Central Limit Theorem Worksheet

The objective of this worksheet is to help you use CLT simulation in order to understand the theory.

Table of Contents

Definition of Central Limit Theorem:

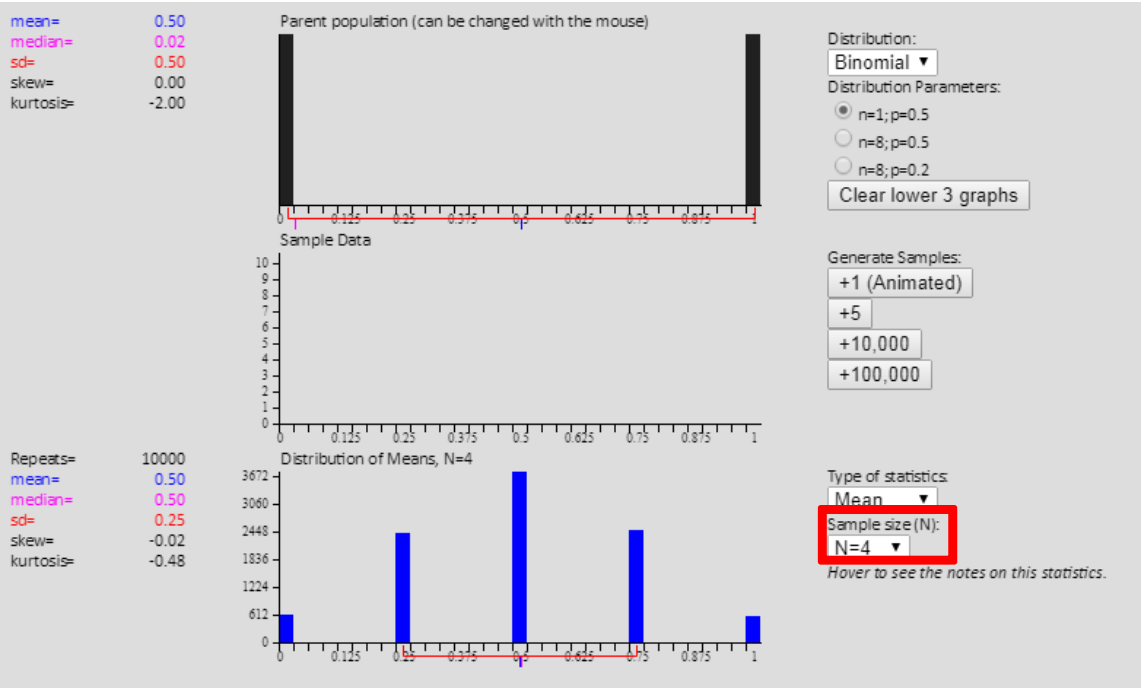
For a large number of identically independently distributed random variables each with finite mean μ and finite variance σ^2, the sample mean will be approximately normally distributed with mean μ and variance σ^2/n, regardless of the population distribution.

Notes for definition:

1. Sample size n should be sufficiently large enough.
2. (Won't test these but these are conditions) "identically" "independently" "random".
3. If population mean is μ with variance σ^2 then mean of sample means is also μ with variance σ^2/n , and $\sigma/\sqrt{n}$ is called sample standard error, and it affects the "spread" of the graph.
4. No matter what population distribution is, if above conditions are met, then CLT holds.

Overview of CLT simulation

The whole idea of this simulation is to show you, when sample size is small (N=4), the sampling distribution of sample mean (the blue figure) is discrete.



When sample size is sufficient (N=32), the sampling distribution of sample mean approximates to normal, regardless of population distribution (here is a Bernoulli

distribution, or a Binomial distribution with n=1 and p=0.5).

