ST5215 Advanced Statistical Theory, Lecture 17

HUANG Dongming

National University of Singapore

15 Oct 2020

Overview

Last time

- Continuous mapping
- Slutsky's theorem
- δ-method
- Strong Law of Large Number (SLLN)

Today

- Weak Law of Large Number (WLLN)
- Weak Convergence of Measures
- Central Limit Theorem

Recap

- Continuous mapping: Let $g : \mathcal{R}^k \to \mathcal{R}$ be continuous. If $X_n \xrightarrow{*} X$, then $g(X_n) \xrightarrow{*} g(X)$ where $*$ could be *a.s.*, P, or D
- Slutsky's theorem: If $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{D} c$ for a constant c, then
  - $X_n + Y_n \xrightarrow{D} X + c$
  - $X_n Y_n \xrightarrow{D} cX$
  - $X_n / Y_n \xrightarrow{D} X/c$ if $c \neq 0$
- δ-method: If $a_n \to \infty$ and $a_n (X_n - c) \xrightarrow{D} Y$ where $c \in \mathcal{R}$. If g is a function from $\mathcal{R}$ to $\mathcal{R}$ and is differentiable at c, then

$$a_n [g(X_n) - g(c)] \xrightarrow{D} g'(c) Y$$

- SLLN: Let $X_1, X_2, \ldots$ be i.i.d. random variables. There exists a constant c s.t. $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{a.s.}} c \Leftrightarrow E|X_1| < \infty$ and $c = EX_1$

Lemma

$E|X| \le \sum_{n=0}^{\infty} P(|X| \ge n) \le 1 + E|X|$

Lemma (Kronecker's lemma)

Suppose $\{x_n\}$ *is a sequence of real numbers, and* $a_n \uparrow \infty$ *and are nonnegative. If* $\sum_{n=1}^{\infty} \frac{x_n}{a_n}$ *converges, then* $a_n^{-1} \sum_{i=1}^{n} x_i \to 0$

Lemma (Kolmogorov's inequality)

Suppose $Z_1, Z_2, \ldots$ *are independent and have* 0 *means and finite variances.*
Let $S_j = \sum_{i=1}^{j} Z_i$*.*
Then

$$P\left(\max_{1 \le j \le n} |S_j| > t\right) \le \frac{1}{t^2} \sum_{i=1}^{n} \mathrm{Var}(Z_i) \tag{1}$$

Proposition (Z)

If Z_n*'s are independent,* $EZ_n = 0$ *for all* n*, and* $\sum_{n=1}^{\infty} \mathrm{Var}(Z_n) < \infty$*, then* $\sum_{n=1}^{\infty} Z_n$ *converges a.s.*

Example

Let f and g be continuous functions on [0,1] satisfying $0 \leq f(x) \leq Cg(x)$ for all x, where $C > 0$ is a constant. Assume that $\int_0^1 g(x)dx \neq 0$
We now show that

$$\lim_{n\to\infty} \int_0^1 \int_0^1 \cdots \int_0^1 \frac{\sum_{i=1}^n f(x_i)}{\sum_{i=1}^n g(x_i)} dx_1 dx_2 \cdots dx_n = \frac{\int_0^1 f(x)dx}{\int_0^1 g(x)dx}$$

Proof.
Let $X_1, X_2, \ldots$ be i.i.d. $Unif[0,1]$. Then

$$E[f(X_1)] = \int_0^1 f(x)dx < \infty, \quad E[g(X_1)] = \int_0^1 g(x)dx < \infty$$

- By the SLLN

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i) \stackrel{a.s.}{\to} E[f(X_1)], \quad \frac{1}{n}\sum_{i=1}^{n} g(X_i) \stackrel{a.s.}{\to} E[g(X_1)]$$

- By the properties of continuous mapping and concatenation,

$$Y_n := \frac{\sum_{i=1}^{n} f(X_i)}{\sum_{i=1}^{n} g(X_i)} \stackrel{a.s.}{\to} \frac{E[f(X_1)]}{E[g(X_1)]}$$

- Note that $Y_n \in [0, C]$. By DCT, we have

$$EY_n \to \frac{E[f(X_1)]}{E[g(X_1)]}$$

Weak Law of Large Number (WLLN)

