Stat 709: Mathematical Statistics
Lecture 15

Jun Shao

Department of Statistics
University of Wisconsin
Madison, WI 53706, USA

Lecture 15: The law of large numbers

- The law of large numbers concerns the limiting behavior of a sum of random variables.
- The weak law of large numbers (WLLN) refers to convergence in probability.
- The strong law of large numbers (SLLN) refers to a.s. convergence.

Lemma 1.6. (Kronecker's lemma)

Let $x_n \in \mathscr{R}$, $a_n \in \mathscr{R}$, $0 < a_n \le a_{n+1}$, $n = 1, 2, ...$, and $a_n \to \infty$. If the series $\sum_{n=1}^{\infty} x_n / a_n$ converges, then $a_n^{-1} \sum_{i=1}^{n} x_i \to 0$.

Our first result gives the WLLN and SLLN for a sequence of independent and identically distributed (i.i.d.) random variables.

Lecture 15: The law of large numbers

- The law of large numbers concerns the limiting behavior of a sum of random variables.
- The weak law of large numbers (WLLN) refers to convergence in probability.
- The strong law of large numbers (SLLN) refers to a.s. convergence.

Lemma 1.6. (Kronecker's lemma)

Let $x_n \in \mathscr{R}$, $a_n \in \mathscr{R}$, $0 < a_n \le a_{n+1}$, $n = 1, 2, ...$, and $a_n \to \infty$.
If the series $\sum_{n=1}^{\infty} x_n/a_n$ converges, then $a_n^{-1} \sum_{i=1}^{n} x_i \to 0$.

Our first result gives the WLLN and SLLN for a sequence of independent and identically distributed (i.i.d.) random variables.

Lecture 15: The law of large numbers

- The law of large numbers concerns the limiting behavior of a sum of random variables.
- The weak law of large numbers (WLLN) refers to convergence in probability.
- The strong law of large numbers (SLLN) refers to a.s. convergence.

Lemma 1.6. (Kronecker's lemma)

Let $x_n \in \mathscr{R}$, $a_n \in \mathscr{R}$, $0 < a_n \leq a_{n+1}$, $n = 1, 2, \ldots$, and $a_n \to \infty$.
If the series $\sum_{n=1}^{\infty} x_n/a_n$ converges, then $a_n^{-1} \sum_{i=1}^{n} x_i \to 0$.

Our first result gives the WLLN and SLLN for a sequence of independent and identically distributed (i.i.d.) random variables.

Theorem 1.13

Let $X_1, X_2, \ldots$ be i.i.d. random variables.

(i) (The WLLN). A necessary and sufficient condition for the existence of a sequence of real numbers $\{a_n\}$ for which

$$\frac{1}{n}\sum_{i=1}^{n} X_i - a_n \to_p 0$$

is that $nP(|X_1| > n) \to 0$, in which case we may take

$$a_n = E(X_1 I_{\{|X_1| \le n\}}).$$

(ii) (The SLLN). A necessary and sufficient condition for the existence of a constant c for which

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to_{a.s.} c$$

is that $E|X_1| < \infty$, in which case $c = EX_1$ and

$$\frac{1}{n}\sum_{i=1}^{n} c_i (X_i - EX_1) \to_{a.s.} 0$$

for any bounded sequence of real numbers $\{c_i\}$.

Proof of Theorem 1.13(i)

We prove the sufficiency.
The proof of necessity can be found in Petrov (1975).
Consider a sequence of random variables obtained by truncating X_j's at n:

$$Y_{nj} = X_j I_{\{|X_j| \leq n\}}.$$

Let

$$T_n = X_1 + \cdots + X_n, \quad Z_n = Y_{n1} + \cdots + Y_{nn}.$$

Then

$$P(T_n \neq Z_n) \leq \sum_{j=1}^{n} P(Y_{nj} \neq X_j) = nP(|X_1| > n) \to 0.$$

For any $\varepsilon > 0$, it follows from Chebyshev's inequality that

$$P\left(\left|\frac{Z_n - EZ_n}{n}\right| > \varepsilon\right) \leq \frac{\mathrm{var}(Z_n)}{\varepsilon^2 n^2} = \frac{\mathrm{var}(Y_{n1})}{\varepsilon^2 n} \leq \frac{EY_{n1}^2}{\varepsilon^2 n},$$

where the last equality follows from the fact that $Y_{nj}, j = 1, ..., n$, are i.i.d.

