

5.1: Introduction to the Central Limit Theorem

Why are we so concerned with means? Two reasons are: they give us a middle ground for comparison, and they are easy to calculate. In this chapter, you will study means and the **Central Limit Theorem**.

The **Central Limit Theorem** is one of the most powerful and useful ideas in all of statistics. The Central Limit Theorem is a theorem which means that it is NOT a theory or just somebody's idea of the way things work. As a theorem it ranks with the Pythagorean Theorem, or the theorem that tells us that the sum of the angles of a triangle must add to 180. These are facts of the ways of the world rigorously demonstrated with mathematical precision and logic. As we will see this powerful theorem will determine just what we can, and cannot say, in inferential statistics. The Central Limit Theorem is concerned with drawing finite samples of size n from a population with a known mean, μ , and a known standard deviation, σ . The conclusion is that if we collect samples of size n with a "large enough n ," calculate each sample's mean, and create a histogram (distribution) of those means, then the resulting distribution will tend to have an approximate normal distribution.

The astounding result is that it does not matter what the distribution of the original population is, or whether you even need to know it. The important fact is that the distribution of sample means tend to follow the normal distribution.


 This is a photo of change a set of keys in a pile. There appear to be five pennies, three quarters, four dimes, and two nickels. The key ring has a bronze whale on it and holds eleven keys.

Figure 5.1.1 If you want to figure out the distribution of the change people carry in their pockets, using the Central Limit Theorem and assuming your sample is large enough, you will find that the distribution is normal. (credit: John Lodder)

The size of the sample, n , that is required in order to be "large enough" depends on the original population from which the samples are drawn (the sample size should be at least 30, or ideally 100 or more). If the original population is far from normal, then more observations are needed for the sample means. Sampling is done randomly and with replacement in the theoretical model.

This page titled [5.1: Introduction to the Central Limit Theorem](#) is shared under a [CC BY](#) license and was authored, remixed, and/or curated by .