STAT 7200

Introduction to Advanced Probability
Lecture 13

Taylor R. Brown

1. Theory of Convergence
   - Convergence Almost Surely
   - Convergence in Probability
   - Law of Large Numbers

Convergence Almost Surely and Convergence in Probability

- We say that $\{Z_n\}$ converges to Z almost surely (or a.s., or with probability 1), if $\mathrm{P}(\{\omega \in \Omega : \lim_{n\to\infty} Z_n(\omega) = Z(\omega)\}) = 1$. This definition is equivalent to $\mathrm{P}(|Z_n - Z| \geq \varepsilon \ i.o.) = 0$ (or $\mathrm{P}(|Z_n - Z| < \varepsilon \ a.a.) = 1$) for each $\varepsilon > 0$,
- By the (first) Borel-Cantelli Lemma, for r.v.s. $Z, Z_1, Z_2, \ldots$, if for each $\varepsilon > 0$, $\sum_n \mathrm{P}(|Z_n - Z| \geq \varepsilon) < \infty$, then $\mathrm{P}(Z_n \to Z) = 1$.

- We say that $\{Z_n\}$ converges to Z in probability, if for all $\varepsilon > 0$, $\mathrm{P}(|Z_n - Z| \geq \varepsilon) \to 0$ as $n \to \infty$.

- One key approach to prove convergence almost surely/ in probability is to apply Markov's (or Chebychev's) inequality to obtain an upper bound of $\mathrm{P}(|Z_n - Z| \geq \varepsilon)$, and to show $\sum_n \mathrm{P}(|Z_n - Z| \geq \varepsilon) < \infty$ (for convergence almost surely) or $\mathrm{P}(|Z_n - Z| \geq \varepsilon) \to 0$ (for convergence in probability).

Weak and Strong Laws of Large Numbers Version 1

Theorem 1 (WLLN V1)

For a sequence of independent random variables $X_1, X_2, \dots$ *with the same mean* μ *and finite variance bounded by* σ^2, *define* $S_n = X_1 + X_2 + \cdots + X_n$. *Then* S_n/n *converges to* μ *in probability.*

Theorem 2 (SLLN V1)

For a sequence of independent random variables $X_1, X_2, \dots$ *with the same mean* μ *and bounded finite fourth central moments* ($\mathrm{E}(X_i - \mu)^4 \le a \le \infty$), *define* $S_n = X_1 + X_2 + \cdots + X_n$, *then* S_n/n *converges to* μ *almost surely.*

Strong Laws of Large Numbers Version 2

Theorem 3 (SLLN V2)

For a sequence of i.i.d. random variables $X_1, X_2, \ldots$ with the finite mean μ, define $S_n = X_1 + X_2 + \cdots + X_n$; then S_n/n converges to μ almost surely.

Corollary 4 (WLLN V2)

For a sequence of i.i.d. random variables $X_1, X_2, \ldots$ with the finite mean μ, define $S_n = X_1 + X_2 + \cdots + X_n$; then S_n/n converges to μ in probability.

The second version of WLLN follows from the fact that convergence almost surely implies convergence in probability.

Proof of SLLN V2: Part I

We now resume the proof to SLLN2 (started last lecture).

If you didn't show this on your own after last lecture, let's do Proposition 4.2.9: if $X \geq 0$, then $\sum_{k=1}^{\infty} \mathrm{P}(X \geq k) = E\lfloor X \rfloor$.

$$
\begin{aligned}
\sum_{k=1}^{\infty} \mathrm{P}(X \geq k) &= \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \mathrm{P}(k + l > X \geq k + l - 1) \\
&= \sum_{l=1}^{\infty} l\mathrm{P}(1 + l > X \geq 1 + l - 1) \\
&= \sum_{l=1}^{\infty} l\mathrm{P}(\lfloor X \rfloor = l) \\
&= \mathrm{E}[\lfloor X \rfloor]
\end{aligned}
$$

Proof OF SLLN V2: Part I

- First, without loss of generality, we may assume that $X \geq 0$. Otherwise, we can let $X_i = X_i^+ - X_i^-$, and apply the law of large number to X_i^+ and X_i^- respectively.
- Second, to prove almost sure convergence, the most reliable route is to use Chebchev's inequality to obtain an upper bound of $\mathrm{P}(|S_n/n - \mu| \geq \varepsilon)$ and then apply the Borel-Cantelli lemma to show that the probability of event $\{|S_n/n - \mu| \geq \varepsilon \ i.o.\}$ equals 0.
- However, the condition of applying Chebchev's inequality is that the variance of X_i exists. For this reason, we need to construct a truncated version of X_i.