Proof of Theorem 1.13(i) (continued)

From integration by parts, we obtain that

$$\frac{EY_{n1}^2}{n} = \frac{1}{n}\int_{[0,n]} x^2 dF_{|X_1|}(x) = \frac{2}{n}\int_0^n xP(|X_1|>x)dx - nP(|X_1|>n),$$

which converges to 0 since $nP(|X_1|>n) \to 0$ (why?).
This proves that $(Z_n - EZ_n)/n \to_p 0$, which together with $P(T_n \neq Z_n) \to 0$ and the fact that $EY_{nj} = E(X_1 I_{\{|X_1|\le n\}})$ imply the result.

Proof of Theorem 1.13(ii)

The proof for sufficiency is given in the textbook.
We prove the necessity.
Suppose that $T_n/n \to_{a.s.}$ holds for some $c \in \mathscr{R}$, $T_n = X_1 + \cdots + X_n$.
Then

$$\frac{X_n}{n} = \frac{T_n}{n} - c - \frac{n-1}{n}\left(\frac{T_{n-1}}{n-1} - c\right) + \frac{c}{n} \to_{a.s.} 0.$$

Proof of Theorem 1.13(i) (continued)

From integration by parts, we obtain that

$$\frac{EY_{n1}^2}{n} = \frac{1}{n}\int_{[0,n]} x^2 dF_{|X_1|}(x) = \frac{2}{n}\int_0^n xP(|X_1| > x)dx - nP(|X_1| > n),$$

which converges to 0 since $nP(|X_1| > n) \to 0$ (why?).
This proves that $(Z_n - EZ_n)/n \to_p 0$, which together with $P(T_n \neq Z_n) \to 0$ and the fact that $EY_{nj} = E(X_1 I_{\{|X_1| \leq n\}})$ imply the result.

Proof of Theorem 1.13(ii)

The proof for sufficiency is given in the textbook.
We prove the necessity.
Suppose that $T_n/n \to_{a.s.}$ holds for some $c \in \mathscr{R}$, $T_n = X_1 + \cdots + X_n$.
Then

$$\frac{X_n}{n} = \frac{T_n}{n} - c - \frac{n-1}{n}\left(\frac{T_{n-1}}{n-1} - c\right) + \frac{c}{n} \to_{a.s.} 0.$$

Proof of Theorem 1.13. (ii)

From Exercise 114, $X_n/n \to_{a.s.} 0$ and the i.i.d. assumption on X_n's imply

$$\sum_{n=1}^{\infty} P(|X_n| \geq n) = \sum_{n=1}^{\infty} P(|X_1| \geq n) < \infty,$$

which implies $E|X_1| < \infty$ (Exercise 54).

From the proved sufficiency, $c = EX_1$.

Remarks

- If $E|X_1| < \infty$, then $a_n = E(X_1 I_{\{|X_1| \leq n\}}) \to EX_1$ and result the WLLN is actually established in Example 1.28 in a much simpler way.
- On the other hand, if $E|X_1| < \infty$, then a stronger result, the SLLN, can be obtained.
- Some results for the case of $E|X_1| = \infty$ can be found in Exercise 148 and Theorem 5.4.3 in Chung (1974).

The next result is for sequences of independent but not necessarily identically distributed random variables.

Proof of Theorem 1.13. (ii)

From Exercise 114, $X_n/n \to_{a.s.} 0$ and the i.i.d. assumption on X_n's imply

$$\sum_{n=1}^{\infty} P(|X_n| \geq n) = \sum_{n=1}^{\infty} P(|X_1| \geq n) < \infty,$$

which implies $E|X_1| < \infty$ (Exercise 54).
From the proved sufficiency, $c = EX_1$.

