

Central Limit Theorem

Suppose X_1, X_2, \dots is an infinite sequence of independent random variables (not necessarily Normal) that each have mean μ and variance σ^2 . Fix "a" then the probability that the sum of the first n of the random variables, suitably scaled and shifted, is less than a will go to the same probability that a standard Normal random variable is less than a (as $n \rightarrow \infty$).

$$\lim P\left(\frac{X_1 + \dots + X_n - n\mu}{\sqrt{n\sigma^2}} \leq a\right) = \int_{-\infty}^a \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = F_z(a)$$

The Central Limit Theorem is a limiting result. In practice, we often apply it to finite but large collections of independent random variables, to model their sum by a standard Normal random variable.

Two caveats: The CLT works better and better, the larger n is, and there is no cutoff for which it works completely above some particular n .

We must be sure to get the scaling and shifting correct, if we are going to apply the CLT properly.