Lecture-16: Weak convergence of random variables

1 Convergence in distribution

Definition 1.1. A sequence $(X_n : n \in \mathbb{N})$ of random variables **converges in distribution** to a random variable X if

$$\lim_n F_{X_n}(x) = F_X(x)$$

at all continuity points x of F_X. Convergence in distribution is denoted by $\lim_n X_n = X$ in distribution.

Proposition 1.2. *Let $(X_n : n \in \mathbb{N})$ be a sequence of random variables and let X be a random variable. Then the following are equivalent:*

(a) $\lim_n X_n = X$ *in distribution.*

(b) $\lim_n \mathbb{E}[g(X_n)] = \mathbb{E}[g(X)]$ *for any bounded continuous function g.*

(c) Characteristic functions converge point-wise, i.e. $\lim_n \Phi_{X_n}(u) \to \Phi_X(u)$ *for each $u \in R$.*

Proof. Let $(X_n : n \in \mathbb{N})$ be a sequence of random variables and let X be a random variable. We will show that $(a) \implies (b) \implies (c) \implies (a)$.

$(a) \implies (b)$: Let $\lim_n X_n = X$ in distribution, then $\lim_n \int_{x \in \mathbb{R}} g(x) dF_{X_n}(x) = \int_{x \in \mathbb{R}} g(x) \lim_n dF_{X_n}(x)$ by the bounded convergence theorem for any bounded continuous function g.

$(b) \implies (c)$: Let $\lim_n \mathbb{E}[g(X_n)] = \mathbb{E}[g(X)]$ for any bounded continuous function g. Taking $g(x) = e^{jux}$, we get the result.

$(c) \implies (a)$: The proof of this part is technical and is omitted.

□

Example 1.3 (Convergence in distribution but not in probability). Consider a sequence of non-degenerate continuous *i.i.d.* random variables $(X_n : n \in \mathbb{N})$ and independent random variable X with the common distribution F_X. Then $F_{X_n} = F_X$ for all $n \in \mathbb{N}$, and hence $\lim_n X_n = X$ in distribution. However, for any $n \in \mathbb{N}$ and $\epsilon > 0$, from the monotonicity of distribution function, we have

$$P\{|X_n - X| > \epsilon\} = \mathbb{E}\mathbb{1}_{\{X_n \notin [X-\epsilon, X+\epsilon]\}} = \mathbb{E}F_X(X+\epsilon) - \mathbb{E}F_X(X-\epsilon) > 0.$$

Lemma 1.4 (Convergence in probability implies in distribution). *Consider a sequence $(X_n : n \in \mathbb{N})$ of random variables and a random variable X, such that* $\lim_n X_n = X$ *in probability, then* $\lim_n X_n = X$ *in distribution.*

Proof. Fix $\epsilon > 0$, and consider the event $E_n \triangleq \{\omega \in \Omega : |X_n - X|(\omega) > \epsilon\} = \{X_n \notin [X-\epsilon, X+\epsilon]\} \in \mathcal{F}$. We further define events $A_n(x) \triangleq \{X_n \leqslant x\}$ and $A(x) \triangleq \{X \leqslant x\}$, then we can write

$$A_n(x) \cap A(x+\epsilon) \subseteq A(x+\epsilon), \qquad A_n(x) \cap A^c(x+\epsilon) \subseteq E_n,$$
$$A(x-\epsilon) \cap A_n(x) \subseteq A_n(x), \qquad A(x-\epsilon) \cap A_n^c(x) \subseteq E_n.$$

From the above set relations, law of total probability, and union bound, we have

$$F(x-\epsilon) - P(E_n) \leqslant F_n(x) \leqslant F(x+\epsilon) + P(E_n).$$

From the convergence in probability, we have $\lim_n P(E_n) = 0$ get

$$F(x-\epsilon) \leqslant \liminf_n F_n(x) \leqslant \limsup_n F_n(x) \leqslant F(x+\epsilon).$$

We get the result at the continuity points of F_X, since the choice of ϵ was arbitrary. □

Theorem 1.5 (Central Limit Theorem). *Consider an* i.i.d. *random sequence $(X_n : n \in \mathbb{N})$ with $\mathbb{E}X_n = \mu$ and* $\mathrm{Var}(X_n) = \sigma^2$, *defined on the probability space $(\Omega, \mathcal{F}, P)$. We define the n-sum as $S_n = \sum_{i=1}^n X_i$ and consider a standard normal random variable Y with density function $f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ for all $y \in \mathbb{R}$. Then,*

$$\lim_n \frac{S_n - n\mu}{\sigma\sqrt{n}} = Y \textit{ in distribution.}$$

Proof. The classical proof is using the characteristic functions. Let $Z_i \triangleq \frac{X_i - \mu}{\sigma}$ for all $i \in \mathbb{N}$, then the shifted and scaled n-sum $\frac{S_n - n\mu}{\sigma\sqrt{n}} = \sum_{i=1}^n Z_i$ We use the third equivalence in Proposition 1.2 to show that the characteristic function of converges to the characteristic function of the standard normal. We define the characteristic functions

$$\Phi_n(u) \triangleq \mathbb{E}\exp\left(ju\frac{(S_n - n\mu)}{\sigma\sqrt{n}}\right), \qquad \Phi_{Z_i}(u) \triangleq \mathbb{E}\exp(juZ_i), \qquad \Phi_Y(u) \triangleq \mathbb{E}\exp(juY).$$