Remarks

- If $E|X_1| < \infty$, then $a_n = E(X_1 I_{\{|X_1| \leq n\}}) \to EX_1$ and result the WLLN is actually established in Example 1.28 in a much simpler way.
- On the other hand, if $E|X_1| < \infty$, then a stronger result, the SLLN, can be obtained.
- Some results for the case of $E|X_1| = \infty$ can be found in Exercise 148 and Theorem 5.4.3 in Chung (1974).

The next result is for sequences of independent but not necessarily identically distributed random variables.

Proof of Theorem 1.13. (ii)

From Exercise 114, $X_n/n \to_{a.s.} 0$ and the i.i.d. assumption on X_n's imply

$$\sum_{n=1}^{\infty} P(|X_n| \geq n) = \sum_{n=1}^{\infty} P(|X_1| \geq n) < \infty,$$

which implies $E|X_1| < \infty$ (Exercise 54).
From the proved sufficiency, $c = EX_1$.

Remarks

- If $E|X_1| < \infty$, then $a_n = E(X_1 I_{\{|X_1| \leq n\}}) \to EX_1$ and result the WLLN is actually established in Example 1.28 in a much simpler way.
- On the other hand, if $E|X_1| < \infty$, then a stronger result, the SLLN, can be obtained.
- Some results for the case of $E|X_1| = \infty$ can be found in Exercise 148 and Theorem 5.4.3 in Chung (1974).

The next result is for sequences of independent but not necessarily identically distributed random variables.

Theorem 1.14

Let $X_1, X_2, \ldots$ be independent random variables with finite expectations.

(i) (The SLLN). If there is a constant $p \in [1,2]$ such that

$$\sum_{i=1}^{\infty} \frac{E|X_i|^p}{i^p} < \infty, \tag{1}$$

then

$$\frac{1}{n} \sum_{i=1}^{n} (X_i - EX_i) \to_{a.s.} 0.$$

(ii) (The WLLN). If there is a constant $p \in [1,2]$ such that

$$\lim_{n\to\infty} \frac{1}{n^p} \sum_{i=1}^{n} E|X_i|^p = 0, \tag{2}$$

then

$$\frac{1}{n} \sum_{i=1}^{n} (X_i - EX_i) \to_p 0.$$

Remarks

- Note that (1) implies (2) (Lemma 1.6).
- The result in Theorem 1.14(i) is called Kolmogorov's SLLN when $p = 2$ and is due to Marcinkiewicz and Zygmund when $1 \le p < 2$.
- An obvious sufficient condition for (1) with $p \in (1,2]$ is $\sup_n E|X_n|^p < \infty$.
- The WLLN and SLLN have many applications in probability and statistics.

Example 1.32

Let f and g be continuous functions on $[0,1]$ satisfying $0 \le f(x) \le Cg(x)$ for all x, where $C > 0$ is a constant.
We now show that

$$\lim_{n\to\infty} \int_0^1 \int_0^1 \cdots \int_0^1 \frac{\sum_{i=1}^n f(x_i)}{\sum_{i=1}^n g(x_i)} dx_1 dx_2 \cdots dx_n = \frac{\int_0^1 f(x)dx}{\int_0^1 g(x)dx} \qquad (3)$$

(assuming that $\int_0^1 g(x)dx \neq 0$).

Remarks

- Note that (1) implies (2) (Lemma 1.6).
- The result in Theorem 1.14(i) is called Kolmogorov's SLLN when $p = 2$ and is due to Marcinkiewicz and Zygmund when $1 \le p < 2$.
- An obvious sufficient condition for (1) with $p \in (1,2]$ is $\sup_n E|X_n|^p < \infty$.
- The WLLN and SLLN have many applications in probability and statistics.

Example 1.32

Let f and g be continuous functions on $[0,1]$ satisfying $0 \le f(x) \le Cg(x)$ for all x, where $C > 0$ is a constant.
We now show that

$$\lim_{n\to\infty} \int_0^1 \int_0^1 \cdots \int_0^1 \frac{\sum_{i=1}^n f(x_i)}{\sum_{i=1}^n g(x_i)} dx_1 dx_2 \cdots dx_n = \frac{\int_0^1 f(x)dx}{\int_0^1 g(x)dx} \qquad (3)$$

(assuming that $\int_0^1 g(x)dx \neq 0$).

