8.2: Law of Large Numbers for Continuous Random Variables

In the previous section we discussed in some detail the Law of Large Numbers for discrete probability distributions. This law has a natural analogue for continuous probability distributions, which we consider somewhat more briefly here.

Chebyshev Inequality

Just as in the discrete case, we begin our discussion with the Chebyshev Inequality.

Theorem $8.2.1$

Let X be a continuous random variable with density function $f(x)$. Suppose X has a finite expected value $\mu = E(X)$ and finite variance $\sigma^2 = V(X)$. Then for any positive number $\epsilon > 0$ we have

$$P(|X-\mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2} \tag{8.2.1}$$

The proof is completely analogous to the proof in the discrete case, and we omit it.

Note that this theorem says nothing if $\sigma^2 = V(X)$ is infinite.

Example $8.2.1$

Let X be any continuous random variable with $E(X) = \mu$ and $V(X) = \sigma^2$. Then, if $\epsilon = k\sigma = k$ standard deviations for some integer k, then

$$P(|X-\mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2} \tag{8.2.2}$$

just as in the discrete case.

Law of Large Numbers

With the Chebyshev Inequality we can now state and prove the Law of Large Numbers for the continuous case.

Theorem $8.2.2$

Let $X_1, X_2, \ldots, X_n$ be an independent trials process with a continuous density function f, finite expected value μ, and finite variance σ^2. Let $S_n = X_1 + X_2 + \cdots + X_n$ be the sum of the X_i. Then for any real number $\epsilon > 0$ we have

$$\lim_{n \to \infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = 0 \tag{8.2.3}$$

or equivalently,

$$\lim_{n \to \infty} P\left(\left|\frac{S_n}{n} - \mu\right| < \epsilon\right) = 1 \tag{8.2.4}$$

Note that this theorem is not necessarily true if σ^2 is infinite (see Example 8.8).

As in the discrete case, the Law of Large Numbers says that the average value of n independent trials tends to the expected value as $n \to \infty$, in the precise sense that, given $\epsilon > 0$, the probability that the average value and the expected value differ by more than ϵ tends to 0 as $n \to \infty$.

Once again, we suppress the proof, as it is identical to the proof in the discrete case.

Uniform Case

Example 8.2.2

Suppose we choose at random n numbers from the interval $[0,1]$ with uniform distribution. Then if X_i describes the i th choice, we have

$$\begin{aligned} \mu &= E\left(X_i\right) = \int_0^1 x\,dx = \frac{1}{2}, \\ \sigma^2 &= V\left(X_i\right) = \int_0^1 x^2\,dx - \mu^2 \\ &= \frac{1}{3} - \frac{1}{4} = \frac{1}{12}. \end{aligned}$$

Hence,

$$\begin{aligned} E\left(\frac{S_n}{n}\right) &= \frac{1}{2}, \\ V\left(\frac{S_n}{n}\right) &= \frac{1}{12n}, \end{aligned}$$

and for any $\epsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \epsilon\right) \leq \frac{1}{12n\epsilon^2} \tag{8.2.5}$$

This says that if we choose n numbers at random from $[0,1]$, then the chances are better than $1 - 1/\left(12n\epsilon^2\right)$ that the difference $|S_n/n - 1/2|$ is less than ϵ. Note that ϵ plays the role of the amount of error we are willing to tolerate: If we choose $\epsilon = 0.1$, say, then the chances that $|S_n/n - 1/2|$ is less than 0.1 are better than $1 - 100/(12n)$. For $n = 100$, this is about .92 , but if $n = 1000$, this is better than .99 and if $n = 10,000$, this is better than .999 .

We can illustrate what the Law of Large Numbers says for this example graphically. The density for $A_n = S_n/n$ is determined by

$$f_{A_n}(x) = n f_{S_n}(nx). \tag{8.2.6}$$

We have seen in Section 7.2, that we can compute the density $f_{S_n}(x)$ for the sum of n uniform random variables. In Figure 8.2 we have used this to plot the density for A_n for various values of n. We have shaded in the area for which A_n would lie between .45 and .55 . We see that as we increase n, we obtain more and more of the total area inside the shaded region. The Law of Large Numbers tells us that we can obtain as much of the total area as we please inside the shaded region by choosing n large enough (see also Figure 8.1).