Theorem

Let $X_1, X_2, \ldots$ be i.i.d. random variables. A necessary and sufficient condition for the existence of a sequence of real numbers $\{a_n\}$ for which

$$\frac{1}{n}\sum_{i=1}^{n} X_i - a_n \xrightarrow{\mathcal{P}} 0$$

is that $nP\left(|X_1| > n\right) \to 0$, in which case we may take $a_n = E\left(X_1 I_{\{|X_1| \leq n\}}\right)$

- "Weak" refers to convergence in probability
- Unlike SLLN, this result does not require $E|X_1| < \infty$

Proof of Sufficiency

- Let $Y_{nj} = X_j I_{\{|X_j| \le n\}}$, $T_n = \sum_{j=1}^n X_j$ and $Z_n = \sum_{j=1}^n Y_{nj}$
- Then

$$P(T_n \neq Z_n) \le \sum_{j=1}^n P(Y_{nj} \neq X_j) = nP(|X_1| > n) \to 0 \qquad (2)$$

- By Chebyshev's inequality, for any $\epsilon > 0$,

$$P\left(\left|\frac{Z_n - EZ_n}{n}\right| > \epsilon\right) \le \frac{\operatorname{Var}(Z_n)}{\epsilon^2 n^2} \le \frac{EY_{n1}^2}{\epsilon^2 n} \qquad (3)$$

- Using integration by part

$$\begin{aligned} \frac{EY_{n1}^2}{n} &= \frac{1}{n}\int_0^n x^2 dF_{|X_1|}(x) \\ &= \frac{2}{n}\int_0^n xP(|X_1| > x)\,dx - nP(|X_1| > n) \to 0 \end{aligned}$$

- (2) and (3) together imply the result with $a_n = EZ_n/n$

Law of Large Number without Identical Distribution

Theorem

Let $X_1, X_2, \ldots$ be independent random variables with finite expectations.

(i) *(The SLLN). If there is a constant $p \in [1,2]$ such that $\sum_{i=1}^{\infty} \frac{E|X_i|^p}{i^p} < \infty$, then*

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - EX_i) \xrightarrow{a.s.} 0$$

(ii) *(The WLLN). If there is a constant $p \in [1,2]$ such that $\lim_{n\to\infty} \frac{1}{n^p}\sum_{i=1}^{n} E|X_i|^p = 0$, then*

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - EX_i) \xrightarrow{\mathcal{P}} 0$$

- The condition for SLLN implies the condition for WLLN (Kronecker's Lemma)
- If $\sup_n E|X_n|^p < \infty$ for some $p \in (1,2]$, then the condition for SLLN holds (since $\sum_n 1/n^p < \infty$)

Remarks on the proof

The proof is left for exercise. Here is a roadmap for (i)

- Consider the same truncation $Y_n = X_n I_{|X_n| \le n}$
- Use the inequality that

$$I_{|X_n| \le n} \le \frac{n^{2-p}}{|X_n|^{2-p}}$$

  and the same idea in the proof of SLLN for i.i.d. (Proposition Z) to show $n^{-1} \sum_{i=1}^{n} (Y_i - EY_i) \overset{a.s.}{\to} 0$
- Show $\frac{1}{n} \sum_{i=1}^{n} X_i - \frac{1}{n} \sum_{i=1}^{n} Y_i \overset{a.s.}{\to} 0$ using Borel-Cantelli lemma
- Show $n^{-1} \sum_{i=1}^{n} |E(X_i - Y_i)| \to 0$ by showing $\sum_{n=1}^{\infty} \frac{|E(X_n - Y_n)|}{n} < \infty$ (with Kronecker's Lemma)

Proof for (ii): use Chebyshev's inequality and

$$E \left| \sum_{i=1}^{n} X_i \right|^p \le C_p \sum_{i=1}^{n} E |X_i|^p$$

for $p \in [1, 2]$

Example

- Let $T_n = \sum_{i=1}^{n} X_i$, where X_n 's are independent random variables satisfying $P\left(X_n = \pm n^\theta\right) = 0.5$ and $\theta > 0$ is a constant.
- For $\theta < 0.5$,

$$\sum_{n=1}^{\infty} \frac{EX_n^2}{n^2} = \sum_{n=1}^{\infty} \frac{n^{2\theta}}{n^2} < \infty$$

- By SLLN, $T_n/n \stackrel{a.s.}{\to} 0$.

