

7.3: Using the Central Limit Theorem

The **Law of Large Numbers**, along with the Central Limit Theorem, provides another critical piece of information to allow us to engage in inferential statistics. In short, the Law of Large Numbers proves that the expected value of the sampling distribution of the sample mean is the population mean:

$$E(\mu_{\bar{x}}) = \mu \quad (7.3.1)$$

The proof is through the concept of large numbers.

Suppose you were to take a sample and calculate a sample mean. Then you take another sample, combine it with the previous sample, and calculate the sample mean of the combined sample. Then you repeat this process over and over, creating bigger and bigger samples and calculating a sample mean each time along the way. The sample means from larger and larger samples will get closer and closer to the population mean, μ . Figure 7.3 shows the running average as more sample means are added and then averaged. The proof of the Law of Large Numbers mathematically was perfected during a period of 20 years and was presented by Jacob Bernoulli in 1713.

LLN : as $n \rightarrow \infty$ the running average of $\bar{X}'s \rightarrow \mu$

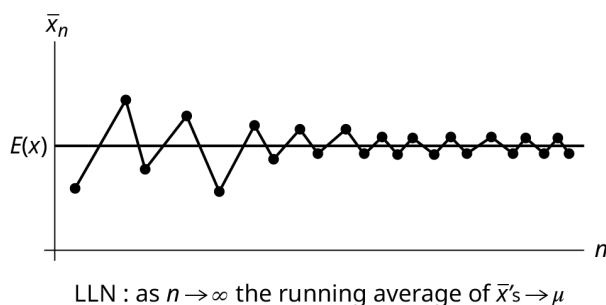


Figure 7.3.1

Stated mathematically

$$\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| > \varepsilon) = 0 \quad (7.3.2)$$

or alternatively presented as

$$P\left(\lim_{n \rightarrow \infty} |\bar{X}_n - \mu| < \varepsilon\right) = 1 \quad (7.3.3)$$

where \bar{X}_n is the running average as additional sample means are added to the previous sample means.

In summary:

There are three critical mathematical conclusions that flow from the Central Limit Theorem and the application of the Law of Large Numbers.

1. By the Central Limit Theorem, for large enough sample sizes, the sampling distribution of sample means tends to be normally distributed regardless of the underlying distribution of the population data.
2. As the sample size, n , gets larger and larger, the sampling distribution standard deviation gets smaller. Remember that the standard deviation for the sampling distribution of \bar{X}_n is $\frac{\sigma}{\sqrt{n}}$. The sample mean, \bar{x} , is more likely to be closer to μ as n increases.
3. By the Law of Large Numbers, the expected value of the sampling distribution of the sample mean is the population mean.

Law of Large Numbers

The law of large numbers says that if you take samples of larger and larger size from any population, then the mean of the sampling distribution, $\mu_{\bar{x}}$ tends to get closer and closer to the true population mean, μ . From the Central Limit Theorem, we know that as n gets larger and larger, the sample means follow a normal distribution. The larger n gets, the smaller the standard deviation of the sampling distribution gets. (Remember that the standard deviation for the sampling distribution of \bar{X} is $\frac{\sigma}{\sqrt{n}}$.) This means

that the sample mean \bar{x} must be closer to the population mean μ as n increases. We can say that μ is the value that the sample means approach as n gets larger. The Central Limit Theorem illustrates the law of large numbers.

Examples of the Central Limit Theorem

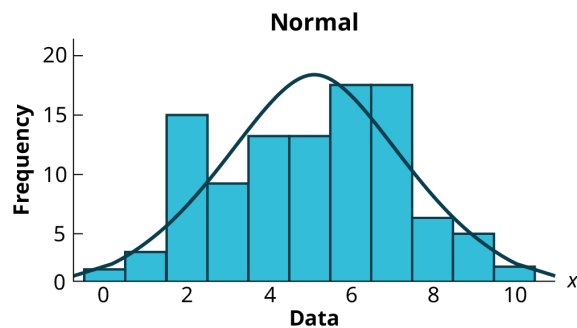
This concept is so important and plays such a critical role in what follows it deserves to be developed further. Indeed, there are two critical issues that flow from the Central Limit Theorem and the application of the Law of Large numbers to it. These are

1. The probability density function of the sampling distribution of means is normally distributed **regardless** of the underlying distribution of the population observations and
2. standard deviation of the sampling distribution decreases as the size of the samples that were used to calculate the means for the sampling distribution increases.

Taking these in order. It would seem counterintuitive that the population may have **any** distribution and the distribution of means coming from it would be normally distributed. With the use of computers, experiments can be simulated that show the process by which the sampling distribution changes as the sample size is increased. These simulations show visually the results of the mathematical proof of the Central Limit Theorem.

