Convergence in Probability

Definition: $X_n \to_p X$ iff $\forall \varepsilon > 0$,
$$\mathbb{P}(|X_n - X| > \varepsilon) \to 0.$$

Remark:

① Indeed, convergence in probability could be metrization, define:
$$d(X, Y) := \mathbb{E}\,\frac{|X-Y|}{|X-Y|+1},$$
then $X_n \to_p X \iff d(X_n, X) \to 0$, which implies that the convergence is similar to other convergence in metric space.

We mark that $X_n \to_p 0$:
$$X_n = o_p(1),$$
and $X_n = o_p(Y_n)$ iff
$$\frac{X_n}{Y_n} = o_p(1).$$

Definition: $X_n = O_p(1)$ iff $\forall \delta > 0$, $\exists$ $M > 0$ s.t.

$$\sup_n \mathbb{P}(|X_n| > M) \leq \delta,$$

which is said to be bounded in probability.

Remark:

① We assume that $X = O_p(1)$.

Pro

If $X_n \to_p X$, then $X_n = O_p(1)$

Pf: $X_n = X_n - X + X$, we just need to prove

$$X - X_n = O_p(1)$$

Since $X - X_n = o_p(1)$, then $\exists k$, for M_0 $\forall \varepsilon$, $\mathbb{P}(|X - X_n| > M_0) < \varepsilon$, $n \geq k$

For $i < k$, $\exists M_i$ s.t.

$$\mathbb{P}(|X - X_i| > M_i) < \varepsilon$$

Let $M = \max\{M_0, \dots, M_{k-1}\}$, then

$$\sup_n \mathbb{P}(|X - X_n| > M) < \varepsilon$$

$\square$

Continuous Mapping Theorem: If g is a continuous function, and $X_n \to_p X$, then

$$g(X_n) \to_p g(X)$$

Pf: $\forall \varepsilon > 0$,

$$\mathbb{P}(|g(X_n) - g(X)| > \varepsilon)$$

$$\leq \mathbb{P}(\{|g(X_n) - g(X)| > \varepsilon\} \text{ and } \{|X_n|, |X| \leq K)$$

$$+ \mathbb{P}(|X| > K) + \mathbb{P}(|X_n| > K).$$

Since X, $X_n = O_p(1)$, we can choose a suitable K s.t. $\forall \delta > 0$,

$$\mathbb{P}(|X| > K) < \delta/3,$$

$$\sup_n \mathbb{P}(|X_n| > K) < \delta/3.$$

Note that g is uniformly continuous on $\bar{B}(0;K)$, which means that $\exists M_\varepsilon$ s.t.

$$\mathbb{P}(|g(X) - g(X_n)| > \varepsilon \mid |X_n|, |X| \leq K)$$

$$\leq \mathbb{P}(|X - X_n| > M_\varepsilon \mid |X_n|, |X| \leq K)$$

$\Rightarrow$

$$\mathbb{P}(|g(X_n) - g(X)| > \varepsilon)$$

$$\leq \mathbb{P}(|X - X_n| > M_\varepsilon) + \frac{2\delta}{3},$$ and $\exists$ k s.t. $\forall n \geq k$,

$$\mathbb{P}(|X - X_n| > M_\varepsilon) < \delta/3.$$

$\square$

Pro

$o_p(1) + o_p(1) = o_p(1)$

$o_p(1) + O_p(1) = O_p(1)$

Lemma

① $X_n = o_p(1)$, $R(h) = o(|h|^p)$, as $h \to 0$, then

$$R(X_n) = o_p(|X_n|^p)$$

② $X_n = O_p(1)$, $R(h) = O(|h|^p)$, then

$$R(X_n) = O_p(|X_n|^p)$$

Pf: ① Let $g(h) := \begin{cases} R(h)/|h|^p, & h \neq 0, \\ 0, & h = 0 \end{cases}$

$g(h) = o(1)$, $h \to 0^+$.

② $g(h) = O(1)$, $h \to 0^+$, then $\forall \delta > 0$, $\exists M$ s.t. $|g(h)| \le M$, if $|h| < \delta$, $h \neq 0$.

Then for $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t.

$$1 - \varepsilon \le \mathbb{P}(|X_n| < \delta) \le \mathbb{P}(|g(X_n)| \le M).$$

$\square$

Weak convergence

Let $F_X(x) = \mathbb{P}(X \le x)$.

Definition: $X_n \to_w X$ iff

$F_{X_n}(x) \to F_X(x)$, where x is a continuous points of F_X.

Remark: X_n, X don't need to be defined in a same probability space, while they should value the same space. (Define on law!)

Pro ① $X_n \to_w X$ if $X_n \to_p X$;

② $X_n \to_w C$ iff $X_n \to_p C$.

Pf: ① $\forall$ continuous point of F_X,

$$\begin{aligned}
&\mathbb{P}(X_n \le x) \\
&= \mathbb{P}(X_n \le x, X > x) + \mathbb{P}(X_n \le x, X \le x) \\
&= \mathbb{P}(X_n \le x, X > x) - \mathbb{P}(X_n > x, X \le x) \\
&\qquad + \mathbb{P}(X \le x)
\end{aligned}$$

Note $\mathbb{P}(x < X \le x+\delta) = o(1)$, as $\delta \to 0^+$, then

$$\begin{aligned}
\mathbb{P}(X_n \le x, X > x) &= \mathbb{P}(X_n \le x, X > x+\delta) \\
&\quad + o(1) \\
&\le \mathbb{P}(|X_n - X| > \delta) + o(1) \\
&= o(1).
\end{aligned}$$

Do the similar step for
$$\mathbb{P}(X_n > x, X \le x),$$
then $\mathbb{P}(X_n \le x) \to \mathbb{P}(X \le x)$.