Proof OF SLLN V2: Part II

- Let $Y_i = X_i 1_{X_i \leq i}$. Then $0 \leq Y_i \leq i$, $Y_i \leq X_i$, $\mathrm{E}(Y_i^k) \leq i^k < \infty$ for any k.

Lemma 5

Define $T_n = Y_1 + \cdots + Y_n$, if T_n/n converges to μ almost surely, S_n/n also converges to μ almost surely

- **Proof:** We only need to show that $(T_n - S_n)/n \to 0$ almost surely.
  - As $\sum_{k=1}^{\infty} \mathrm{P}(X_k \neq Y_k) = \sum_{k=1}^{\infty} \mathrm{P}(X_k > k) \leq \sum_{k=1}^{\infty} \mathrm{P}(X_1 \geq k) \leq \mathrm{E}(X_1) = \mu < \infty$ (see Proposition 4.2.9). By the Borel-Cantelli Lemma, $\mathrm{P}(X_k \neq Y_k \ i.o.) = 0$. Thus $\mathrm{P}(X_k - Y_k = 0 \ a.a) = 1$.
  - For any $\omega \in \{\omega : X_k(\omega) - Y_k(\omega) = 0 \ a.a\}$, there is an $N \in \mathrm{N}$ so that for any $n > N$, $X_n(\omega) = Y_n(\omega)$. Correspondingly, for $n > N$, $(T_n(\omega) - S_n(\omega))/n = \sum_{i=1}^{N}(Y_i(\omega) - X_i(\omega))/n \to 0$ as $n \to \infty$. Thus $\mathrm{P}(\lim_n (T_n - S_n)/n = 0) \geq \mathrm{P}(X_k - Y_k = 0 \ a.a) = 1$.

Proof OF SLLN V2: Part III

- Another trick we would like to use is to focus on a subsequence.

Lemma 6

For $\alpha > 1$, let $a_k = \lfloor \alpha^k \rfloor$, the greatest integer less than or equal to α^k. If for any $\alpha > 1$, T_{a_n}/a_n converges to μ almost surely, then T_n/n also converges to μ almost surely.

- **Proof:** For any k, we can find $n_k = n$ such that $a_n \leq k < a_{n+1}$:
  $\frac{a_n}{a_{n+1}} \frac{T_{a_n}}{a_n} = \frac{T_{a_n}}{a_{n+1}} \leq \frac{T_k}{k} \leq \frac{T_{a_{n+1}}}{a_n} = \frac{a_{n+1}}{a_n} \frac{T_{a_{n+1}}}{a_{n+1}}$
  - As $k \to \infty$, $\frac{a_n}{a_{n+1}} \to \frac{1}{\alpha}$ and $\frac{a_{n+1}}{a_n} \to \alpha$.
  - Goal:
    $\mu - \varepsilon \leq \frac{\mu}{(1+\delta)\alpha} \leq \frac{a_n}{a_{n+1}} \frac{T_{a_n}}{a_n} \leq \frac{T_k}{k} \leq \frac{a_{n+1}}{a_n} \frac{T_{a_{n+1}}}{a_{n+1}} \leq \mu(1+\delta)\alpha < \mu + \varepsilon$

Proof OF SLLN V2: Part III (continued)

- Goal:
  $\mu - \varepsilon \leq \frac{\mu}{(1+\delta)\alpha} \leq \frac{a_n}{a_{n+1}} \frac{T_{a_n}}{a_n} \leq \frac{T_k}{k} \leq \frac{a_{n+1}}{a_n} \frac{T_{a_{n+1}}}{a_{n+1}} \leq \mu(1+\delta)\alpha < \mu + \varepsilon$
- Pick $\varepsilon > 0$. Pick $\alpha > 1$ so that $\mu\alpha^2 < \mu + \varepsilon$. Then pick δ such that $(1+\delta) < \alpha$. These two together imply $\mu(1+\delta)\alpha < \mu + \varepsilon$.
- Pick N_1 such that $n > N_1$ implies $a_{n+1}/a_n < \alpha(1+\delta)^{1/2}$. Pick N_2 such that $n > N_2$ implies $T_{a_{n+1}}/a_{n+1} < \mu(1+\delta)^{1/2}$. Pick N_3 such that $n > N_3$ implies $T_{a_n}/a_n > \mu/(1+\delta)^{1/2}$

Proof OF SLLN V2: Part IV

- Here we will show that, for $a_k = \lfloor \alpha^k \rfloor$ $(\alpha > 1)$, T_{a_n}/a_n converges to μ almost surely.

