

9.3: Central Limit Theorem for Continuous Independent Trials

We have seen in Section 1.2 that the distribution function for the sum of a large number n of independent discrete random variables with mean μ and variance σ^2 tends to look like a normal density with mean $n\mu$ and variance $n\sigma^2$. What is remarkable about this result is that it holds for *any* distribution with finite mean and variance. We shall see in this section that the same result also holds true for continuous random variables having a common density function.

Let us begin by looking at some examples to see whether such a result is even plausible.

Standardized Sums

✓ Example 9.3.1

Suppose we choose n random numbers from the interval $[0, 1]$ with uniform density. Let X_1, X_2, \dots, X_n denote these choices, and $S_n = X_1 + X_2 + \dots + X_n$ their sum.

We saw in Example [7.2.12] that the density function for S_n tends to have a normal shape, but is centered at $n/2$ and is flattened out. In order to compare the shapes of these density functions for different values of n , we proceed as in the previous section: we S_n by defining

$$S_n^* = \frac{S_n - n\mu}{\sqrt{n}\sigma}. \quad (9.3.1)$$

Then we see that for all n we have

$$\begin{aligned} E(S_n^*) &= 0, \\ V(S_n^*) &= 1. \end{aligned}$$

The density function for S_n^* is just a standardized version of the density function for S_n (see Figure [fig 9.7]).

✓ Example 9.3.1

Let us do the same thing, but now choose numbers from the interval $[0, +\infty)$ with an exponential density with parameter λ . Then (see Example 6.21)

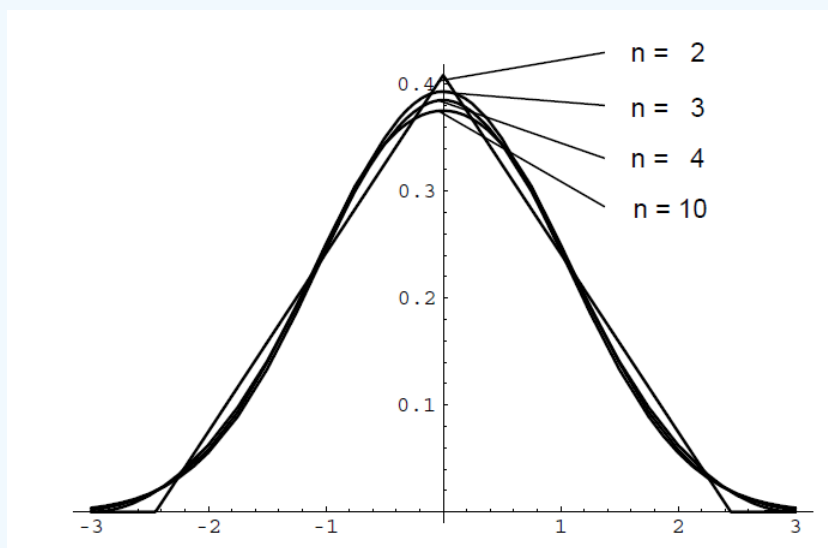


Figure 9.3.1: Density function for S_n^* (uniform case, $n=2,3,4,10$).

$$\mu = E(X_i) = \frac{1}{\lambda},$$

$$\sigma^2 = V(X_j) = \frac{1}{\lambda^2}.$$

Here we know the density function for S_n explicitly (see Section [sec 7.2]). We can use Corollary [cor 5.1] to calculate the density function for S_n^* . We obtain

$$f_{S_n}(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{n-1}}{(n-1)!},$$

$$f_{S_n^*}(x) = \frac{\sqrt{n}}{\lambda} f_{S_n}\left(\frac{\sqrt{n}x + n}{\lambda}\right).$$

The graph of the density function for S_n^* is shown in Figure [fig 9.9].

These examples make it seem plausible that the density function for the normalized random variable S_n^* for large n will look very much like the normal density with mean 0 and variance 1 in the continuous case as well as in the discrete case. The Central Limit Theorem makes this statement precise.

