Probability Foundations for Electrical Engineers **July-November 2015**

Lecture 29: The Laws of Large Numbers

Lecturer: Dr. Krishna Jagannathan *Scribe: Ravi Kolla, Vishakh Hegde and Arjun Bhagoji*

In this lecture, we study the laws of large numbers (LLNs), which are arguably the single most important class of theorems, which form the backbone of probability theory. In particular, the LLNs provide an intuitive interpretation for the expectation of a random variable as the 'average value' of the random variable. In the case of i.i.d. random variables that we consider in this lecture, the LLN roughly says that the sample average of a large number of i.i.d. random variables converges to the expected value. The sense of convergence in the weak law of large numbers is convergence in probability. The strong law of large numbers, as the name suggests, asserts the stronger notion of almost sure convergence.

29.1 Weak Law of Large Numbers

The earliest available proof of the weak law of large number dates to the year 1713, in the posthumously published work of Jacob Bernoulli. It asserts convergence in probability of the sample average to the expected value.

Theorem 29.1 (Weak Law of Large numbers) *Let $X_1, X_2, \ldots$ be i.i.d random variables with finite mean, $\mathbb{E}[X]$. Let $S_n = \sum_{i=1}^{n} X_i$. Then,*

$$\frac{S_n}{n} \xrightarrow{i.p} \mathbb{E}[X].$$

Proof: First, we give a *partial* proof by assuming the variance of X to be finite *i.e.,* $\sigma_X^2 < \infty$. Since X_i's are i.i.d, $\mathbb{E}[S_n] = n\mathbb{E}[X]$, $Var\left(S_n\right) = nVar(X) \Rightarrow \mathbb{E}\left[\frac{S_n}{n}\right] = \mathbb{E}[X], Var\left(\frac{S_n}{n}\right) = \frac{\sigma_X^2}{n}$.

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left[X\right]\right| > \epsilon\right) &\leq \lim_{n\to\infty} \frac{Var\left(\frac{S_n}{n}\right)}{\epsilon^2} \quad \text{(By Chebyshev's Inequality)},\\
&= \lim_{n\to\infty} \frac{\sigma_X^2}{n\epsilon^2},\\
&= 0.
\end{aligned}$$

■

Next, we give a general proof using characteristic functions.

Proof: Assume that X_i (where $i = 1, 2, \ldots, n, \ldots$) are i.i.d random variables. The characteristic function of X_i be $C_{X_i}(t) \equiv C_X(t)$ for any $i \in \{1, 2, \ldots, n\}$. Let $S_n = X_1 + X_2 + \ldots X_n$ be the sum of these n i.i.d random variables. The following can be easily verified:

$$\begin{aligned}
C_{S_n} &= [C_X(t)]^n = \mathbb{E}[e^{itS_n}],\\
&= \mathbb{E}[e^{\frac{itnS_n}{n}}],\\
&= C_{\frac{S_n}{n}}(nt).
\end{aligned}$$

This implies that,

$$C_{\frac{S_n}{n}}(t) = [C_X\left(\frac{t}{n}\right)]^n,$$
$$= \left[1 + \frac{i\mathbb{E}[X]t}{n} + o\left(\frac{t}{n}\right)\right]^n.$$

As $n \to \infty$, we have,

$$C_{\frac{S_n}{n}}(t) \to e^{i\mathbb{E}[X]t}, \ \forall t \in \mathbb{R}.$$

Note that, $e^{i\mathbb{E}[X]t}$ is a valid characteristic function. In fact, it is a characteristic function of a constant random variable which takes the value $\mathbb{E}[X]$. From the theorem on convergence of characteristic functions, we have

$$\frac{S_n}{n} \xrightarrow{D} \mathbb{E}[X].$$

Since $\mathbb{E}[X]$ is a constant[1], we have,

$$\frac{S_n}{n} \xrightarrow{i.p} \mathbb{E}[X].$$

■

29.2 Strong Law of Large Numbers

The Strong Law of Large Numbers (SLLN) gives us the condition when the sample average $\left(\frac{S_n}{n}\right)$ converges almost surely to the expected value.

