The CLT is a general result for sums of random variables. In statistics, it helps to study the distribution of the sample mean. Let $X_1, X_2, \ldots, X_n$ be now n independent and identically distributed (i.i.d.) random variables from any probability distribution with mean μ and variance σ^2 both finite.

$\lim_{n\to\infty} \frac{S_n - n\mu}{\sqrt{n}\sigma} \sim N(0,1)$. Can also be written as:

$\lim_{n\to\infty} \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1), \overline{X} = \frac{S_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$ is the sample mean.

Implications for the Sample Mean: The CLT thus implies that for large, but finite n, $\overline{X} \approx N(\mu, \frac{\sigma^2}{n})$

- A rule of thumb **for 'large n" is often n ≥ 30.**
- Amazingly, this result holds irrespective of the distribution of the $\{X_i\}$ (and including discrete random variables). CLT thus demonstrates that statistical regularity can arise even from the combination of highly-diverse random phenomena.
- If $X_i \sim N(\mu, \sigma^2), \forall i$, the result becomes exact even for finite n, since the sum of independent normal random variables is normally distributed.