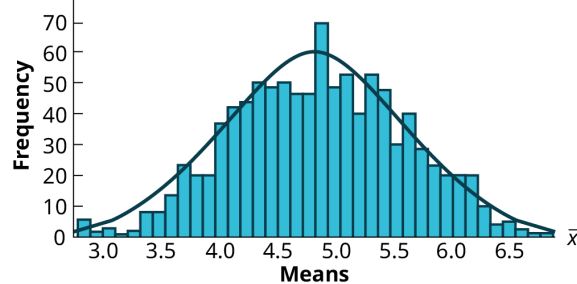
Here are three examples of very different population distributions and the evolution of the sampling distribution to a normal distribution as the sample size increases. The top panel in these cases represents the histogram for the original data. The three panels show the histograms for 1,000 randomly drawn samples for different sample sizes: $n = 10$, $n = 25$ and $n = 50$. As the sample size increases, and the number of samples taken remains constant, the distribution of the 1,000 sample means becomes closer to the smooth line that represents the normal distribution.

Figure 7.3.2 is for a normal distribution of individual observations and we would expect the sampling distribution to converge on the normal quickly. The results show this and show that even at a very small sample size the distribution is close to the normal distribution.



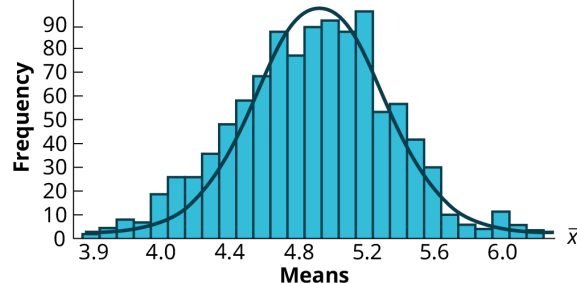
Sample Size $n = 10$

Histogram (with Normal Curve) of Mean



Sample Size $n = 25$

Histogram (with Normal Curve) of Mean



Sample Size $n = 50$

Histogram (with Normal Curve) of Mean

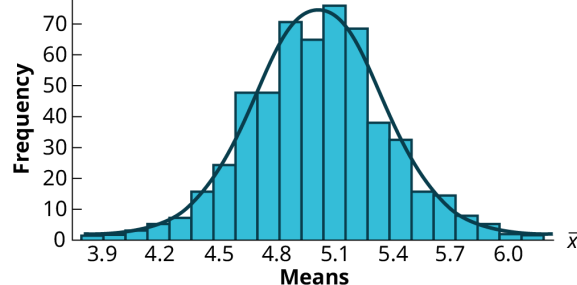
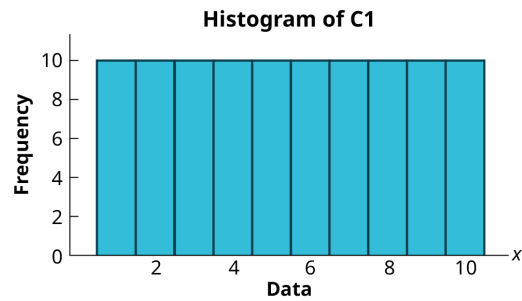


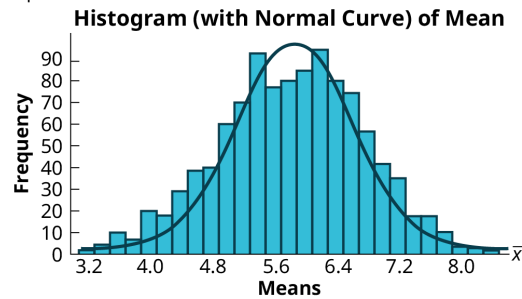
Figure 7.3.2: Copy and Paste Caption here. (Copyright; author via source)

Figure 7.3.3 is a uniform distribution which, a bit amazingly, quickly approached the normal distribution even with only a sample of 10.