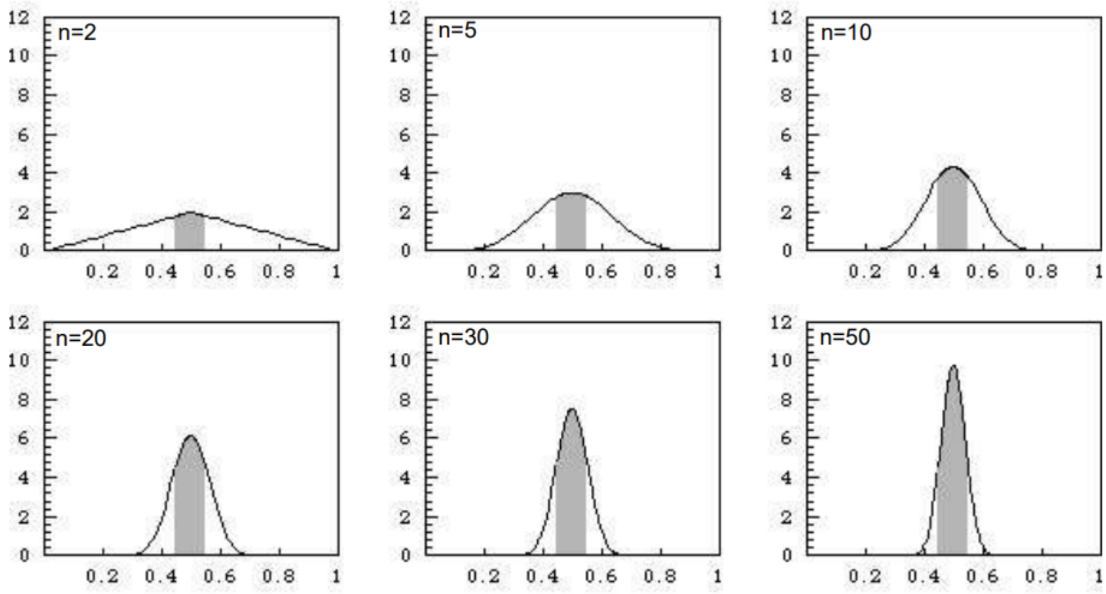

Figure 8.2.1: Illustration of Law of Large Numbers — uniform case

Normal Case

Example 8.2.3

Suppose we choose n real numbers at random, using a normal distribution with mean 0 and variance 1 . Then

$$\begin{aligned}\mu &= E\left(X_i\right)=0, \\ \sigma^2 &= V\left(X_i\right)=1.\end{aligned}$$

Hence,

$$\begin{aligned}E\left(\frac{S_n}{n}\right) &= 0 \\ V\left(\frac{S_n}{n}\right) &= \frac{1}{n}\end{aligned}$$

and, for any $\epsilon>0$,

$$P\left(\left|\frac{S_n}{n}-0\right| \geq \epsilon\right) \leq \frac{1}{n \epsilon^2}. \tag{8.2.7}$$

In this case it is possible to compare the Chebyshev estimate for $P\left(\left|S_n / n-\mu\right| \geq \epsilon\right)$ in the Law of Large Numbers with exact values, since we know the density function for S_n / n exactly (see Example 7.9). The comparison is shown in Table 8.1, for $\epsilon=.1$. The data in this table was produced by the program LawContinuous. We see here that the Chebyshev estimates are in general not very accurate.

Table 8.2.1: Chebyshev estimates.

n	$P\left(\left\vert S_n / n\right\vert \geq .1\right)$	Chebyshev
100	.31731	1.00000
200	.15730	.50000
300	.08326	.33333
400	.04550	.25000

n	P\left(\left|S_n / n\right| \geq .1\right)	Chebyshev
500	.02535	.20000
600	.01431	.16667
700	.00815	.14286
800	.00468	.12500
900	.00270	.11111
1000	.00157	.10000

Monte Carlo Method

Here is a somewhat more interesting example.

Example 8.2.4

Let $g(x)$ be a continuous function defined for $x \in [0,1]$ with values in $[0,1]$. In Section 2.1, we showed how to estimate the area of the region under the graph of $g(x)$ by the Monte Carlo method, that is, by choosing a large number of random values for x and y with uniform distribution and seeing what fraction of the points $P(x,y)$ fell inside the region under the graph (see Example 2.2).

