspace{2cm}})$ by the Central Limit Theorem for Means.

$\mu_X = 2$, $\sigma_X = 0.5$, $n = 50$, and $X \sim N\left(2, \frac{0.5}{\sqrt{50}}\right)$

Find $P(1.8 < \bar{x} < 2.3)$. Draw a graph.

$P(1.8 < \bar{x} < 2.3) = 0.9977$

$\text{normalcdf}\left(1.8, 2.3, 2, \frac{.5}{\sqrt{50}}\right) = 0.9977$

The probability that the mean time is between 1.8 hours and 2.3 hours is $\underline{\hspace{2cm}}$.

Glossary

Definition 1: Average

A number that describes the central tendency of the data. There are a number of specialized averages, including the arithmetic mean, weighted mean, median, mode, and geometric mean.

Definition 2: Central Limit Theorem

Given a random variable (RV) with known mean μ and known standard deviation σ . We are sampling with size n and we are interested in two new RVs - the sample mean, \bar{X} , and the sample sum, ΣX . If the size n of the sample is sufficiently large, then $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ and $\Sigma X \sim N(n\mu, \sqrt{n}\sigma)$. If the size n of the sample is sufficiently large, then the distribution of the sample means and the distribution of the sample sums will approximate a normal distribution regardless of the shape of the population. The mean of the sample means will equal the population mean and the mean of the sample sums will equal n times the population mean. The standard deviation of the distribution of the sample means, $\frac{\sigma}{\sqrt{n}}$, is called the standard error of the mean.

Definition 3: Normal Distribution

A continuous random variable (RV) with pdf $f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$, where μ is the mean of the distribution and σ is the standard deviation. Notation: $X \sim N(\mu, \sigma)$. If $\mu = 0$ and $\sigma = 1$, the RV is called **the standard normal distribution**.

Definition 4: Standard Error of the Mean

The standard deviation of the distribution of the sample means, $\frac{\sigma}{\sqrt{n}}$.