Lecture 16: November 16

Lecturer: Persi Diaconis   *Scribes: Kenneth Tay*

16.1 Weak Convergence

Definition 16.1 *If F_n, F are distribution functions on $\mathbb{R}$, then F_n **converges weakly** to F if $F_n(x) \to F(x)$ for all continuity points x of F.*

If F_n converges weakly to F, we write $F_n \Rightarrow F$.

Example: Take F_n corresponding to a point mass at $\frac{1}{n}$, F corresponding to a point mass at 0. Then $\sup_x |F_n(x) - F(x)| = 1$ for all n, but $F_n \Rightarrow F$.

Below is a lemma which we will use at some point:

Lemma 16.2 *Let $\mathcal{C}_b^\infty = \{f : f : \mathbb{R} \to \mathbb{R}, f \text{ has bounded derivatives of all order}\}$.*

If F_n, F are distribution functions on $\mathbb{R}$ such that $F_n \Rightarrow F$, then

$$\int_{-\infty}^{\infty} f \, dF_n \to \int_{-\infty}^{\infty} f \, dF$$

for all $f \in \mathcal{C}_b^\infty$. The converse is true as well.

Proof: We will just prove the converse statement.

Want $\mathbb{E}_{F_n}(g) \to \mathbb{E}_F(g)$, where $g(x) = 1$ if $x \le x_0$, 0 otherwise.

Let

$$\psi(x) = \begin{cases} 1 & \text{if } x \le 0, \\ 0 & \text{if } x \ge 1, \\ \frac{1}{z} \int_x^1 \exp\left[-\frac{1}{s(1-s)}\right] ds & \text{if } 0 < x < 1, \end{cases}$$

where $z = \int_0^1 \exp\left[-\frac{1}{s(1-s)}\right] ds$. It can be checked that ψ is differentiable of all orders.

For $u > 0$, let $\psi_u(x - y) = \psi(u(x - y))$. As a function of x, we have

$$\psi_u(x - y) = \begin{cases} 1 & \text{if } x \le y, \\ 0 & \text{if } x \ge y + \frac{1}{u}, \\ \text{``smooth''} & \text{if } y < x < y + \frac{1}{u}. \end{cases}$$

Fix y. Note that

$$F_n(y) = \int_{-\infty}^{y} F_n(dx) = \int_{-\infty}^{\infty} \delta_{(-\infty, y]} F_n(dx) \le \int_{-\infty}^{\infty} \psi_u(x - y) F_n(dx)$$

for all n and u. Hence,

$$\limsup_n F_n(y) \leq \limsup_n \int_{-\infty}^{\infty} \psi_u(x - y) F_n(dx)$$

$$= \int_{-\infty}^{\infty} \psi_u(x - y) F(dx)$$

$$\leq \int_{-\infty}^{\infty} \delta_{(-\infty, y + \frac{1}{u}]} F(dx)$$

$$= F\left(y + \frac{1}{u}\right).$$

By a similar argument using $\psi_u\left(x - y + \frac{1}{u}\right)$, we can obtain

$$\liminf_n F_n(y) \geq F\left(y - \frac{1}{u}\right).$$

Hence, for all $u > 0$,

$$F\left(y - \frac{1}{u}\right) \leq \liminf_n F_n(y) \leq \limsup_n F_n(y) \leq F\left(y + \frac{1}{u}\right).$$

If x is a continuity point, $\lim_{u \to \infty} F\left(y - \frac{1}{u}\right) = \lim_{u \to \infty} F\left(y + \frac{1}{u}\right) = F(y)$, implying that $F_n(y) \to F(y)$, as required.