Central Limit Theorem

Theorem Central Limit Theorem 9.3.1

(Let $S_n = X_1 + X_2 + \cdots + X_n$ be the sum of n independent continuous random variables with common density function p having expected value μ and variance σ^2 . Let $S_n^* = (S_n - n\mu)/\sqrt{n}\sigma$. Then we have, for all $a < b$,

$$\lim_{n \rightarrow \infty} P(a < S_n^* < b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx. \quad (9.3.2)$$

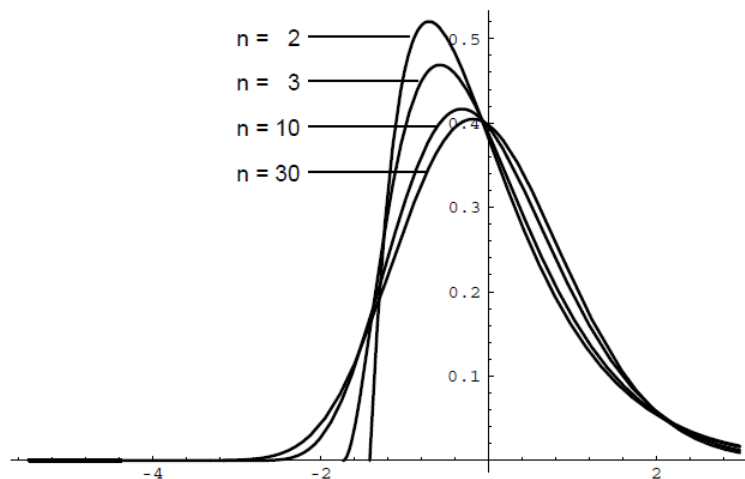


Figure 9.3.1: Density function for S_n^* (exponential case, $\lambda=1$)

We shall give a proof of this theorem in Section 10.3. We will now look at some examples.

✓ Example 9.3.2

Suppose a surveyor wants to measure a known distance, say of 1 mile, using a transit and some method of triangulation. He knows that because of possible motion of the transit, atmospheric distortions, and human error, any one measurement is apt to be slightly in error. He plans to make several measurements and take an average. He assumes that his measurements are

independent random variables with a common distribution of mean $\mu = 1$ and standard deviation $\sigma = .0002$ (so, if the errors are approximately normally distributed, then his measurements are within 1 foot of the correct distance about 65% of the time). What can he say about the average?

He can say that if n is large, the average S_n/n has a density function that is approximately normal, with mean $\mu = 1$ mile, and standard deviation $\sigma = .0002/\sqrt{n}$ miles.

How many measurements should he make to be reasonably sure that his average lies within .0001 of the true value? The Chebyshev inequality says

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq .0001\right) \leq \frac{(.0002)^2}{n(10^{-8})} = \frac{4}{n}, \quad (9.3.3)$$

so that we must have $n \geq 80$ before the probability that his error is less than .0001 exceeds .95.

We have already noticed that the estimate in the Chebyshev inequality is not always a good one, and here is a case in point. If we assume that n is large enough so that the density for S_n is approximately normal, then we have

$$\begin{aligned} P\left(\left|\frac{S_n}{n} - \mu\right| < .0001\right) &= P(-.5\sqrt{n} < S_n^* < .5\sqrt{n}) \\ &\approx \frac{1}{\sqrt{2\pi}} \int_{-.5\sqrt{n}}^{+.5\sqrt{n}} e^{-x^2/2} dx, \end{aligned}$$

and this last expression is greater than .95 if $.5\sqrt{n} \geq 2$. This says that it suffices to take $n = 16$ measurements for the same results. This second calculation is stronger, but depends on the assumption that $n = 16$ is large enough to establish the normal density as a good approximation to S_n^* , and hence to S_n . The Central Limit Theorem here says nothing about how large n has to be. In most cases involving sums of independent random variables, a good rule of thumb is that for $n \geq 30$, the approximation is a good one. In the present case, if we assume that the errors are approximately normally distributed, then the approximation is probably fairly good even for $n = 16$.

Estimating the Mean

✓ Example 9.3.3

(Continuation of Example ✓ 9.3.2) Now suppose our surveyor is measuring an unknown distance with the same instruments under the same conditions. He takes 36 measurements and averages them. How sure can he be that his measurement lies within .0002 of the true value?