Weak Convergency

- Convergence in distribution is about the convergence of CDFs, not really about random variables (they are dummy variables)
- CDFs are probability measures

Definition (Convergence of probability measures)

A sequence of probability measures ν_n converges weakly to ν if $\int f \; \mathrm{d}\nu_n \to \int f \; \mathrm{d}\nu$ for every bounded and continuous real function f

Proposition

Suppose X_n's and X are random k-vectors. $X_n \xrightarrow{\mathcal{D}} X$ is equivalent to any one of the following conditions:

(a) $E\left[h\left(X_n\right)\right] \to E[h(X)]$ *for every bounded continuous function* h

(b) $\limsup_n P_{X_n}(C) \leq P_X(C)$ *for any closed set* $C \subset \mathcal{R}^k$

(c) $\liminf_n P_{X_n}(O) \geq P_X(O)$ *for any open set* $O \subset \mathcal{R}^k$.

Convergence in distribution can be characterized by characteristic functions

Theorem (Lévy continuity)

$\{X_n\}$ converges in distribution to X iff the corresponding characteristic functions $\{\phi_n\}$ converges pointwise to ϕ_X.

Example

- Let $X_1, \ldots, X_n$ be i.i.d. with Lebesgue p.d.f. $f(x) = (1 - \cos x)/\left(\pi x^2\right)$
- The ch.f. of X_1 is $\max\{1 - |t|, 0\}$ and the ch.f. of $T_n/n = (X_1 + \cdots + X_n)/n$ is

$$\left(\max\left\{1 - \frac{|t|}{n}, 0\right\}\right)^n \to e^{-|t|}, n \to \infty$$

  for $t \in \mathcal{R}$
- Since $e^{-|t|}$ is the ch.f. of the Cauchy distribution, we conclude that $T_n/n \xrightarrow{\mathcal{D}} X$, where X has the Cauchy distribution

If X has a p.d.f. f and X_n has a p.d.f. f_n, we have another way to check whether $X_n \xrightarrow{\mathcal{D}} X$

Theorem (Scheffés theorem)

Let $\{f_n\}$ be a sequence of p.d.f.'s on $\mathcal{R}^k$ w.r.t. a measure ν. Suppose that $\lim_{n\to\infty} f_n(x) = f(x)$ a.e. ν and $f(x)$ is a p.d.f. w.r.t. ν. Then $\lim_{n\to\infty} \int |f_n(x) - f(x)|\, d\nu = 0$ and $P_{f_n} \Rightarrow P_f$

- Let $g_n(x) = [f(x) - f_n(x)]\, I_{\{f \geq f_n\}}(x), n = 1, 2, \ldots$ Then

$$\int |f_n(x) - f(x)|\, d\nu = 2 \int g_n(x) d\nu$$

- Since $0 \leq g_n(x) \leq f(x)$ for all x and $g_n \to 0$ a.e. ν, the result follows from DCT.
- Let F_n and F be the c.d.f. of f_n and f. For any $x \in \mathcal{R}^k$, let $A = \{y \in \mathcal{R}^k : y_i \leq x_i, i = 1, \ldots, k\}$, then

$$\left| \int_A f_n \, \mathrm{d}\nu - \int_A f \, \mathrm{d}\nu \right| \leq \int |f_n - f| \, \mathrm{d}\nu \to 0, \qquad (4)$$

which implies $F_n(x) \to F(x)$

Remarks on Scheffés theorem

- ν is usually the Lebesgue measure or counting measure
- e.g. $X_n \sim \mathrm{Binom}(n, p_n)$ and $np_n \to \lambda$, then $X_n \xrightarrow{D} X \sim \mathrm{Poisson}(\lambda)$
- e.g. $X_n \sim t_n$ then $X_n \xrightarrow{\mathcal{D}} X \sim N(0, 1)$

Central Limit Theorem

Sometimes, we need to find the asymptotic distributions of a statistic to make inference

- e.g. asymptotic hypothesis test, confidence intervals

Theorem (Classical CLT)

Let $\{X_n\}_{n=1}^{\infty}$ *be a sequence of i.i.d. random k-vectors. Suppose* $\Sigma = \mathrm{Var} X_1$ *is finite, then*

$$\frac{\sum_{i=1}^{n}(X_i - EX_i)}{\sqrt{n}} \xrightarrow{\mathcal{D}} N(0, \mathrm{Var}(X_1)) \tag{5}$$