- First, as $Y_n = X_n 1_{X_n \leq n}$, and X_is are i.i.d. random variables. $\mathrm{E}(Y_n) = \mathrm{E}(X_n 1_{X_n \leq n}) = \mathrm{E}(X_1 1_{X_1 \leq n}) \to \mathrm{E}(X_1) = \mu$ by the monotone convergence theorem.
- Second, as $n \to \infty$, $a_n \to \infty$, $\mathrm{E}(T_{a_n})/a_n = \sum_{i=1}^{a_n} \mathrm{E}(Y_i)/a_n \to \mu$. Thus, we only need to show $(T_{a_n} - \mathrm{E}(T_{a_n}))/a_n \to 0$ almost surely.
- Our goal is then to verify that for any $\varepsilon > 0$

$$\sum_{n=1}^{\infty} \mathrm{P}\left(\left| \frac{T_{a_n} - \mathrm{E}(T_{a_n})}{a_n} \right| \geq \varepsilon \right) \leq \sum_{n=1}^{\infty} \frac{\mathrm{Var}(T_{a_n})}{a_n^2 \varepsilon^2} < \infty$$

Proof OF SLLN V2: Part V

- To show $\sum_{n=1}^{\infty} \frac{\mathrm{Var}(T_{a_n})}{a_n^2 \varepsilon^2} < \infty$, note that:

$$\mathrm{Var}(T_{a_n}) = \sum_{k=1}^{a_n} \mathrm{Var}(Y_k) \leq \sum_{k=1}^{a_n} \mathrm{E}(Y_k^2)$$
$$= \sum_{k=1}^{a_n} \mathrm{E}(X_k^2 1_{X_k \leq k}) = \sum_{k=1}^{a_n} \mathrm{E}(X_1^2 1_{X_1 \leq k}) \leq a_n \mathrm{E}(X_1^2 1_{X_1 \leq a_n})$$

- So we have

$$\sum_{n=1}^{\infty} \frac{\mathrm{Var}(T_{a_n})}{a_n^2 \varepsilon^2} \leq \sum_{n=1}^{\infty} \frac{\mathrm{E}(X_1^2 1_{X_1 \leq a_n})}{a_n \varepsilon^2} = \frac{1}{\varepsilon^2} \mathrm{E}(X_1^2 \sum_{n=1}^{\infty} \frac{1}{a_n} 1_{a_n \geq X_1})$$

- We will show that $\sum_{n=1}^{\infty} \frac{1}{a_n} 1_{a_n \geq x} \leq \frac{2/x}{1-\alpha^{-1}}$, so that

$$\mathrm{E}(X_1^2 \sum_{n=1}^{\infty} \frac{1}{a_n} 1_{a_n \geq X_1}) \leq \mathrm{E}(X_1^2 \frac{2/X_1}{1-\alpha^{-1}}) = \mathrm{E}(\frac{2X_1}{1-\alpha^{-1}}) = \frac{2\mu}{1-\alpha^{-1}} < \infty$$

Proof OF SLLN V2: Part VI

- We still need to show that $\sum_{n=1}^{\infty} \frac{1}{a_n} 1_{a_n \geq x} \leq \frac{2/x}{1-\alpha^{-1}}$ for $a_k = \lfloor \alpha^k \rfloor$ $(\alpha > 1)$.
- We can verify that $a_n \geq \alpha^n/2$, then

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{1}{a_n} 1_{a_n \geq x} &= \sum_{a_n \geq x} \frac{1}{a_n} \leq \sum_{\alpha^n \geq x} \frac{1}{a_n} \leq \sum_{\alpha^n \geq x} \frac{2}{\alpha^n} \\
&\leq \sum_{k=0}^{\infty} \frac{2}{\alpha^k x} \\
&= \frac{2/x}{1-\alpha^{-1}}
\end{aligned}$$

Proof OF SLLN V2: Part VII

- **Summary** We first assume $X \geq 0$, then we define $Y_i = X_i 1_{X_i \leq i}$, then for $\alpha > 0$, we define the index of a subsequence as $a_k = \lfloor \alpha^k \rfloor$.

1) We show that $(T_{a_n} - \mathrm{E}(T_{a_n}))/a_n \to 0$ almost surely,
2) $T_{a_n}/a_n \to \mu$ almost surely.
3) $T_n/n \to \mu$ almost surely.
4) $S_n/n \to \mu$ almost surely.
5) For general X, $\sum_{i=1}^n X_i^+/n \to \mathrm{E}(X^+)$ and $\sum_{i=1}^n X_i^-/n \to \mathrm{E}(X^-)$ almost surely, then $\sum_{i=1}^n X_i/n \to \mu$ almost surely.