Again using the normal approximation, we get

$$\begin{aligned} P\left(\left|\frac{S_n}{n} - \mu\right| < .0002\right) &= P(|S_n^*| < .5\sqrt{n}) \\ &\approx \frac{2}{\sqrt{2\pi}} \int_{-3}^3 e^{-x^2/2} dx \\ &\approx .997. \end{aligned}$$

This means that the surveyor can be 99.7 percent sure that his average is within .0002 of the true value. To improve his confidence, he can take more measurements, or require less accuracy, or improve the quality of his measurements (i.e., reduce the variance σ^2). In each case, the Central Limit Theorem gives quantitative information about the confidence of a measurement process, assuming always that the normal approximation is valid.

Now suppose the surveyor does not know the mean or standard deviation of his measurements, but assumes that they are independent. How should he proceed?

Again, he makes several measurements of a known distance and averages them. As before, the average error is approximately normally distributed, but now with unknown mean and variance.

Sample Mean

If he knows the variance σ^2 of the error distribution is .0002, then he can estimate the mean μ by taking the or of, say, 36 measurements:

$$\bar{\mu} = \frac{x_1 + x_2 + \cdots + x_n}{n}, \quad (9.3.4)$$

where $n = 36$. Then, as before, $E(\bar{\mu}) = \mu$. Moreover, the preceding argument shows that

$$P(|\bar{\mu} - \mu| < .0002) \approx .997. \quad (9.3.5)$$

The interval $(\bar{\mu} - .0002, \bar{\mu} + .0002)$ is called for μ (see Example 9.3.1).

Sample Variance

If he does not know the variance σ^2 of the error distribution, then he can estimate σ^2 by the :

$$\bar{\sigma}^2 = \frac{(x_1 - \bar{\mu})^2 + (x_2 - \bar{\mu})^2 + \cdots + (x_n - \bar{\mu})^2}{n}, \quad (9.3.6)$$

where $n = 36$. The Law of Large Numbers, applied to the random variables $(X_i - \bar{\mu})^2$, says that for large n , the sample variance $\bar{\sigma}^2$ lies close to the variance σ^2 , so that the surveyor can use $\bar{\sigma}^2$ in place of σ^2 in the argument above.

Experience has shown that, in most practical problems of this type, the sample variance is a good estimate for the variance, and can be used in place of the variance to determine confidence levels for the sample mean. This means that we can rely on the Law of Large Numbers for estimating the variance, and the Central Limit Theorem for estimating the mean.

We can check this in some special cases. Suppose we know that the error distribution is with unknown mean and variance. Then we can take a sample of n measurements, find the sample mean $\bar{\mu}$ and sample variance $\bar{\sigma}^2$, and form

$$T_n^* = \frac{S_n - n\bar{\mu}}{\sqrt{n\bar{\sigma}}}, \quad (9.3.7)$$

where $n = 36$. We expect T_n^* to be a good approximation for S_n^* for large n .

t-Density

The statistician W. S. Gosset¹³ has shown that in this case T_n^* has a density function that is not normal but rather a with n degrees of freedom. (The number n of degrees of freedom is simply a parameter which tells us which t -density to use.) In this case we can use the t -density in place of the normal density to determine confidence levels for μ . As n increases, the t -density approaches the normal density. Indeed, even for $n = 8$ the t -density and normal density are practically the same (see Figure 9.3.3).

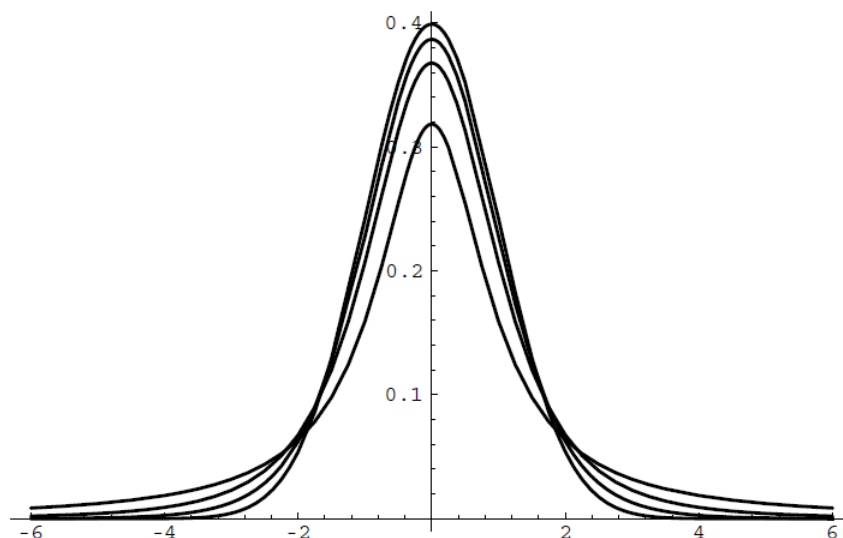


Figure 9.3.3: Graph of t -density for $n = 1, 3, 8$ and the normal density with $\mu = 0, \sigma = 1$