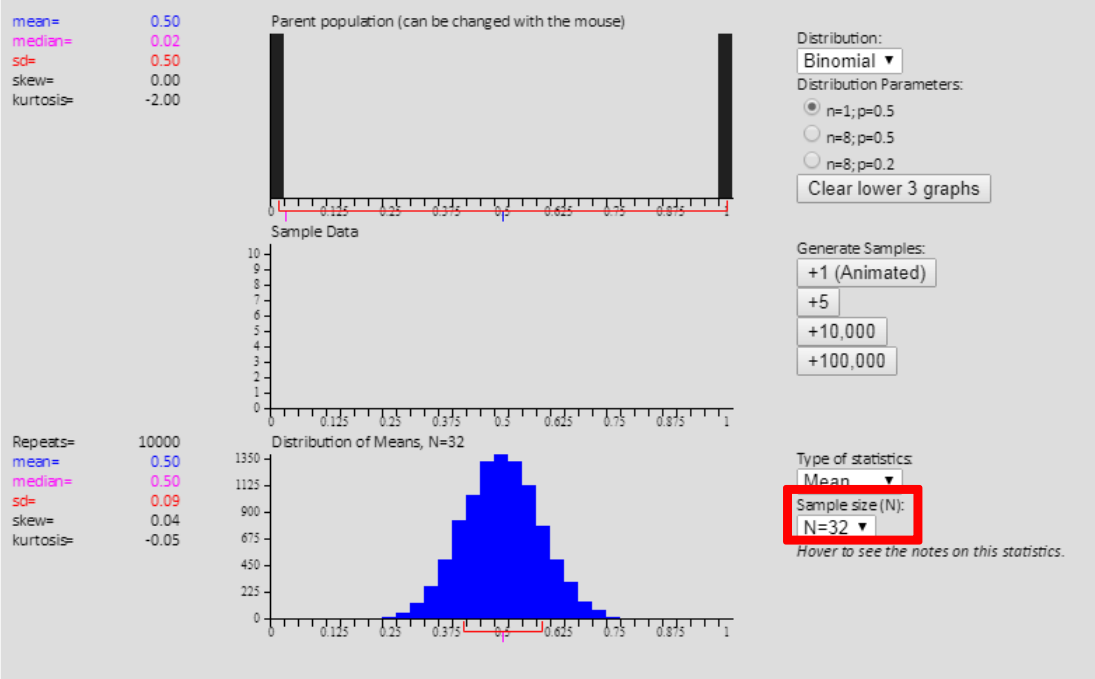
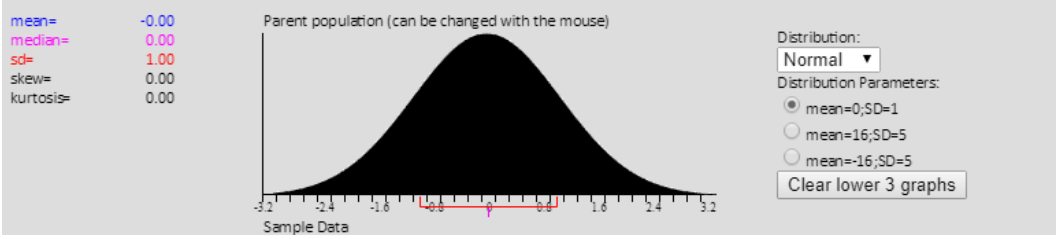
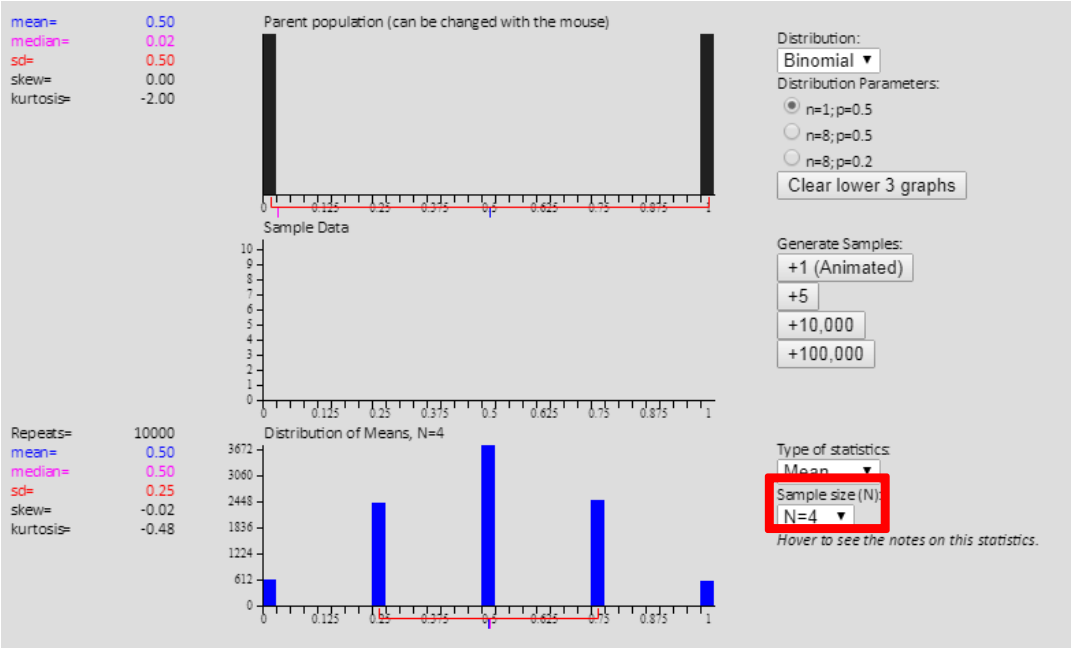