We can compute the characteristic function of the standard normal as

$$\Phi_Y(u) = \frac{1}{\sqrt{2\pi}} \int_{y \in \mathbb{R}} e^{-\frac{u^2}{2}} \exp\left(-\frac{(y-ju)^2}{2}\right) dy = e^{-\frac{u^2}{2}}.$$

Since $(Z_i : n \in \mathbb{N})$ is a zero mean *i.i.d.* sequence, and using the Taylor expansion of the characteristic function, we have

$$\Phi_n(u) = \prod_{i=1}^n \mathbb{E}\exp\left(ju\frac{(X_i - \mu)}{\sigma\sqrt{n}}\right) = \left[\Phi_{Z_1}\left(\frac{u}{\sqrt{n}}\right)\right]^n = \left[1 - \frac{u^2}{2n} + o\left(\frac{u^2}{n}\right)\right]^n.$$

For any $u \in \mathbb{R}$, taking limit $n \in \mathbb{N}$, we get the result. □

2 Strong law of large numbers

Definition 2.1. For a random sequence $(X_n : n \in \mathbb{N})$ with bounded mean $\mathbb{E}|X_n| < \infty$, we define the n-sum as $S_n \triangleq \sum_{i=1}^n X_i$ and the empirical n-mean $\frac{S_n}{n}$ for each $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, we define event

$$E_n \triangleq \{\omega \in \Omega : |S_n - \mathbb{E}S_n| > n\epsilon\} \in \mathcal{F}.$$

Theorem 2.2 (L^4 strong law of large numbers). *Let $(X_n : n \in \mathbb{N})$ be a sequence of pair-wise uncorrelated random variables with bounded mean $\mathbb{E}X_n$ and uniformly bounded fourth central moment $\mathbb{E}(X_n - \mathbb{E}X_n)^4 \leqslant B < \infty$ for all $n \in \mathbb{N}$. Then, the empirical n-mean converges to $\lim_n \frac{\mathbb{E}S_n}{n}$ almost surely.*

Proof. From the Markov's and Minkowski's inequality, we have $P(E_n) \leqslant \frac{\sum_{i=1}^n \mathbb{E}(X_i - \mu)^4}{n^4\epsilon^2} \leqslant \frac{B + \mu^4}{n^3\epsilon^2}$. It follows that the $\sum_{n \in \mathbb{N}} P(E_n) < \infty$, and hence by Borel Canteli Lemma, we have $P\{E_n^c \text{ for all but finitely many } n\} = 1$. Since, the choice of ϵ was arbitrary, the result follows. □

Theorem 2.3 (L^2 strong law of large numbers). *Let $(X_n : n \in \mathbb{N})$ be a sequence of pair-wise uncorrelated random variables with mean $\mathbb{E}X_n$ and uniformly bounded variance* $\mathrm{Var}(X_n) \leqslant B < \infty$ *for all $n \in \mathbb{N}$. Then, the empirical n-mean converges to $\lim_n \frac{\mathbb{E}S_n}{n}$ almost surely.*

Proof. For each $n \in \mathbb{N}$, we define events $F_n \triangleq E_{n^2}$, and

$$G_n \triangleq \left\{ \max_{n^2 \leqslant k < (n+1)^2} |S_k - S_{n^2} - \mathbb{E}(S_k - S_{n^2})| > n^2 \epsilon \right\} = \bigcup_{n^2 \leqslant k < (n+1)^2} \left\{ \omega \in \Omega : \left| \sum_{i=n^2}^{k} X_i - \mathbb{E} X_i \right| > n^2 \epsilon \right\}.$$

From the Markov's inequality and union bound, we have

$$P(F_n) \leqslant \frac{\sum_{i=1}^{n^2} \mathrm{Var}(X_i)}{n^4 \epsilon^2} \leqslant \frac{B}{n^2 \epsilon^2}, \qquad P(G_n) \leqslant \sum_{k=0}^{2n} \frac{kB}{n^4 \epsilon} \leqslant \frac{(2n+1)B}{n^3 \epsilon}.$$

Therefore, $\sum_{n \in \mathbb{N}} P(F_n) < \infty$ and $\sum_{n \in \mathbb{N}} P(G_n) < \infty$, and hence by Borel Canteli Lemma, we have

$$\lim_n \frac{S_{n^2} - \mathbb{E} S_{n^2}}{n^2} = \lim_n \frac{S_k - S_{n^2} - \mathbb{E}(S_k - S_{n^2})}{n^2} = 0 \text{ a.s.}$$

The result follows from the fact that for any $k \in \mathbb{N}$, there exists $n \in \mathbb{N}$ such that $k \in \left\{ n^2, \dots, (n+1)^2 - 1 \right\}$ and hence

$$\frac{|S_k - \mathbb{E} S_k|}{k} \leqslant \left(\frac{|S_{n^2} - \mathbb{E} S_{n^2}|}{n^2} + \frac{|S_k - S_{n^2} - \mathbb{E}(S_k - S_{n^2})|}{n^2} \right).$$

□

Theorem 2.4 (L^1 strong law of large numbers). *Let $(X_n : n \in \mathbb{N})$ be a sequence of pair-wise uncorrelated random variables such that $\mathbb{E}|X_n| \leqslant B < \infty$ for all $n \in \mathbb{N}$. Then, the empirical n-mean converges to $\lim_n \frac{\mathbb{E} S_n}{n}$ almost surely.*