

### 3.11: The Variance of the Average- The Central Limit Theorem

The **central limit theorem** establishes very important relationships between the statistics for two distributions that are related in a particular way. It enables us to understand some important features of physical systems.

The central limit theorem concerns the distribution of averages. If we have some original distribution and sample it three times, we can calculate the average of these three data points. Call this average  $A_{3,1}$ . We could repeat this activity and obtain a second average of three values,  $A_{3,2}$ . We can do this repeatedly, generating averages  $A_{3,3}, \dots, A_{3,n}$ . Several things will be true about these averages:

- The set of all of the possible averages-of-three,  $\{A_{3,i}\}$ , is itself a distribution. This averages-of-three distribution is different from the original distribution. Each average-of-three is a value of the random variable associated with the averages-of-three distribution.
- Each of the  $A_{3,i}$  is an estimate of the mean of the original distribution.
- The distribution of the  $A_{3,i}$  will be less spread out than the original distribution.

There is nothing unique about averaging three values. We could sample the original distribution seven times and compute the average of these seven values, calling the result  $A_{7,1}$ . Repeating, we could generate averages  $A_{7,2}, \dots, A_{7,m}$ . All of the things we say about the averages-of-three are also true of these averages-of-seven. However, we can now say something more: The distribution of the  $A_{7,i}$  will be less spread out than the distribution of the  $A_{3,i}$ . The corresponding probability density functions are sketched in Figure 13.

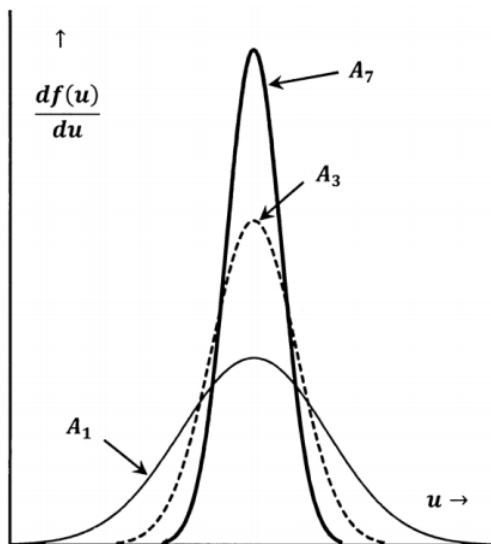


Figure 13. The Variance of an Average of  $N$  is proportional to  $1/N$ .

The central limit theorem relates the mean and variance of the distribution of averages to the mean and variance of the original distribution:

***If random samples of  $N$  values are taken from a distribution, whose mean is  $\mu$  and whose variance is  $\sigma^2$ , averages of these  $N$  values,  $A_N$ , are approximately normally distributed with a mean of  $\mu$  and a variance of  $\sigma^2/N$ . The approximation to the normal distribution becomes better as  $N$  becomes larger.***

It turns out that the number,  $N$ , of trials that is needed to get a good estimate of the variance is substantially larger than the number required to get a good estimate of the mean.

This page titled [3.11: The Variance of the Average- The Central Limit Theorem](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Paul Ellgen](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.