Here is a better way to estimate the same area (see Figure 8.3). Let us choose a large number of independent values X_n at random from $[0,1]$ with uniform density, set $Y_n = g(X_n)$, and find the average value of the Y_n. Then this average is our estimate for the area. To see this, note that if the density function for X_n is uniform,

$$\begin{aligned} \mu = E(Y_n) &= \int_0^1 g(x) f(x) dx \\ &= \int_0^1 g(x) dx \\ &= \text{ average value of } g(x), \end{aligned}$$

while the variance is

$$\sigma^2 = E\left((Y_n - \mu)^2\right) = \int_0^1 (g(x) - \mu)^2 dx < 1 \tag{8.2.8}$$

since for all x in [0,1], $g(x)$ is in $[0,1]$, hence μ is in $[0,1]$, and so $|g(x) - \mu| \leq 1$. Now let $A_n = (1/n)(Y_1 + Y_2 + \cdots + Y_n)$. Then by Chebyshev's Inequality, we have

$$P(|A_n - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2} < \frac{1}{n\epsilon^2}. \tag{8.2.9}$$

This says that to get within ϵ of the true value for $\mu = \int_0^1 g(x) dx$ with probability at least p, we should choose n so that $1/n\epsilon^2 \leq 1 - p$ (i.e., so that $n \geq 1/\epsilon^2(1-p)$). Note that this method tells us how large to take n to get a desired accuracy.

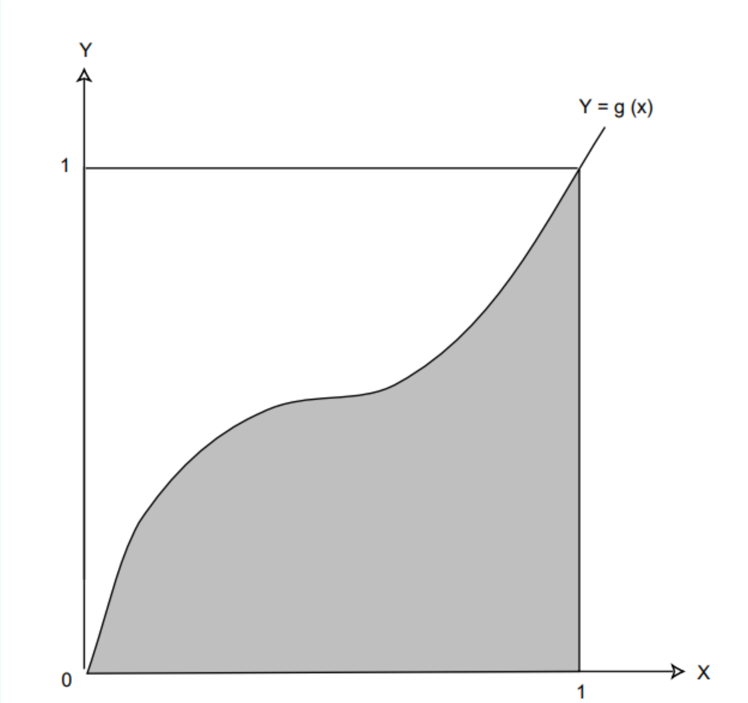

Figure +Area problem.

The Law of Large Numbers requires that the variance σ^2 of the original underlying density be finite: $\sigma^2 < \infty$. In cases where this fails to hold, the Law of Large Numbers may fail, too. An example follows.

Cauchy Case

Example 8.2.1

Suppose we choose n numbers from $(-\infty, +\infty)$ with a Cauchy density with parameter $a = 1$. We know that for the Cauchy density the expected value and variance are undefined (see Example 6.28). In this case, the density function for

$$A_n = \frac{S_n}{n} \tag{8.2.10}$$

is given by (see Example 7.6)

$$f_{A_n}(x) = \frac{1}{\pi(1+x^2)} \tag{8.2.11}$$

that is, the density function for A_n is the same for all n. In this case, as n increases, the density function does not change at all, and the Law of Large Numbers does not hold.