Example 1.32 (continued)

$X_1, X_2, \ldots$ be i.i.d. random variables having the uniform distribution on $[0,1]$.

By Theorem 1.2,

$$E[f(X_1)] = \int_0^1 f(x)dx < \infty, \quad E[g(X_1)] = \int_0^1 g(x)dx < \infty.$$

By the SLLN (Theorem 1.13(ii)),

$$\frac{1}{n}\sum_{i=1}^n f(X_i) \to_{a.s.} E[f(X_1)], \quad \frac{1}{n}\sum_{i=1}^n g(X_i) \to_{a.s.} E[g(X_1)],$$

By Theorem 1.10(i),

$$\frac{\sum_{i=1}^n f(X_i)}{\sum_{i=1}^n g(X_i)} \to_{a.s.} \frac{E[f(X_1)]}{E[g(X_1)]}. \tag{4}$$

Since the random variable on the left-hand side of (4) is bounded by C, result (3) follows from the dominated convergence theorem and the fact that the left-hand side of (3) is the expectation of the random variable on the left-hand side of (4).

Example

Let $T_n = \sum_{i=1}^n X_i$, where X_n's are independent random variables satisfying $P(X_n = \pm n^\theta) = 0.5$ and $\theta > 0$ is a constant.
We want to show that $T_n/n \to_{a.s.} 0$ when $\theta < 0.5$.
For $\theta < 0.5$,

$$\sum_{n=1}^\infty \frac{EX_n^2}{n^2} = \sum_{n=1}^\infty \frac{n^{2\theta}}{n^2} < \infty.$$

By the Kolmogorov strong law of large numbers, $T_n/n \to_{a.s.} 0$.

Example (Exercise 165)

Let $X_1, X_2, \ldots$ be independent random variables.
Suppose that

$$\frac{1}{\sigma_n} \sum_{j=1}^n (X_j - EX_j) \to_d N(0,1),$$

where $\sigma_n^2 = \mathrm{var}(\sum_{j=1}^n X_j)$.

Example

Let $T_n = \sum_{i=1}^n X_i$, where X_n's are independent random variables satisfying $P(X_n = \pm n^\theta) = 0.5$ and $\theta > 0$ is a constant.

We want to show that $T_n/n \to_{a.s.} 0$ when $\theta < 0.5$.

For $\theta < 0.5$,

$$\sum_{n=1}^{\infty} \frac{EX_n^2}{n^2} = \sum_{n=1}^{\infty} \frac{n^{2\theta}}{n^2} < \infty.$$

By the Kolmogorov strong law of large numbers, $T_n/n \to_{a.s.} 0$.

Example (Exercise 165)

Let $X_1, X_2, ...$ be independent random variables.

Suppose that

$$\frac{1}{\sigma_n} \sum_{j=1}^{n} (X_j - EX_j) \to_d N(0,1),$$

where $\sigma_n^2 = \mathrm{var}(\sum_{j=1}^n X_j)$.

Example (Exercise 165)

We want to show that

$$\frac{1}{n}\sum_{j=1}^{n}(X_j - EX_j) \to_p 0 \quad \text{iff} \quad \sigma_n/n \to 0.$$

If $\sigma_n/n \to 0$, then by Slutsky's theorem,

$$\frac{1}{n}\sum_{j=1}^{n}(X_j - EX_j) = \frac{\sigma_n}{n}\frac{1}{\sigma_n}\sum_{j=1}^{n}(X_j - EX_j) \to_d 0.$$

Assume now σ_n/n does not converge to 0 but $n^{-1}\sum_{j=1}^{n}(X_j - EX_j) \to_p 0$.
Without loss of generality, assume that $\sigma_n/n \to c \in (0,\infty]$.
By Slutsky's theorem,

$$\frac{1}{\sigma_n}\sum_{j=1}^{n}(X_j - EX_j) = \frac{n}{\sigma_n}\frac{1}{n}\sum_{j=1}^{n}(X_j - EX_j) \to_p 0.$$

This contradicts the fact that $\sum_{j=1}^{n}(X_j - EX_j)/\sigma_n \to_d N(0,1)$.
Hence, $n^{-1}\sum_{j=1}^{n}(X_j - EX_j)$ does not converge to 0 in probability.