

---

This tutorial talks about the central limit theorem. The central limit theorem tells us that under certain conditions, the mean of repeated samples tends to be normally distributed. Now this explains why so many real world processes are approximately normally distributed.

There are three conditions for the central limit theorem. One is that the sample size needs to be large enough, typically more than 30. Second, smaller samples will still work if the underlying distribution isn't skewed. So if the underlying distribution is fairly normal, then the smaller sample size is still going to be OK, even if that's under 30.

And then thirdly, the population standard deviation needs to be finite. For example, if we're counting islands, you would need to set some minimum size to what we define to be an island. Otherwise, depending on how we count, we can keep getting larger and larger numbers of islands, and then we wouldn't have a well-defined mean or standard deviation.

Now with the central limit theorem, in practice, getting a sample size of 30 or more is what statisticians want. So this first condition, that the sample size be large enough more than 30, is typically what most statisticians look for. However, those smaller sample sizes, again, can still work.

Now with the central limit theorem, that's what's letting us know why so many real world processes are approximately normally distributed. This has been your tutorial on the central limit theorem.