UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Discussion 06
Spring 2024

1. **The Weak Law of Large Numbers**

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with common mean μ and MGF M_X. We assume that $M_X(s)$ is finite when $s \in (-d, d)$ for some $d > 0$. Let

$$\bar{X}_n := \frac{X_1 + \ldots + X_n}{n}.$$

a. Show that the transform (or MGF) associated with $\bar{X}_n$ satisfies

$$M_{\bar{X}_n}(s) = M_X(s/n)^n.$$

b. Suppose that the transform $M_X(s)$ has a first-order Taylor series expansion around $s = 0$ of the form

$$M_X(s) = a + bs + o(s),$$

where $o(s)$ is a function that satisfies $\lim_{s\to 0} o(s)/s = 0$. Find a and b in terms of μ.

c. Show that for all $s \in (-d, d)$,

$$\lim_{n\to\infty} M_{\bar{X}_n}(s) = e^{\mu s}.$$

Hint: If $(a_n)_{n\in\mathbb{N}}$ is a sequence of real numbers converging to a, then $\lim_{n\to\infty}(1+\frac{a_n}{n})^n = e^a$.

d. Deduce that $\bar{X}_n \xrightarrow{d} \mu$. Note that the pointwise convergence of MGFs is equivalent to convergence in distribution.

Solution:

a. Let $S_n = X_1 + \cdots + X_n$. Then, as the sequence of random variables is i.i.d., we have that

$$M_{S_n}(s) = M_X(s)^n.$$

In addition,

$$M_{\bar{X}_n}(s) = M_{S_n/n}(s) = \mathbb{E}\left(\exp\left(s\frac{S_n}{n}\right)\right) = M_{S_n}(s/n) = M_X(s/n)^n.$$

b. $a = 1$ and $b = \mu$.

c. Taking the hint given,

$$\lim_{n\to\infty} M_{\bar{X}_n}(s) = \lim_{n\to\infty}\left(1 + \frac{\mu s + s\frac{o(1/n)}{1/n}}{n}\right)^n = e^{\mu s}.$$

d. Let $Y = \mu$ be a constant random variable, and observe that $M_Y(s) = e^{s\mu}$. Since $M_{\bar{X}_n} \to M_Y$ as $n \to \infty$, as the MGF uniquely determines the distribution, we can deduce that $\bar{X}_n \xrightarrow{d} Y = \mu$ as $n \to \infty$.

2. **Borel–Cantelli and the Strong Law**

In this problem, we walk through a proof of the strong law (assuming finite 4th moments) that relies only on basic probability. In class we covered the *Borel-Cantelli lemma*, which states that for events $(A_n)_{n=1}^{\infty}$, if $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, then

$$\mathbb{P}(A_n \text{ i.o.}) = 0,$$

where we define the event $\{A_n \text{ i.o.}\} = \cap_{n\geq 1} \cup_{m\geq n} A_n$ as the event where infinitely many A_n occur.

a. Let $X_1, X_2, \ldots$ be i.i.d. with $\mathbb{E}\, X_i = 0$ and $\mathbb{E}\, X_i^4 < \infty$ (and so we also have finite second and third moments). Let $S_n = X_1 + \cdots + X_n$, and compute $\mathbb{E}[S_n^4]$. Write your answer in terms of the moments $\mathbb{E}[X_i^2], \mathbb{E}[X_i^3], \mathbb{E}[X_i^4]$.

b. Fix an $\varepsilon > 0$, and use Markov's inequality to show that, for any n,

$$\mathbb{P}(|S_n/n| > \varepsilon) \leq O(n^{-2}).$$

c. Finally, use Borel-Cantelli to conclude that $\mathbb{P}(\lim_{n\to\infty} S_n/n = 0) = 1$. This a weaker (the full theorem assumes only finite first moments) form of the *strong law of large numbers*.

Solution:

a. We expand:

$$\mathbb{E}\, S_n^4 = \mathbb{E}\left(\sum_{i=1}^{n} X_i\right)^4 = \mathbb{E} \sum_{1\leq i,j,k,l\leq n} X_i X_j X_k X_l.$$

Terms of the form $\mathbb{E}[X_i^3 X_j]$, $\mathbb{E}[X_i^2 X_j X_k]$, and $\mathbb{E}[X_i X_j X_k X_l]$ are just 0 by independence. We are left with

$$\mathbb{E}\left(\sum_{i=1}^{n} X_i^4\right) + \mathbb{E}\left[\sum_{i\neq j} X_i^2 X_j^2\right] = n\,\mathbb{E}[X_1^4] + 3n(n-1)\,\mathbb{E}[X_1^2]\,\mathbb{E}[X_2^2].$$

b. By Markov's inequality and the previous part, we have

$$\mathbb{P}(|S_n/n| > \varepsilon) < \varepsilon^{-4}\,\mathbb{E}(S_n/n)^4 = O(\varepsilon^{-4} n^{-2}).$$

c. Letting $A_n = \{|S_n/n| > \varepsilon\}$, we get from the Borel-Cantelli lemma that $\mathbb{P}(|S_n/n| > \varepsilon \text{ i.o.}) = 0$. Since ε is arbitrary, this implies almost sure convergence.