Exercises

Notes of computer problems

(a) Simulation: Recall (see Corollary 5.2) that

$$X = F^{-1}(rnd) \quad (9.3.8)$$

will simulate a random variable with density $f(x)$ and distribution

$$F(X) = \int_{-\infty}^x f(t) dt. \quad (9.3.9)$$

In the case that $f(x)$ is a normal density function with mean μ and standard deviation σ , where neither F nor F^{-1} can be expressed in closed form, use instead

$$X = \sigma \sqrt{-2 \log(rnd) \cos 2\pi(rnd)} + \mu. \quad (9.3.10)$$

(b) Bar graphs: you should aim for about 20 to 30 bars (of equal width) in your graph. You can achieve this by a good choice of the range $[x_{\min}, x_{\max}]$ and the number of bars (for instance, $[\mu - 3\sigma, \mu + 3\sigma]$ with 30 bars will work in many cases). Experiment!

Exercises 9.3.1

Let X be a continuous random variable with mean $\mu(X)$ and variance $\sigma^2(X)$, and let $X^* = (X - \mu)/\sigma$ be its standardized version. Verify directly that $\mu(X^*) = 0$ and $\sigma^2(X^*) = 1$.

Exercises 9.3.2

Let $\{X_k\}$, $1 \leq k \leq n$, be a sequence of independent random variables, all with mean 0 and variance 1, and let S_n , S_n^* , and A_n be their sum, standardized sum, and average, respectively. Verify directly that $S_n^* = S_n/\sqrt{n} = \sqrt{n}A_n$.

Exercises 9.3.3

Let $\{X_k\}$, $1 \leq k \leq n$, be a sequence of random variables, all with mean μ and variance σ^2 , and $Y_k = X_k^*$ be their standardized versions. Let S_n and T_n be the sum of the X_k and Y_k , and S_n^* and T_n^* their standardized version. Show that $S_n^* = T_n^* = T_n/\sqrt{n}$.

Exercises 9.3.4

Suppose we choose independently 25 numbers at random (uniform density) from the interval $[0, 20]$. Write the normal densities that approximate the densities of their sum S_{25} , their standardized sum S_{25}^* , and their average A_{25} .

Exercises 9.3.5

Write a program to choose independently 25 numbers at random from $[0, 20]$, compute their sum S_{25} , and repeat this experiment 1000 times. Make a bar graph for the density of S_{25} and compare it with the normal approximation of Exercise 9.3.4. How good is the fit? Now do the same for the standardized sum S_{25}^* and the average A_{25} .

Exercises 9.3.6

In general, the Central Limit Theorem gives a better estimate than Chebyshev's inequality for the average of a sum. To see this, let A_{25} be the average calculated in Exercise 9.3.5, and let N be the normal approximation for A_{25} . Modify your program in Exercise [exer 9.4.5] to provide a table of the function $F(x) = P(|A_{25} - 10| \geq x)$ = fraction of the total of 1000 trials for which $|A_{25} - 10| \geq x$. Do the same for the function $f(x) = P(|N - 10| \geq x)$. (You can use the normal table, Table [tabl 9.1], or the procedure **NormalArea** for this.) Now plot on the same axes the graphs of $F(x)$, $f(x)$, and the Chebyshev function $g(x) = 4/(3x^2)$. How do $f(x)$ and $g(x)$ compare as estimates for $F(x)$?

Exercises 9.3.7

The Central Limit Theorem says the sums of independent random variables tend to look normal, no matter what crazy distribution the individual variables have. Let us test this by a computer simulation. Choose independently 25 numbers from the interval $[0, 1]$ with the probability density $f(x)$ given below, and compute their sum S_{25} . Repeat this experiment 1000 times, and make up a bar graph of the results. Now plot on the same graph the density $\phi(x) = \text{normal } \backslash, \backslash, \backslash, (x, \mu(S_{25}), \sigma(S_{25}))$. How well does the normal density fit your bar graph in each case?