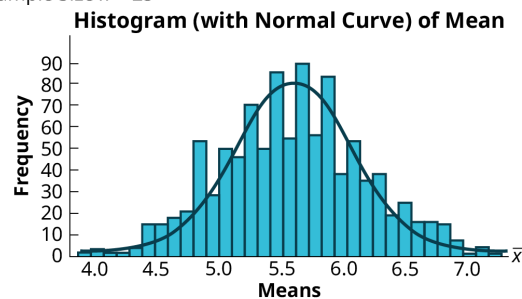
Distribution of Random Variable



Sample Size $n = 10$



Sample Size $n = 25$



Sample Size $n = 50$

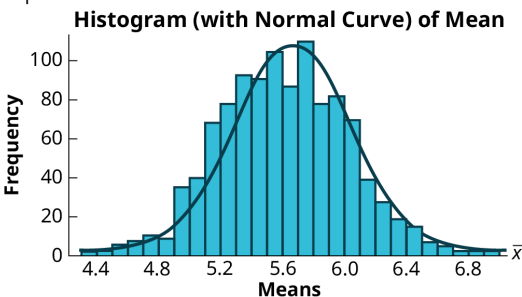
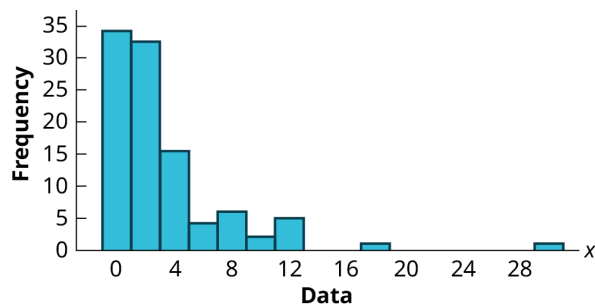


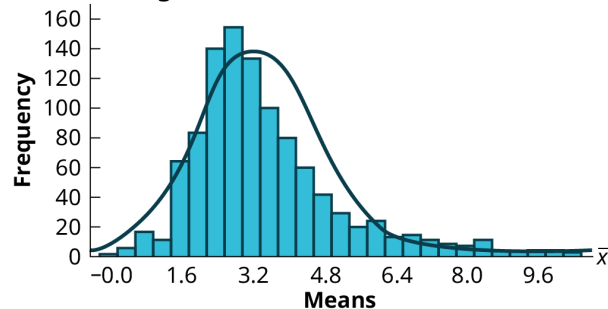
Figure 7.3.3: Copy and Paste Caption here. (Copyright; author via source)

Figure 7.3.4 is a skewed distribution. This last one could be an exponential, geometric, or binomial with a small probability of success creating the skew in the distribution. For skewed distributions our intuition would say that this will take larger sample sizes to move to a normal distribution and indeed that is what we observe from the simulation. Nevertheless, at a sample size of 50, not considered a very large sample, the distribution of sample means has very decidedly gained the shape of the normal distribution.



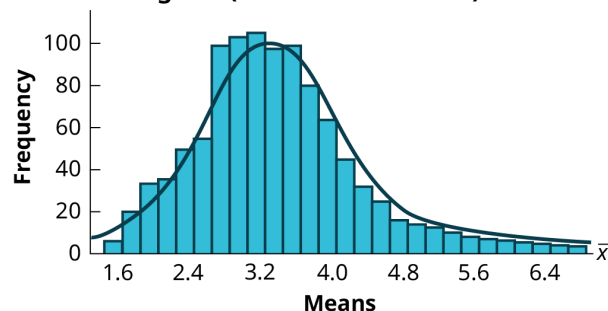
Distribution of Sample means with $n = 10$

Histogram (with Normal Curve) of Mean



Distribution of Sample means with $n = 25$

Histogram (with Normal Curve) of Mean



Distribution of Sample means with $n = 50$

Histogram (with Normal Curve) of Mean

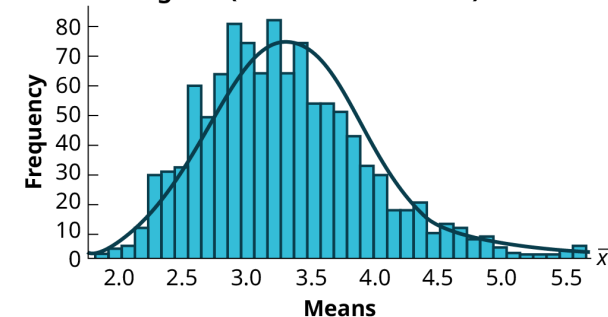


Figure 7.3.4

The Central Limit Theorem provides more than the proof that the sampling distribution of means is normally distributed. It also provides us with the mean and standard deviation of this distribution. Further, as discussed above, the expected value of the mean, $\mu_{\bar{x}}$, is equal to the mean of the population of the original data which is what we are interested in estimating from the sample we took. We have already inserted this conclusion of the Central Limit Theorem into the formula we use for standardizing from the sampling distribution to the standard normal distribution. And finally, the Central Limit Theorem has also provided the standard deviation of the sampling distribution, $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, and this is critical to have to calculate probabilities of values of the new random variable, \bar{x} .