- If $k = 1$, proved in Example 1.28 in the textbook using Lévy continuity by computing the limit of ch.f.'s
- For general k, use Cramér-Wold device

CLT for Triangular Arrays

Theorem (Lindeberg's CLT)

For each n, let $\{X_{nj}, j = 1, \ldots, k_n\}$ *be a set of independent random variables. Suppose* $k_n \to \infty$ *as* $n \to \infty$ *and*

$$0 < \sigma_n^2 = \mathrm{Var}\left(\sum_{j=1}^{k_n} X_{nj}\right) < \infty, \qquad n = 1, 2, \ldots.$$

If

$$\frac{1}{\sigma_n^2} \sum_{j=1}^{k_n} E\left\{(X_{nj} - EX_{nj})^2 I_{\{|X_{nj} - EX_{nj}| > \epsilon\sigma_n\}}\right\} \to 0 \tag{6}$$

for any $\epsilon > 0$*, then*

$$\frac{1}{\sigma_n} \sum_{j=1}^{k_n} (X_{nj} - EX_{nj}) \xrightarrow{D} N(0, 1).$$

Remarks

- Condition (6) controls the tails of X_{nj}, and is called *Lindeberg's condition*.
- Condition (6) is implied by either of the following
  - Lyapunov condition:

$$\frac{1}{\sigma_n^{2+\delta}} \sum_{j=1}^{k_n} E|X_{nj} - EX_{nj}|^{2+\delta} \to 0 \text{ for some } \delta > 0. \tag{7}$$

  - Uniform boundedness: if $|X_{nj}| \leq M$ for all n and j and $\sigma_n^2 = \sum_{j=1}^{k_n} \mathrm{Var}(X_{nj}) \to \infty$.
- In general, Condition (6) is NOT necessary for the convergence result.
- But if we assume the *Feller's condition*:

$$\lim_{n\to\infty} \max_{j \leq k_n} \frac{\sigma_{nj}^2}{\sigma_n^2} = 0,$$

then Condition (6) is not only sufficient but also necessary

Example: Asymptotic Distribution of Empirical Variance

- Let $X_1, \ldots, X_n$ be i.i.d. such that $EX_1^4 < \infty$.
- Denote $\sigma^2 = \mathrm{Var}(X_1)$, $\mu = EX_1$, and $m_2 = EX_1^2$.
- Let $\hat{\mu} = \overline{X} = n^{-1} \sum_{i=1}^n X_i$ and $\hat{\sigma}^2 = n^{-1} \sum_{i=1}^n (X_i - \overline{X})^2$.

Now we derive the asymptotic distribution of $\sqrt{n}(\hat{\sigma}^2 - \sigma^2)$.

- Note that $\hat{\sigma}^2 = \hat{m}_2 - \hat{\mu}^2$, where $\hat{m}_2 = n^{-1} \sum_{i=1}^n X_i^2$.
- This motivates us to define $g(y_1, y_2) = y_2 - y_1^2$.
- By multivariate CLT, for $Y_n = (\hat{\mu}, \hat{m}_2)^\top$, we have $\sqrt{n}(Y_n - c) \xrightarrow{D} N(0, \Sigma)$, where $c = (\mu, m_2)$ and $\Sigma = \mathrm{Cov}([X_1, X_1^2]^\top)$.
- Observe that $\nabla g(y_1, y_2) = (-2y_1, 1)^\top \neq 0$.
- By δ-method,

$$\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{D} N\left(0, (-2\mu, 1)\Sigma(-2\mu, 1)^\top\right). \tag{8}$$

Tutorial

1. Suppose that $X_n \xrightarrow{\text{D}} X$. Then, for any $r > 0$
$$\lim_{n\to\infty} E|X_n|^r = E|X|^r < \infty$$
if and only if $\{|X_n|^r\}$ is uniformly integrable in the sense that
$$\lim_{t\to\infty} \sup_n E\left(|X_n|^r I_{\{|X_n|>t\}}\right) = 0$$

2. Let $X, X_1, X_2, \ldots$ be random variables. Show that if $\lim_n X_n = X$ a.s., then $Y_n := \sup_{m>n} |X_m|$ is bounded in probability.