Theorem 29.2 *If $\{X_i, i \geq 1\}$ is a sequence of i.i.d RVs with $\mathbb{E}[|X_i|] < \infty$, then $\frac{S_n}{n} \xrightarrow{a.s.} \mathbb{E}[X]$, i.e.,* $\mathbb{P}\left(\omega | \frac{S_n(\omega)}{n} \to \mathbb{E}[X]\right) = 1.$

Here, $S_n(\omega)$ is just $X_1(\omega) + X_2(\omega) \cdots + X_n(\omega)$. Thus, for a fixed $\omega \in \Omega$, $\left\{\frac{S_n(\omega)}{n}, n \geq 1\right\}$ is a sequence of real numbers. Then, there are the following three possibilities regarding the convergence of this sequence:

1. The sequence $\frac{S_n(\omega)}{n}$ does not converge as $n \to \infty$.
2. The sequence $\frac{S_n(\omega)}{n}$ converges to a value other than $\mathbb{E}[X]$, as $n \to \infty$.
3. The sequence $\frac{S_n(\omega)}{n}$ converges to $\mathbb{E}[X]$ as $n \to \infty$.

The SLLN asserts that the set of $\omega \in \Omega$ where the third possibility holds has a probability of 1. Also, the SLLN implies the WLLN because almost sure convergence implies convergence in probability. From Theorem 28.16, we obtain another way of stating the SLLN as given below

$$\lim_{n\to\infty} \mathbb{P}\left(\bigcup_{m\geq n}\left\{\omega : \left|\frac{S_m(\omega)}{m} - \mathbb{E}[X]\right| > \epsilon\right\}\right) = 0, \quad \forall \epsilon > 0. \tag{29.1}$$

A general proof of the SLLN is rather long, so we will restrict ourselves to two partial proofs, each of which makes a stronger assumption than needed about the moments of the random variable X.

[1]Recall that convergence in probability is equivalent to convergence in distribution, when the limit is a constant.

29.3 Partial Proof 1 (assuming finite fourth moment)

Proof: Assume $\mathbb{E}[X_i^4] = \eta < \infty$ and without loss of generality, $\mathbb{E}[X] = 0$. The second assumption is not crucial. We want to show that $\frac{S_n}{n} \overset{a.s.}{\to} 0$.

Now,

$$\begin{aligned}\mathbb{E}[S_n^4] &= \mathbb{E}[(X_1 + X_2 + \cdots X_n)^4],\\ &= n\eta + \binom{4}{2}\binom{n}{2}\mathbb{E}[X_1^2 X_2^2], && (29.2)\\ &= n\eta + 6\binom{n}{2}\sigma^4,\\ &\leq n\eta + 3n^2\sigma^4. && (29.3)\end{aligned}$$

In (29.2), the coefficient of η is n because there are n terms of the form X_i^4. Terms of the form $X_i^3 X_j$ are not present as our assumption that $\mathbb{E}[X] = 0$ ensures that these terms go to zero. For the other surviving terms of the form $X_i^2 X_j^2$, the coefficient arises because there are $\binom{n}{2}$ ways to choose the distinct indices i and j, after which one can choose X_i from 2 out of the 4 terms being multiplied together, in which case X_j will come from the other two terms.

Now, we make use of the Markov inequality and substitute the inequality for $\mathbb{E}\left[S_n^4\right]$ from (29.3).

$$\begin{aligned}\mathbb{P}\left(\left|\frac{S_n}{n}\right|^4 > \epsilon\right) &\leq \frac{\mathbb{E}[S_n^4]}{n^4\epsilon},\\ &\leq \frac{n\eta + 3n^2\sigma^4}{n^4\epsilon},\\ &= \frac{\eta}{n^3\epsilon} + \frac{3\sigma^4}{n^2\epsilon}. && (29.4)\end{aligned}$$

Then, from (29.4),

$$\sum_{n=1}^{\infty}\mathbb{P}\left(\left|\frac{S_n}{n}\right|^4 > \epsilon\right) \leq \sum_{n=1}^{\infty}\frac{\eta}{n^3\epsilon} + \frac{3\sigma^4}{n^2\epsilon} < \infty. \quad (29.5)$$

Using the first Borel-Cantelli lemma, we can conclude

$$\begin{aligned}&\left|\frac{S_n}{n}\right|^4 \xrightarrow{a.s.} 0,\\ \Rightarrow &\frac{S_n}{n} \xrightarrow{a.s.} 0.\end{aligned}$$