Exercises

Example 8.2.1:

Let X be a continuous random variable with mean $\mu = 10$ and variance $\sigma^2 = 100/3$. Using Chebyshev's Inequality, find an upper bound for the following probabilities. (a) $P(|X-10| \geq 2)$.

(b) $P(|X-10| \geq 5)$.

(c) $P(|X-10| \geq 9)$.

(d) $P(|X-10| \geq 20)$.

Example 8.2.2:

Let X be a continuous random variable with values unformly distributed over the interval $[0,20]$.

(a) Find the mean and variance of X.

(b) Calculate $P(|X-10| \geq 2), P(|X-10| \geq 5), P(|X-10| \geq 9)$, and $P(|X-10| \geq 20)$ exactly. How do your answers compare with those of Exercise 1? How good is Chebyshev's Inequality in this case?

Example 8.2.3:

Let X be the random variable of Exercise 2 .

(a) Calculate the function $f(x)=P(|X-10| \geq x)$.

(b) Now graph the function $f(x)$, and on the same axes, graph the Chebyshev function $g(x)=100/\left(3 x^{2}\right)$. Show that $f(x) \leq g(x)$ for all $x>0$, but that $g(x)$ is not a very good approximation for $f(x)$.

Example 8.2.4:

Let X be a continuous random variable with values exponentially distributed over $[0, \infty)$ with parameter $\lambda=0.1$.

(a) Find the mean and variance of X.

(b) Using Chebyshev's Inequality, find an upper bound for the following probabilities: $P(|X-10| \geq 2), P(|X-10| \geq 5), P(|X-10| \geq 9)$, and $P(|X-10| \geq 20)$.

(c) Calculate these probabilities exactly, and compare with the bounds in (b).

Example 8.2.5:

Let X be a continuous random variable with values normally distributed over $(-\infty,+\infty)$ with mean $\mu=0$ and variance $\sigma^{2}=1$.

(a) Using Chebyshev's Inequality, find upper bounds for the following probabilities: $P(|X| \geq 1), P(|X| \geq 2)$, and $P(|X| \geq 3)$.

(b) The area under the normal curve between -1 and 1 is .6827 , between -2 and 2 is .9545 , and between -3 and 3 it is .9973 (see the table in Appendix A). Compare your bounds in (a) with these exact values. How good is Chebyshev's Inequality in this case?

Example 8.2.6:

If X is normally distributed, with mean μ and variance σ^{2}, find an upper bound for the following probabilities, using Chebyshev's Inequality.

(a) $P(|X-\mu| \geq \sigma)$.

(b) $P(|X-\mu| \geq 2 \sigma)$.

(c) $P(|X-\mu| \geq 3 \sigma)$. (d) $P(|X-\mu| \geq 4 \sigma)$.

Now find the exact value using the program NormalArea or the normal table in Appendix A, and compare.

Example 8.2.7:

If X is a random variable with mean $\mu \neq 0$ and variance σ^{2}, define the relative deviation D of X from its mean by

$$D=\left|\frac{X-\mu}{\mu}\right| \tag{8.2.12}$$

(a) Show that $P(D \geq a) \leq \sigma^{2}/\left(\mu^{2} a^{2}\right)$.

(b) If X is the random variable of Exercise 1, find an upper bound for $P(D \geq .2), P(D \geq .5), P(D \geq .9)$, and $P(D \geq 2)$.

Example 8.2.8:

Let X be a continuous random variable and define the standardized version X^{*} of X by:

$$X^{*}=\frac{X-\mu}{\sigma}. \tag{8.2.13}$$

(a) Show that $P\left(\left|X^{*}\right| \geq a\right) \leq 1/a^{2}$.

(b) If X is the random variable of Exercise 1, find bounds for $P(|X^*| \geq 2)$, $P(|X^*| \geq 5)$, and $P(|X^*| \geq 9)$.

Example 8.2.9:

(a) Suppose a number X is chosen at random from $[0, 20]$ with uniform probability. Find a lower bound for the probability that X lies between 8 and 12, using Chebyshev's Inequality.

(b) Now suppose 20 real numbers are chosen independently from $[0, 20]$ with uniform probability. Find a lower bound for the probability that their average lies between 8 and 12 .

(c) Now suppose 100 real numbers are chosen independently from $[0, 20]$. Find a lower bound for the probability that their average lies between 8 and 12.