Exercise 1

Suppose that $X_n \stackrel{D}{\to} X$. Then, for any $r > 0$

$$\lim_{n\to\infty} E|X_n|^r = E|X|^r < \infty$$

if and only if $\{|X_n|^r\}$ is uniformly integrable in the sense that

$$\lim_{t\to\infty} \sup_n E\left(|X_n|^r I_{\{|X_n|>t\}}\right) = 0$$

Proof: By Skorohod's theorem, we can assume that $X_n \stackrel{a.s.}{\to} X$
Sufficiency: Assume that $\{|X_n|^r\}$ is uniformly integrable

- Choose t large but fixed s.t. $\sup_n E\,|X_n|^r \le t^r + \sup_n E\left(|X_n|^r I_{\{|X_n|>t\}}\right) < t^r + 1 < \infty$
- By Fatou's lemma, $E|X|^r \le \liminf_n E\,|X_n|^r < \infty$
- We only need to show

$$\limsup_n E\,|X_n|^r \le E|X|^r$$

Notation: For any $\epsilon > 0$ and $t > 0$, let $A_n = \{|X_n - X| \leq \epsilon\}$ and $B_n = \{|X_n| > t\}$.

- Then

$$\begin{aligned} E\left|X_n\right|^r &= E\left(\left|X_n\right|^r I_{A_n^c \cap B_n}\right) + E\left(\left|X_n\right|^r I_{A_n^c \cap B_n^c}\right) + E\left(\left|X_n\right|^r I_{A_n}\right) \\ &\leq E\left(\left|X_n\right|^r I_{B_n}\right) + t^r P\left(A_n^c\right) + E\left|X_n I_{A_n}\right|^r \end{aligned}$$

- For $r \leq 1, \left|X_n I_{A_n}\right|^r \leq \left(\left|X_n - X\right|^r + \left|X\right|^r\right) I_{A_n}$ and

$$E\left|X_n I_{A_n}\right|^r \leq E\left[\left(\left|X_n - X\right|^r + \left|X\right|^r\right) I_{A_n}\right] \leq \epsilon^r + E|X|^r$$

- For $r > 1$, an application of Minkowski's inequality leads to

$$\begin{aligned} E\left|X_n I_{A_n}\right|^r &= E\left|\left(X_n - X\right) I_{A_n} + X I_{A_n}\right|^r \\ &\leq \left\{\left[E\left|\left(X_n - X\right) I_{A_n}\right|^r\right]^{1/r} + \left[E\left|X I_{A_n}\right|^r\right]^{1/r}\right\} \\ &\leq \left\{\epsilon + \left[E|X|^r\right]^{1/r}\right\}^r \end{aligned}$$

- In the following, we assume $r \leq 1$. The case for $r > 1$ is essentially the same

- Since $P\left(A_n^c\right) \to 0$, we have

$$\begin{aligned}
\limsup_n E\left|X_n\right|^r &\leq \limsup_n E\left(\left|X_n\right|^r I_{B_n}\right) + t^r \lim_{n\to\infty} P\left(A_n^c\right) \\
&\quad + \limsup_n E\left|X_n I_{A_n}\right|^r \\
&\leq \sup_n E\left(\left|X_n\right|^r I_{\{|X_n|>t\}}\right) + E|X|^r + \epsilon^r
\end{aligned}$$

- Since $\{|X_n|^r\}$ is uniformly integrable, by letting $t \to \infty$, we have

$$\limsup_n E\left|X_n\right|^r \leq E|X|^r + \epsilon^r$$

- Since ϵ is arbitrary, we conclude that $\limsup_n E\left|X_n\right|^r \leq E|X|^r$

Exercise 1, Necessity

Suppose $\lim_{n\to\infty} E|X_n|^r = E|X|^r < \infty$

- First fixed t. Let $B_{n,t} = \{|X_n| > t\}$ and let $\xi_n = |X_n|^r I_{B^c_{n,t}} - |X|^r I_{B^c_{n,t}}$.
- Then $\xi_n \overset{a.s.}{\to} 0$ and $|\xi_n| \le t^r + |X|^r$, which is integrable.
  By the dominated convergence theorem, $E\xi_n \to 0$.
- Since $E|X_n|^r - E|X|^r \to 0$,

$$E\left(|X_n|^r I_{B_{n,t}}\right) - E\left(|X|^r I_{B_{n,t}}\right) \to 0$$

- Note that $B_{n,t} \subset \{|X_n - X| > t/2\} \cup \{|X| > t/2\}$,

$$\begin{aligned}\limsup_n E\left(|X_n|^r I_{B_{n,t}}\right) &\le \limsup_n E\left(|X|^r I_{B_{n,t}}\right)\\ &\le E\left(|X|^r I_{\{|X|>t/2\}}\right) + \\ &\qquad \limsup_n E\left(|X|^r I_{\{|X_n - X|>t/2\}}\right)\end{aligned}$$