■

16.2   Central Limit Theorem

The gist of the Central Limit Theorem is as follows: Let $\{X_i\}_{i=1}^{\infty}$ be random variables with finite mean μ_i and finite variance σ_i^2. If they are "not too dependent" and "no few terms dominate", then for every $x \in \mathbb{R}$,

$$P\left\{\frac{S_n - \sum_{i=1}^{n} \mu_i}{s_n} \leq x\right\} \to \Phi(x),$$

where $s_n^2 = \sum_{i=1}^{n} \sigma_i^2$, and $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{t^2}{2}\right] dt$.

There are several versions of the Central Limit Theorem; in this lecture, we will go through Lindeberg's version.

Definition 16.3 *A **triangular array** is a series of random variables $\{X_{ni}\}$ where the random variables in the nth row are $X_{n1}, X_{n2}, \ldots, X_{nk_n}$.*

Theorem 16.4 (Lindeberg's CLT) *Let $\{X_{ni}\}$ be a triangular array. Suppose that $\mathbb{E}(X_{ni}) = 0$ for all n and i, and that $Var\, X_{ni} = \sigma_{ni}^2 < \infty$. Let $S_n = \sum_{i=1}^{k_n} X_{ni}$, $s_n^2 = \sum_{i=1}^{k_n} \sigma_{ni}^2$ (i.e. sum of the nth row and its variance.)*

Suppose for every n, $\{X_{ni}\}_{i=1}^{k_n}$ is independent. Suppose also that the Lindeburg condition holds: i.e. for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{i=1}^{k_n} \int_{\{|X_{ni}|>\varepsilon s_n\}} X_{ni}^2 dP = 0.$$

Then for every $x \in \mathbb{R}$,

$$P\left\{\frac{S_n}{s_n} \le x\right\} \to \Phi(x).$$

16.2.1 Example: iid setting

Suppose X_i, $1 \le i < \infty$ are iid, with mean 0 and finite variance σ^2. Prove that

$$P\left\{\frac{S_n}{\sigma\sqrt{n}} \le x\right\} \to \Phi(x).$$

Proof: Make a triangular array

$$X_1$$
$$X_1, X_2$$
$$X_1, X_2, X_3$$
$$\vdots$$

We have $s_n^2 = n\sigma^2$. For fixed $\varepsilon > 0$,

$$\frac{1}{s_n^2} \sum_{i=1}^{k_n} \int_{\{|X_{ni}|>\varepsilon s_n\}} X_{ni}^2 dP = \frac{1}{n\sigma^2} \cdot n \cdot \int_{\{|X_1|>\varepsilon\sigma\sqrt{n}\}} X_1^2 dP$$

$$= \frac{1}{\sigma^2} \int_{\{|X_1|>\varepsilon\sigma\sqrt{n}\}} X_1^2 dP$$

$$\to 0$$

as $n \to \infty$. Hence, Lindeberg's condition holds and we can conclude by using Lindeberg's CLT. ∎

16.2.2 Example: Card guessing, scoring system 1

Recall the card guessing set-up with complete feedback (i.e. guessing cards one by one, I tell you what the card actually was). For a deck of n cards, let Y_{ni} be 1 if you guess the ith card correctly, 0 otherwise, i.e.

$$Y_{ni} = \begin{cases} 1 & \text{with probability } \frac{1}{n-i+1}, \\ 0 & \text{otherwise.} \end{cases}$$

Let $X_{ni} = Y_{ni} - \frac{1}{n-i+1}$ so that $\mathbb{E}X_{ni} = 0$. We can compute

$$\sigma_{ni}^2 = \frac{1}{n-i+1}\left(1 - \frac{1}{n-i+1}\right),$$

$$s_n^2 = \sum_{i=1}^n \frac{1}{i}\left(1 - \frac{1}{i}\right)$$

$$= \log n - \gamma - \pi^2/6 + O(1/n)$$

$$\sim \log n.$$

Let's check that the Lindeburg condition holds: since X_{ni} is bounded by 1, for fixed $\varepsilon > 0$,

$$\frac{1}{s_n^2}\sum_{i=1}^{k_n}\int_{\{|X_{ni}|>\varepsilon s_n\}} X_{ni}^2 dP \approx \frac{1}{\log n}\sum_{i=1}^n\int_{\{|X_{ni}|>\varepsilon\sqrt{\log n}\}} X_{ni}^2 dP$$

$$= 0$$

exactly for large n.