Illustration of CLT

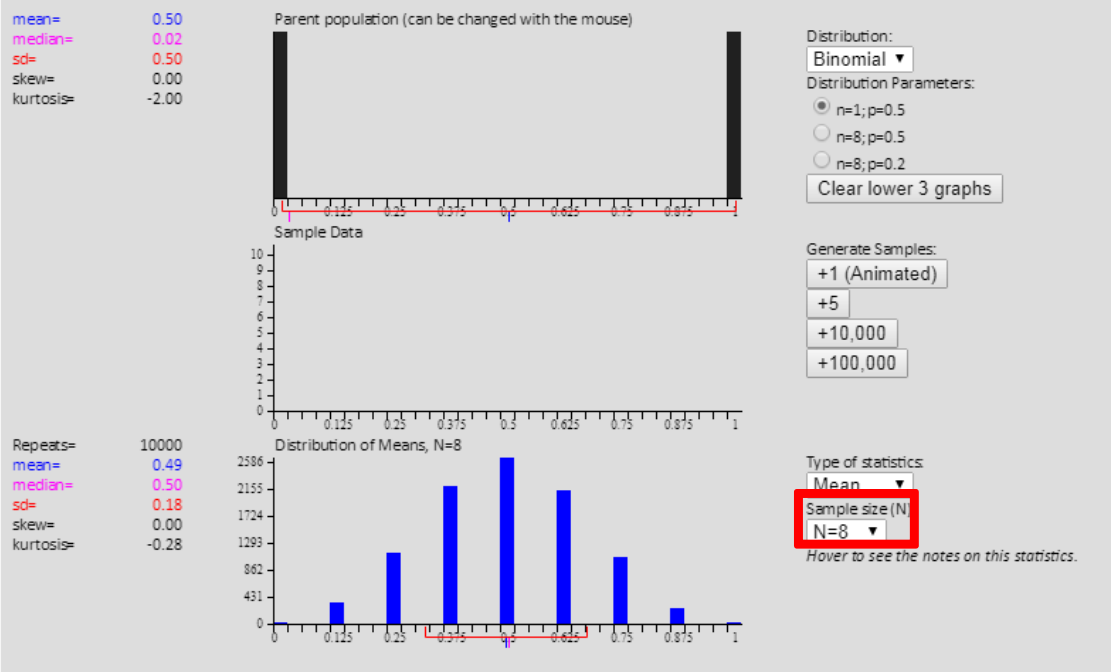
Let's recap what a Normal distribution looks like.