3. **The CLT Implies the WLLN**

   a. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables. Show that if X_n converges in distribution to a constant c, then X_n converges in probability to c.

   b. Now let $(X_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables with mean μ and finite variance σ^2. Show that the CLT implies the WLLN: that is,

$$\frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n}(X_i-\mu)\xrightarrow{\mathsf{d}} Z\sim\mathcal{N}(0,1) \implies \frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{\mathbb{P}}\mu,$$

   where $\xrightarrow{\mathsf{d}}$ is short for "converges in distribution" and $\xrightarrow{\mathbb{P}}$ for "converges in probability."

Solution:

   a. Since X_n converges in distribution to c, we know that for all $\varepsilon>0$,

$$\lim_{n\to\infty}F_{X_n}(c-\varepsilon)=F_c(c-\varepsilon)=0$$
$$\lim_{n\to\infty}F_{X_n}(c+\tfrac{\varepsilon}{2})=F_c(c+\tfrac{\varepsilon}{2})=1.$$

   Using these limits, we have convergence in probability:

$$\begin{aligned}\lim_{n\to\infty}\mathbb{P}(|X_n-c|\ge\varepsilon)&=\lim_{n\to\infty}\mathbb{P}(X_n\le c-\varepsilon)+\lim_{n\to\infty}\mathbb{P}(X_n\ge c+\varepsilon)\\&\le\lim_{n\to\infty}\mathbb{P}(X_n\le c-\varepsilon)+\lim_{n\to\infty}\mathbb{P}(X_n> c+\tfrac{\varepsilon}{2})\\&=\lim_{n\to\infty}F_{X_n}(c-\varepsilon)+\lim_{n\to\infty}1-F_{X_n}(c+\tfrac{\varepsilon}{2})\\&=0+1-1=0.\end{aligned}$$

   (The reason we take $c+\frac{\varepsilon}{2}$ instead of $c+\varepsilon$ is because $1-F_{X_n}(x)=\mathbb{P}(X_n>x)$, but we have $\mathbb{P}(X_n\ge c+\varepsilon)$, which is not a strict inequality.)

   b. From the CLT, we know that

$$Z_n:=\frac{\sqrt{n}}{\sigma}\left(\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right)\text{ converges to } Z\sim\mathcal{N}(0,1)\text{ in distribution.}$$

   Additionally, $a_n:=\frac{\sigma}{\sqrt{n}}\to 0$. Then $Y_n:=a_nZ_n=\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\to 0$ in distribution. By part a, since $c=0$ is a constant, Y_n also converges to 0 in probability. In other words,

$$\frac{1}{n}\sum_{i=1}^{n}X_i\to\mu\text{ in probability,}$$

   which is precisely the Weak Law of Large Numbers.

   Note. The claim that "if $Z_n\to Z$ in distribution and $a_n\to 0$ as constants, then $a_nZ_n\to 0$ in distribution" requires proof, which we present below.

   For $x<0$ and any $N\ge 1$, we know that $\frac{x}{a_n}\le -N$ eventually, so

$$\lim_{n\to\infty}\mathbb{P}(a_nZ_n\le x)=\lim_{n\to\infty}\mathbb{P}(Z_n\le\tfrac{x}{a_n})\le\lim_{n\to\infty}\mathbb{P}(Z_n\le -N)=\mathbb{P}(Z\le -N).$$

The left-hand side does not depend on N, so taking the limit as $N \to \infty$ of both sides, by continuity from above, we find that

$$\lim_{n\to\infty} \mathbb{P}(a_n Z_n \leq x) \leq \mathbb{P}(Z = -\infty) = 0.$$

Similarly, for $x > 0$, we know that $\frac{x}{a_n} \geq N$ eventually for any N, so

$$\lim_{n\to\infty} \mathbb{P}(a_n Z_n > x) \leq \lim_{n\to\infty} \mathbb{P}(Z_n > N) = \mathbb{P}(Z > N),$$

and taking the limit as $N \to \infty$, we find that $\lim_{n\to\infty} \mathbb{P}(a_n Z_n > x) \leq 0$, or $\lim_{n\to\infty} \mathbb{P}(a_n Z_n \leq x) = 1$. In other words, we have shown that

$$\lim_{n\to\infty} \mathbb{P}(a_n Z_n \leq x) = \mathbb{1}\{0 \leq x\},$$

i.e. $a_n Z_n$ converges to 0 in distribution. This is a specific case of a more general result called *Slutsky's theorem.*