Therefore, we can apply Lindeberg's CLT to obtain

$$P\left\{\frac{S_n}{\log n} \le x\right\} \to \Phi(x).$$

16.2.3   Example: Card guessing, scoring system 2

As in the previous example, we have a deck of cards labeled 1 to n, you try to guess the cards. In the previous example, I gave you a score of 1 for each card you got right. In this example, I give you a score equal to the reciprocal of the probability of getting the card right. (This makes the game fair.) Let

$$Y_{ni} = \begin{cases} i & \text{with probability } \frac{1}{i}, \\ 0 & \text{otherwise.} \end{cases}$$

Let $X_{ni} = Y_{ni} - 1$ so that $\mathbb{E}X_{ni} = 1$. We can compute

$$\sigma_{ni}^2 = (i-1)^2 \cdot \frac{1}{i} + (-1)^2\left(1 - \frac{1}{i}\right) = i - 1,$$

$$s_n^2 = \sum_{i=1}^n (i-1) = \frac{n(n-1)}{2}.$$

Let's check if Lindeberg's condition holds:

$$\frac{1}{s_n^2}\sum_{i=1}^{k_n}\int_{\{|X_{ni}|>\varepsilon s_n\}} X_{ni}^2 dP = \frac{2}{n(n-1)}\sum_{i=1}^n\int_{\{|X_{ni}|>\varepsilon\sqrt{n(n-1)/2}\}} X_{ni}^2 dP$$

$$\ge \frac{2}{n^2}\sum_{i>n\varepsilon}\frac{(i-1)^2}{i}$$

which does not go to zero. Hence, the Lindeberg condition does not hold.

In fact, the CLT fails in this case. In turns out that $P\{S_n/n \leq x\} \to G(x)$ where G is some (interesting) distribution function, but which is *not* Gaussian.

16.2.4 Example: Does someone have ESP?

Consider a model where

$$X_i = \begin{cases} 1 & \text{with probability } \dfrac{1}{i^\theta}, \\ 0 & \text{otherwise}, \end{cases}$$

where $0 \leq \theta \leq 1$. The case of $\theta = 1$ corresponds to the model in Section 16.2.2, i.e. guessing cards with complete feedback. The case of $\theta < 1$ corresponds to the setting where someone has greater than the usual odds of guessing the cards right (e.g. due to ESP).

In testing $H_0 : \theta = 1$ vs. $H_1 : \theta < 1$, we look at the likelihood ratio

$$\frac{P_\theta(\text{data})}{P_0(\text{data})} = \frac{\prod_{i=1}^n (1/i^\theta)^{X_i}(1-1/i^\theta)^{1-X_i}}{\prod_{i=1}^n (1/i)^{X_i}(1-1/i)^{1-X_i}}$$

$$= \prod_{i=1}^n \frac{(i^\theta - 1)^{1-X_i}}{i^\theta} \cdot \frac{i}{(i-1)^{1-X_i}}$$

$$= C \prod_{i=1}^n \left(\frac{i-1}{i^\theta - 1} \right)^{X_i}.$$

We reject the null hypothesis if the ratio is large, and don't reject if the ratio is small.

We could also look at the log-likelihood instead:

$$\log\left(\frac{P_\theta(\text{data})}{P_0(\text{data})} \right) = \log C + \sum_{i=1}^n X_i \log\left(\frac{i-1}{i^\theta - 1} \right)$$

$$= \log C + \sum_{i=1}^n X_i w_i,$$

where $w_i = \log\left(\dfrac{i-1}{i^\theta - 1} \right)$. We can check that the Lindeberg condition holds, and so we can apply CLT here.