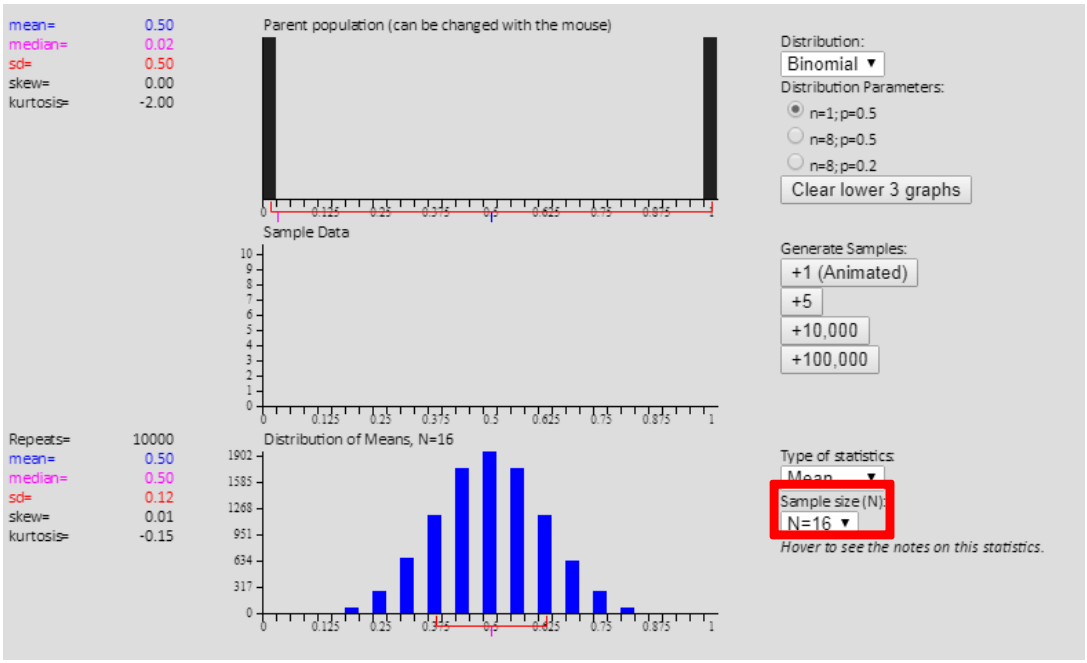
As we can see, a Normal distribution is a continuous distribution, and it is bell shaped with mean equals to median which lies in the middle.



Above is a sampling distribution of sample mean based on a Binomial distribution with n=1 and p=0.5. When N=4 (here N represents sample size), sampling distribution looks discrete and non-Normal (separated blue bars).

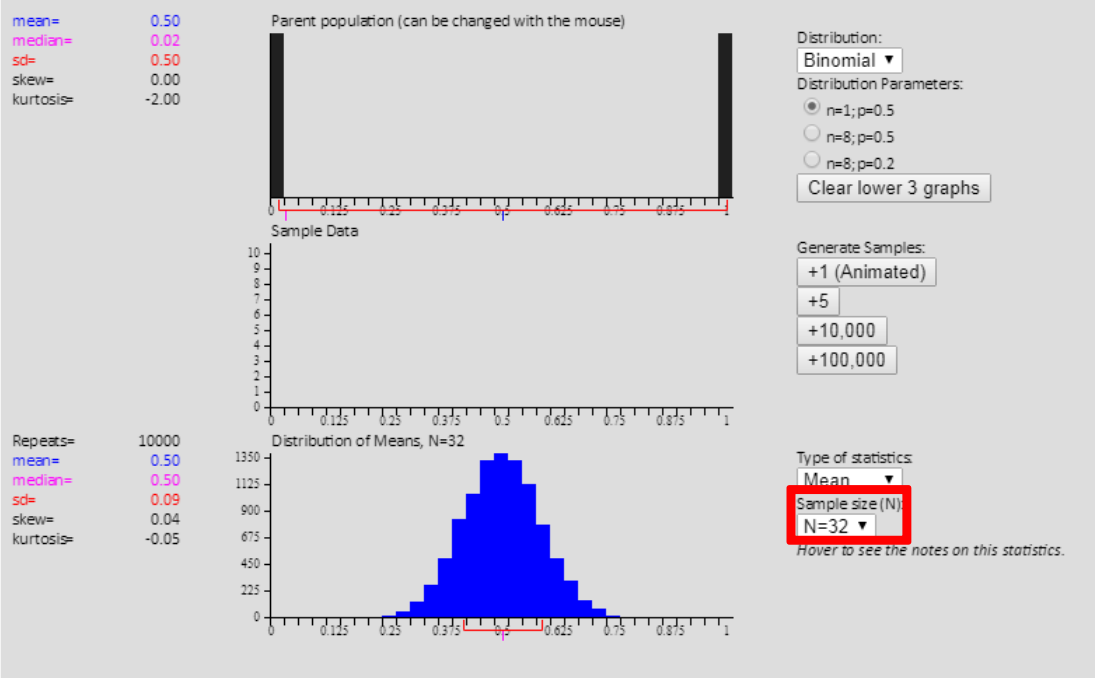


When N=8, sampling distribution looks less discrete than N=4 but still non-Normal



When N=16, sampling distribution starts to look close to a Normal distribution

[Simulation] Central Limit Theorem Simulation with notes



When N=32, sampling distribution approximates to a Normal distribution although the population distribution is not normal. We generally consider n>30 as sufficient sample size.