- $f(x) = 1$.
- $f(x) = 2x$.
- $f(x) = 3x^2$.
- $f(x) = 4|x - 1/2|$.
- $f(x) = 2 - 4|x - 1/2|$.

Exercises 9.3.8

Repeat the experiment described in Exercise 9.3.7 but now choose the 25 numbers from $[0, \infty)$, using $f(x) = e^{-x}$.

Exercises 9.3.9

How large must n be before $S_n = X_1 + X_2 + \cdots + X_n$ is approximately normal? This number is often surprisingly small. Let us explore this question with a computer simulation. Choose n numbers from $[0, 1]$ with probability density $f(x)$, where $n = 3, 6, 12, 20$, and $f(x)$ is each of the densities in Exercise 9.3.7. Compute their sum S_n , repeat this experiment 1000 times, and make up a bar graph of 20 bars of the results. How large must n be before you get a good fit?

Exercises 9.3.10

A surveyor is measuring the height of a cliff known to be about 1000 feet. He assumes his instrument is properly calibrated and that his measurement errors are independent, with mean $\mu = 0$ and variance $\sigma^2 = 10$. He plans to take n measurements and form the average. Estimate, using (a) Chebyshev's inequality and (b) the normal approximation, how large n should be if he wants to be 95 percent sure that his average falls within 1 foot of the true value. Now estimate, using (a) and (b), what value should σ^2 have if he wants to make only 10 measurements with the same confidence?

Exercises 9.3.11

The price of one share of stock in the Pilsdorff Beer Company (see Exercise 8.2.12) is given by Y_n on the n th day of the year. Finn observes that the differences $X_n = Y_{n+1} - Y_n$ appear to be independent random variables with a common distribution having mean $\mu = 0$ and variance $\sigma^2 = 1/4$. If $Y_1 = 100$, estimate the probability that Y_{365} is

- ≥ 100 .
- ≥ 110 .
- ≥ 120 .

Exercises 9.3.1

Test your conclusions in Exercise [exer 9.4.11] by computer simulation. First choose 364 numbers X_i with density $f(x) = \text{normal}(x, 0, 1/4)$. Now form the sum $Y_{365} = 100 + X_1 + X_2 + \cdots + X_{364}$, and repeat this experiment 200 times. Make up a bar graph on $[50, 150]$ of the results, superimposing the graph of the approximating normal density. What does this graph say about your answers in Exercise 9.3.11?

Exercises 9.3.1

Physicists say that particles in a long tube are constantly moving back and forth along the tube, each with a velocity V_k (in cm/sec) at any given moment that is normally distributed, with mean $\mu = 0$ and variance $\sigma^2 = 1$. Suppose there are 10^{20} particles in the tube.

- Find the mean and variance of the average velocity of the particles.
- What is the probability that the average velocity is $\geq 10^{-9}$ cm/sec?

Exercises 9.3.1

An astronomer makes n measurements of the distance between Jupiter and a particular one of its moons. Experience with the instruments used leads her to believe that for the proper units the measurements will be normally distributed with mean d , the true distance, and variance 16. She performs a series of n measurements. Let

$$A_n = \frac{X_1 + X_2 + \cdots + X_n}{n} \quad (9.3.11)$$

be the average of these measurements.

- Show that

$$P\left(A_n - \frac{8}{\sqrt{n}} \leq d \leq A_n + \frac{8}{\sqrt{n}}\right) \approx .95. \quad (9.3.12)$$

- b. When nine measurements were taken, the average of the distances turned out to be 23.2 units. Putting the observed values in (a) gives the for the unknown distance d . Compute this interval.
- c. Why not say in (b) more simply that the probability is .95 that the value of d lies in the computed confidence interval?
- d. What changes would you make in the above procedure if you wanted to compute a 99 percent confidence interval?

Exercises 9.3.1

Plot a bar graph similar to that in Figure [fig 9.61] for the heights of the mid-parents in Galton's data as given in Appendix B and compare this bar graph to the appropriate normal curve.

This page titled [9.3: Central Limit Theorem for Continuous Independent Trials](#) is shared under a [GNU Free Documentation License 1.3](#) license and was authored, remixed, and/or curated by [Charles M. Grinstead & J. Laurie Snell](#) ([American Mathematical Society](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.