29.4 Partial Proof 2 (assuming finite variance)

Assume $\sigma^2 < \infty$ and $\mathbb{E}[X] = \mu$. We begin by proving the SLLN for $X_i \geq 0$. From the partial proof of the Weak Law of Large Numbers, we have

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \epsilon\right) \leq \frac{\sigma_X^2}{n\epsilon^2}. \quad (29.6)$$

To obtain a.s. convergence, consider a deterministic subsequence $n_i = i^2$, $i \geq 1$. Thus we get,

$$\mathbb{P}\left(\left|\frac{S_{i^2}}{i^2} - \mu\right| > \epsilon\right) \leq \frac{\sigma_X^2}{i^2\epsilon^2},$$

which implies that

$$\sum_{i=1}^{\infty} \mathbb{P}\left(\left|\frac{S_{i^2}}{i^2} - \mu\right| > \epsilon\right) < \infty, \ \forall \epsilon > 0,$$

Using Borel-Cantelli lemma 1 we conclude that

$$\frac{S_{i^2}}{i^2} \xrightarrow{a.s.} \mu \quad \text{as } i \to \infty.$$

Let n be such that $i^2 \leq n \leq (i+1)^2$. Since $X_i \geq 0$,

$$S_{i^2} \leq S_n \leq S_{(i+1)^2},$$
$$\Rightarrow \frac{S_{i^2}}{(i+1)^2} \leq \frac{S_n}{n} \leq \frac{S_{(i+1)^2}}{i^2}.$$

Multiplying the expression on the left by i^2 in both the numerator and denominator, and similarly for the expression on the right, except by $(i+1)^2$, we get

$$\frac{S_{i^2}}{(i+1)^2}\frac{i^2}{i^2} \leq \frac{S_n}{n} \leq \frac{S_{(i+1)^2}}{i^2}\frac{(i+1)^2}{(i+1)^2},$$
$$\frac{S_{i^2}}{i^2}\frac{i^2}{(i+1)^2} \leq \frac{S_n}{n} \leq \frac{S_{(i+1)^2}}{(i+1)^2}\frac{(i+1)^2}{i^2},$$

As $i \to \infty$, we have

$$\mu \leq \frac{S_n}{n} \leq \mu.$$

Thus, by the sandwich theorem, we get

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu.$$

To generalise to arbitrary RVs with a finite variance, we just write $X_n = X_n^+ - X_n^-$ and proceed as above since both X_n^+ and X_n^- have a finite variance and are non-negative.

■

29.5 Exercises

1. *[Gallager]* A town starts a mosquito control program and the random variable Z_n is the number of mosquitoes at the end of the n^{th} year (n = 0, 1, ...). Let X_n be the growth rate of mosquitoes in the year n i.e. $Z_n = X_n Z_{n-1}$, $n \geq 1$. Assume that $\{X_n, n \geq 1\}$ is a sequence of i.i.d. random variables with the PMF $\mathbb{P}(X = 2) = \frac{1}{2}$, $\mathbb{P}(X = \frac{1}{2}) = \frac{1}{4}$ and $\mathbb{P}(X = \frac{1}{4}) = \frac{1}{4}$. Suppose Z_0, the initial number of mosquitoes, is a known constant and assume, for simplicity and consistency, that Z_n can take non-integer values.

   (a) Find $\mathbb{E}[Z_n]$ and $\lim_{n\to\infty} \mathbb{E}[Z_n]$.

(b) Based on your answer to part (a), can you conclude whether or not the mosquito control program is successful? What would your conclusion be?

(c) Let $W_n = \log_2 X_n$. Find $\mathbb{E}[W_n]$ and $\mathbb{E}[\log_2 \frac{Z_n}{Z_0}]$.

(d) Show that there exists a constant α such that $\lim_{n\to\infty} \frac{1}{n}\log_2 \frac{Z_n}{Z_0} = \alpha$ almost surely.

(e) Show that there is a constant β such that $\lim_{n\to\infty} Z_n = \beta$ almost surely.

(f) Based on your answer to part (e), can you conclude whether or not the mosquito control program is successful? What would your conclusion be?

(g) How do you reconcile your answers to parts (b) and (f)?

2. Imagine a world in which the value of π is unknown. It is known that area of a circle is proportional to the square of the radius, but the constant of proportionality is unknown. Suppose you are given a uniform random variable generator, and you can generate as many i.i.d. samples as you need, devise a method to estimate the value of the proportionality constant without actually measuring the area/circumference of the circle.