

STAT 7200

Introduction to Advanced Probability

Lecture 23

Taylor R. Brown

1 Weak Convergence

- Characteristic Function
- Continuity Theorem
- Method of Moments
- Central Limit Theorem

Central Limit Theorem

Theorem 1 (Central Limit Theorem)

If X_1, X_2, \dots , are i.i.d. random variables with finite mean μ and finite variance σ^2 . Let $S_n = X_1 + X_2 + \dots + X_n$. Then as $n \rightarrow \infty$, $\mathcal{L}(\frac{S_n - n\mu}{\sqrt{n\sigma^2}})$ converges weakly to $N(0, 1)$. If we use $\Phi(x)$ to represent the CDF of $N(0, 1)$, then the CLT can be written as:
for all $x \in \mathbb{R}$, $\lim_{n \rightarrow \infty} P(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq x) = \Phi(x)$.

- The proof uses the continuity theorem. The following facts are important:

- 1) The characteristic function of $N(0, 1)$ is $e^{-\frac{t^2}{2}}$.
- 2) For any random variables X with finite second moments:

$$\phi(t) = 1 + itE(X) - \frac{t^2}{2}E(X^2) + o(|t^2|).$$

where “small o” notation $o(|t^2|)$ represents a certain quantity so that $o(|t^2|)/(|t^2|) \rightarrow 0$ as $t \rightarrow 0$.

Central Limit Theorem: Proof

- **Proof:** First, without loss of generality, we assume $\mu = 0$, $\sigma^2 = 1$. Then our task is to show that, the distribution of $\frac{S_n - n\mu}{\sqrt{n\sigma^2}} = \frac{S_n}{\sqrt{n}}$ converges weakly to $N(0, 1)$ distribution. And it would be sufficient to show $\phi_{S_n/\sqrt{n}}(t) \rightarrow e^{-t^2/2}$ for all $t \in \mathbb{R}$.
- Since X_1, X_2, \dots are i.i.d, let $\phi(t) = E(e^{itX_1})$, then

$$\begin{aligned}\phi_{S_n/\sqrt{n}}(t) &= \prod_{i=1}^n \phi_{X_i/\sqrt{n}} = [\phi(t/\sqrt{n})]^n \\ &= \left[1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right]^n \\ &\rightarrow e^{-\frac{t^2}{2}}\end{aligned}$$

Convergence Speed of CLT

- When we discussed the law of large numbers, we considered large deviations theory, which studies how rapidly the average converges to the mean. The answer is exponentially fast, as long as the moment generating function exists.
- Here we can also discuss similar issues for the CLT. What conclusion can we draw for $P\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq x\right) - \Phi(x)$?

Convergence Speed of CLT

Theorem 2 (Berry-Esseen Theorem)

Assume X_1, X_2, \dots are i.i.d. random variables with a finite mean μ , finite variance σ^2 and a finite third central moment. If we denote $\rho = E|X_1 - \mu|^3$, and $S_n = X_1 + X_2 + \dots + X_n$, then as $n \rightarrow \infty$,

$$\sup_x \left| P\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq x\right) - \Phi(x) \right| \leq \frac{C\rho}{\sigma^3\sqrt{n}},$$

where C is a constant $((2\pi)^{-1/2} \leq C < 0.8)$

Convergence Speed of CLT: Example

- For instance, in practice, we often approximate a $\text{Bin}(n, p)$ with $N(np, np(1-p))$. This is because we can regard $S_n \sim \text{Bin}(n, p)$ as the sum of n i.i.d. $\text{Bern}(p)$ random variables. Then by the central limit theorem, we have:

$$P\left(\frac{S_n - np}{\sqrt{np(1-p)}} \leq x\right) \rightarrow \Phi(x).$$

- Now if we apply Berry-Essen Theorem, the third central moment of $\text{Bern}(p)$ is $\rho = (1-p)^3 p + p^3(1-p) = p(1-p)[p^2 + (1-p)^2]$. Then we have:

$$\sup_x \left| P\left(\frac{S_n - np}{\sqrt{np(1-p)}} \leq x\right) - \Phi(x) \right| \leq C \frac{p^2 + (1-p)^2}{\sqrt{np(1-p)}}.$$

- Then above result suggests that the approximation of the CDF of $\frac{S_n - np}{\sqrt{np(1-p)}}$ by the standard normal CDF can be poor when p or $1-p$ is small (the distribution of $\text{Bin}(n, p)$ is very asymmetric). In these case, it would usually better to approximate the Binomial distribution with Poisson distribution.