Exercise 1, Necessity (Cont.)

$$\limsup_n E\left(|X_n|^r I_{B_{n,t}}\right) \leq E\left(|X|^r I_{\{|X|>t/2\}}\right) + \limsup_n E\left(|X|^r I_{\{|X_n - X|>t/2\}}\right)$$

- By DCT, $\lim_n E\left(|X|^r I_{\{|X_n-X|>t/2\}}\right) = 0$
- By DCT, $E\left(|X|^r I_{\{|X|>t/2\}}\right) \to 0$ as $t \to \infty$
- So

$$\lim_{t\to\infty} \limsup_n E\left(|X_n|^r I_{B_{n,t}}\right) \leq 0.$$

- Note that for any n, $I_{B_{n,t}}$ is decreasing in t.
- For any ϵ, we can pick t_1 and N_1 large enough s.t.

$$\sup_{t\geq t_1} \sup_{n\geq N_1} E\left(|X_n|^r I_{B_{n,t_1}}\right) < \epsilon$$

  and then take t_2 large enough s.t.

$$\sup_{n\leq N_1} E\left(|X_n|^r I_{B_{n,t_2}}\right) < \epsilon$$

- So if $t > \max(t_1, t_2)$, we have $\sup_n E\left(|X_n|^r I_{B_{n,t}}\right) < \epsilon$

Exercise 2

Let $X, X_1, X_2, \ldots$ be random variables. Show that if $\lim_n X_n = X$ a.s., then $Y_n := \sup_{m>n} |X_m|$ is bounded in probability.

Proof:

- Since $\sup_{m \geq n} |X_m| \leq \sup_{m \geq 1} |X_m|$ for any n, it suffices to show that for any $\epsilon > 0$, there is a $C > 0$ such that
$$P\left(\sup_{n \geq 1} |X_n| > C\right) \leq \epsilon.$$
- Note that $X_n \stackrel{a.s.}{\to} X \Leftrightarrow \lim_{n\to\infty} P\left(\bigcup_{m=n}^{\infty} \{|X_m - X| > \epsilon\}\right) = 0$. For any $\epsilon > 0$ and any fixed $c_1 > 0$, there exists a sufficiently large N such that $P\left(\cup_{n=N}^{\infty} \{|X_n - X| > c_1\}\right) < \epsilon/3$
- For this fixed N, there exist constants $c_2 > 0$ and $c_3 > 0$ such that
$$\sum_{n=1}^{N} P\left(|X_n| > c_2\right) < \frac{\epsilon}{3} \quad \text{and} \quad P\left(|X| > c_3\right) < \frac{\epsilon}{3}$$

Exercise 2 (Cont.)

$P\left(\cup_{n=N}^{\infty}\{|X_n - X| > c_1\}\right) < \epsilon/3$, $\sum_{n=1}^{N} P\left(|X_n| > c_2\right) < \frac{\epsilon}{3}$, and $P\left(|X| > c_3\right) < \frac{\epsilon}{3}$

Let $C = \max\{c_1 + c_3, c_2\}$. Then the result follows from

$$
\begin{aligned}
P\left(\sup_{n\geq 1}|X_n| > C\right) &= P\left(\bigcup_{n=1}^{\infty}\{|X_n| > C\}\right) \\
&\leq \sum_{n=1}^{N} P\left(|X_n| > C\right) + P\left(\bigcup_{n=N}^{\infty}\{|X_n| > C\}\right) \\
&\leq \frac{\epsilon}{3} + P\left(|X| > c_3\right) + P\left(\bigcup_{n=N}^{\infty}\{|X_n| > C, |X| \leq c_3\}\right) \\
&\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + P\left(\bigcup_{n=N}^{\infty}\{|X_n - X| > c_1\}\right) \\
&\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
\end{aligned}
$$