Figure 7.3.5 shows a sampling distribution. The mean has been marked on the horizontal axis of the \bar{x} 's and the standard deviation has been written to the right above the distribution. Notice that the standard deviation of the sampling distribution is the original standard deviation of the population, divided by the sample size. We have already seen that as the sample size increases the sampling distribution becomes closer and closer to the normal distribution. As this happens, the standard deviation of the sampling distribution changes in another way; the standard deviation decreases as n increases. At very large n , the standard deviation of the sampling distribution becomes very small and at infinity it collapses on top of the population mean. This is what it means that the expected value of $\mu_{\bar{x}}$ is the population mean, μ .

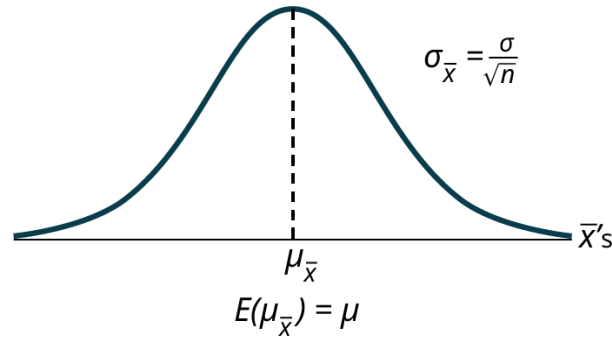


Figure 7.3.5

At non-extreme values of n , this relationship between the standard deviation of the sampling distribution and the sample size plays a very important part in our ability to estimate the parameters we are interested in.

Figure 7.3.6 shows three sampling distributions. The only change that was made is the sample size that was used to get the sample means for each distribution. As the sample size increases, n goes from 10 to 30 to 50, the standard deviations of the respective sampling distributions decrease because the sample size is in the denominator of the standard deviations of the sampling distributions.

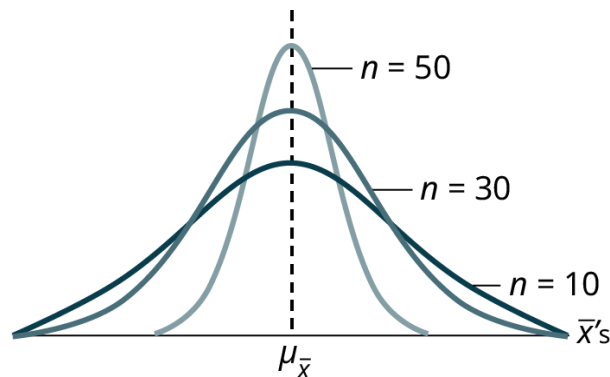


Figure 7.3.6: Copy and Paste Caption here. (Copyright; author via source)

The implications for this are very important. Figure 7.3.7 shows the effect of the sample size on the confidence we will have in our estimates. These are two sampling distributions from the same population. One sampling distribution was created with samples of size 10 and the other with samples of size 50. All other things constant, the sampling distribution with sample size 50 has a smaller standard deviation that causes the graph to be higher and narrower. The important effect of this is that for the same probability of one standard deviation from the mean, this distribution covers much less of a range of possible values than the other distribution. One standard deviation is marked on the \bar{X} axis for each distribution. This is shown by the two arrows that are plus or minus one standard deviation for each distribution. If the probability that the true mean is one standard deviation away from the mean, then for the sampling distribution with the smaller sample size, the possible range of values is much greater. A simple question is, would you rather have a sample mean from the narrow, tight distribution, or the flat, wide distribution as the estimate of the population mean? Your answer tells us why people intuitively will always choose data from a large sample rather than a small sample. The sample mean they are getting is coming from a more compact distribution. This concept will be the foundation for what will be called level of confidence in the next unit.

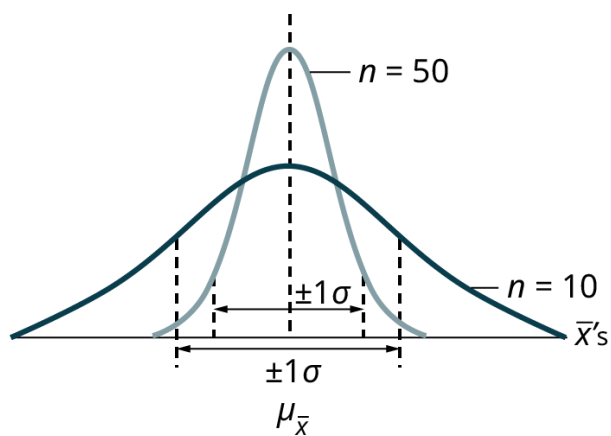


Figure 7.3.7

This page titled [7.3: Using the Central Limit Theorem](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [OpenStax](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.