Example 8.2.10:

A student's score on a particular calculus final is a random variable with values of $[0, 100]$, mean 70 , and variance 25 .

(a) Find a lower bound for the probability that the student's score will fall between 65 and 75 .

(b) If 100 students take the final, find a lower bound for the probability that the class average will fall between 65 and 75 .

Example 8.2.11:

The Pilsdorff beer company runs a fleet of trucks along the 100 mile road from Hangtown to Dry Gulch, and maintains a garage halfway in between. Each of the trucks is apt to break down at a point X miles from Hangtown, where X is a random variable uniformly distributed over $[0, 100]$.

(a) Find a lower bound for the probability $P(|X-50| \leq 10)$. (b) Suppose that in one bad week, 20 trucks break down. Find a lower bound for the probability $P(|A_{20}-50| \leq 10)$, where A_{20} is the average of the distances from Hangtown at the time of breakdown.

Example 8.2.12:

A share of common stock in the Pilsdorff beer company has a price Y_n on the nth business day of the year. Finn observes that the price change $X_n = Y_{n+1} - Y_n$ appears to be a random variable with mean $\mu = 0$ and variance $\sigma^2 = 1/4$. If $Y_1 = 30$, find a lower bound for the following probabilities, under the assumption that the X_n 's are mutually independent.

(a) $P(25 \leq Y_2 \leq 35)$.

(b) $P(25 \leq Y_{11} \leq 35)$.

(c) $P(25 \leq Y_{101} \leq 35)$.

Example 8.2.13:

Suppose one hundred numbers $X_1, X_2, \ldots, X_{100}$ are chosen independently at random from $[0, 20]$. Let $S = X_1 + X_2 + \cdots + X_{100}$ be the sum, $A = S/100$ the average, and $S^* = (S-1000)/(10/\sqrt{3})$ the standardized sum. Find lower bounds for the probabilities

(a) $P(|S-1000| \leq 100)$.

(b) $P(|A-10| \leq 1)$.

(c) $P(|S^*| \leq \sqrt{3})$.

Example 8.2.14:

Let X be a continuous random variable normally distributed on $(-\infty, +\infty)$ with mean 0 and variance 1 . Using the normal table provided in Appendix A, or the program NormalArea, find values for the function $f(x) = P(|X| \geq x)$ as x increases from 0 to 4.0 in steps of .25. Note that for $x \geq 0$ the table gives $NA(0, x) = P(0 \leq X \leq x)$ and thus $P(|X| \geq x) = 2(.5 - NA(0, x)$. Plot by hand the graph of $f(x)$ using these values, and the graph of the Chebyshev function $g(x) = 1/x^2$, and compare (see Exercise 3).

Example 8.2.15:

Repeat Exercise 14, but this time with mean 10 and variance 3. Note that the table in Appendix A presents values for a standard normal variable. Find the standardized version X^* for X, find values for $f^*(x) = P(|X^*| \geq x)$ as in Exercise 14, and then rescale these values for $f(x) = P(|X-10| \geq x)$. Graph and compare this function with the Chebyshev function $g(x) = 3/x^2$.

Example 8.2.16:

Let $Z = X/Y$ where X and Y have normal densities with mean 0 and standard deviation 1 . Then it can be shown that Z has a Cauchy density.

(a) Write a program to illustrate this result by plotting a bar graph of 1000 samples obtained by forming the ratio of two standard normal outcomes. Compare your bar graph with the graph of the Cauchy density. Depending upon which computer language you use, you may or may not need to tell the computer how to simulate a normal random variable. A method for doing this was described in Section 5.2. (b) We have seen that the Law of Large Numbers does not apply to the Cauchy density (see Example 8.8). Simulate a large number of experiments with Cauchy density and compute the average of your results. Do these averages seem to be approaching a limit? If so can you explain why this might be?

Example 8.2.17:

Show that, if $X \geq 0$, then $P(X \geq a) \leq E(X)/a$.

Example 8.2.18:

(Lamperti [9]) Let X be a non-negative random variable. What is the best upper bound you can give for $P(X \geq a)$ if you know

(a) $E(X) = 20$.

(b) $E(X) = 20$ and $V(X) = 25$.

(c) $E(X) = 20, V(X) = 25$, and X is symmetric about its mean.

[9] Private communication.