Extension of Central Limit Theorem: Triangular Array

- The central limit theorem we studied requires that all the random variables are i.i.d., which are very limited. Here we will discuss the possible ways to relax this condition. First, we will introduce the concept of "triangular array" that is used to reformulate the problem.
- **Triangular Arrays:** A collection of random variables $\{Z_{nk} : n \geq 1, 1 \leq k \leq r_n\}$ so that the r.v.s. in each row are independent.

$$\begin{array}{ccccccc} Z_{11}, & \cdots, & & Z_{1r_1} & & & \\ Z_{21}, & \cdots, & & \cdots, & & Z_{2r_2} & \\ \cdots & & & & & & \\ Z_{n1}, & \cdots, & & \cdots, & & \cdots & Z_{nr_n} \\ \cdots & & & & & & \end{array}$$

- We will assume that $E(Z_{nk}) = 0$ and $Var(Z_{nk}) = \sigma_{nk}^2$. Set $S_n = Z_{n1} + \cdots + Z_{nr_n}$, and $s_n^2 = Var(S_n) = \sigma_{n1}^2 + \cdots + \sigma_{nr_n}^2$. Then the question is, under what condition, the distribution of S_n/s_n would converge to $N(0, 1)$?

Lindeberg and Lyapunov Condition

Theorem 3 (Lindeberg Central Limit Theorem)

For triangular array $\{Z_{nk}\}$, if the following Lindeberg condition holds:

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{k=1}^{r_n} E(Z_{nk}^2 1_{|Z_{nk}| \geq \varepsilon s_n}) = 0,$$

for any $\varepsilon > 0$, then $\mathcal{L}(S_n/s_n)$ converges weakly to $N(0, 1)$.

Theorem 4 (Lyapunov Central Limit Theorem)

For triangular array $\{Z_{nk}\}$, if for some $\delta > 0$

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{r_n} \frac{E(Z_{nk}^{2+\delta})}{s_n^{2+\delta}} = 0,$$

then $\mathcal{L}(S_n/s_n)$ converges weakly to $N(0, 1)$.

Lindeberg Condition and Lyapunov Condition: Example

- **Example:** Let Y_1, Y_2, \dots be i.i.d. random variables with mean 0 and variance 1, and $\{c_n\}$ be a constant sequence. If $u_n = \frac{\max_{1 \leq k \leq n} c_k^2}{c_1^2 + \dots + c_n^2} \rightarrow 0$ as $n \rightarrow \infty$, then $\mathcal{L}\left(\frac{\sum_{k=1}^n c_k Y_k}{\sqrt{c_1^2 + \dots + c_n^2}}\right)$ converges weakly to $N(0, 1)$.
- We can construct triangular array $\{Z_{nk}\}$ so that for $1 \leq k \leq n$, $Z_{nk} = c_k Y_k$, then $S_n = \sum_{k=1}^n c_k Y_k$ and $s_n^2 = c_1^2 + \dots + c_n^2$. Then the Lindeberg condition can be verified:

$$\begin{aligned} \frac{1}{s_n^2} \sum_{k=1}^{r_n} E(Z_{nk}^2 1_{|Z_{nk}| \geq \varepsilon s_n}) &= \frac{1}{s_n^2} \sum_{k=1}^n E(c_k^2 Y_k^2 1_{c_k^2 |Y_k|^2 \geq \varepsilon^2 s_n^2}) \\ &\leq \frac{1}{s_n^2} \sum_{k=1}^n c_k^2 E(Y_k^2 1_{|Y_k|^2 \geq \varepsilon^2 \frac{1}{u_n}}) \\ &= \frac{\sum_{k=1}^n c_k^2}{s_n^2} E(Y_1^2 1_{|Y_1|^2 \geq \varepsilon^2 \frac{1}{u_n}}) \\ &= E(Y_1^2 1_{|Y_1|^2 \geq \varepsilon^2 \frac{1}{u_n}}) \rightarrow 0. \end{aligned}$$

Lindeberg Condition and Lyapunov Condition: Example

- **Example:** Let Y_1, Y_2, \dots be i.i.d. random variables with mean 0 and variance 1, and $\{c_n\}$ be a constant sequence. If $u_n = \frac{\max_{1 \leq k \leq n} c_k^2}{c_1^2 + \dots + c_n^2} \rightarrow 0$ as $n \rightarrow \infty$, then $\mathcal{L}\left(\frac{\sum_{k=1}^n c_k Y_k}{\sqrt{c_1^2 + \dots + c_n^2}}\right)$ converges weakly to $N(0, 1)$.
- If for some $\delta > 0$, $E(Y_1^{2+\delta}) = M < \infty$, we can also verify the Lyapunov condition.

$$\begin{aligned}\sum_{k=1}^{r_n} \frac{E(Z_{nk}^{2+\delta})}{s_n^{2+\delta}} &= \sum_{k=1}^n \frac{c_k^{2+\delta} E(Y_k^{2+\delta})}{s_n^{2+\delta}} \\&= M \sum_{k=1}^n \frac{c_k^{2+\delta}}{s_n^{2+\delta}} = M \sum_{k=1}^n \frac{c_k^2}{s_n^2} \left(\frac{c_k^2}{s_n^2}\right)^{\delta/2} \\&\leq M \sum_{k=1}^n \frac{c_k^2}{s_n^2} (u_n)^{\delta/2} = M(u_n)^{\delta/2} \rightarrow 0.\end{aligned}$$