② $\forall \varepsilon > 0$,
$$\begin{aligned} &\mathbb{P}(|X_n - C| > \varepsilon) \\ = &F_{X_n}(C-\varepsilon) + 1 - F_{X_n}(C+\varepsilon) \\ \to &F_X(C-\varepsilon) + 1 - F_X(C+\varepsilon) \\ = &\mathbb{P}(|X - c| > \varepsilon) = 0. \end{aligned}$$
$\square$

Skorokhod's Representation Theorem

If $X_n \to_w X$, then $\exists \hat{X}_n, \hat{X}$, which is defined in the same probability space s.t.
$$F_{\hat{X}_n} = F_{X_n},$$
$$F_{\hat{X}} = F_X,$$
and $\hat{X}_n \to \hat{X}$ a.s. $(X_n, X \in \mathbb{R}^p)$

Pf: Let $\hat{X}_n = F_n^{-1}(U)$, $\hat{X} = F^{-1}(U)$, $U \sim \text{Unif}[0,1]$. Claim that:
$$\hat{X}_n \to \hat{X} \text{ a.s.}$$
Let $p=1$, note that $\hat{X}_n, \hat{X}$ is defined in $([0,1], \mathcal{B}([0,1]))$.

Assume that $C_x = \{ y \in \mathbb{R};$ y is continuous point of $F \}$.

$\forall x \in (0,1)$, $\varepsilon > 0$, $\exists y \in C_x$ s.t.

$$F^{-1}(x) - \varepsilon < y < F^{-1}(x)$$

$\Rightarrow$ $F(y) < x$, then $\exists k$ s.t.

$F_n(y) < x$, $n \geq k$

$\Rightarrow$ $y < F_n^{-1}(x)$

$\Rightarrow$ $F^{-1}(x) - \varepsilon < F_n^{-1}(x)$ ①

Similarly, pick $x' \in (0,1)$, $y \in C_x$, s.t. $x < x'$,

$$F^{-1}(x') < y < F^{-1}(x') + \varepsilon$$

$\Rightarrow$ $x' < F_n(y)$, if n is large enough

$\Rightarrow$ $F_n^{-1}(x') < y$

$\Rightarrow$ $F_n^{-1}(x) < F_n^{-1}(x') < F^{-1}(x') + \varepsilon$ ②

① + ②:

$$F^{-1}(x) \leq \underline{\lim}_n F_n^{-1}(x) \leq \overline{\lim}_n F_n^{-1}(x)$$

(continuous) $\leq F^{-1}(x')$,

Let $x' \downarrow x$, claim holds. $\square$

Remark: If $X_n \to_w X$, g is continuous,

$\exists$ $\hat{X}_n \to \hat{X}$ a.s. $\Rightarrow g(\hat{X}_n) \to_w g(\hat{X})$

$\Rightarrow g(X_n) \to_w g(X)$.

Portmanteau Lemma

The following statements are equivalent.

(1) $\mathbb{P}(X_n \le x) \to \mathbb{P}(X \le x)$ for all continuous points $x \mapsto \mathbb{P}(X \le x)$.

(2) $\mathbb{E} f(X_n) \to \mathbb{E} f(X)$ for all bound, Lipschitz function f.

(3) $\underline{\lim}_n \mathbb{P}(X_n \in G) \ge \mathbb{P}(X \in G)$ for $\forall$ open G.

(4) $\overline{\lim}_n \mathbb{P}(X_n \in F) \le \mathbb{P}(X \in F)$ for $\forall$ closed F.

(5) $\mathbb{P}(X_n \in B) \to \mathbb{P}(X \in B)$ for $\forall$ B s.t. $\mathbb{P}(X \in \partial B) = 0$.

Pf: (1) $\Rightarrow$ (2):

$$X_n \to_w X \Rightarrow g(X_n) \to_w g(X)$$
$$\Rightarrow \mathbb{E} g(X_n) \to \mathbb{E} g(X)$$

(2) $\Rightarrow$ (3)

Let $f_m(x) = (m\, d(x, G^c)) \wedge 1$

$\Rightarrow \quad f_m(x) \uparrow \mathbb{1}_G(x), \ \forall \varepsilon > 0$

$$\underline{\lim}_n \mathbb{P}(X_n \in G)$$
$$= \underline{\lim}_n \lim_m \mathbb{E} f_m(X_n)$$
$$\geq \underline{\lim}_n \mathbb{E} f_m(X_n) = \mathbb{E} f_m(X)$$
$$\Rightarrow \underline{\lim}_n \mathbb{P}(X_n \in G) \geq \mathbb{P}(X \in G)$$

(3) $\Leftrightarrow$ (4) : Easy

(3) + (4) $\Rightarrow$ (5):

$$\underline{\lim}_n \mathbb{P}(X_n \in \mathring{B}) \geq \mathbb{P}(X \in \mathring{B})$$
$$\overline{\lim}_n \mathbb{P}(X_n \in \overline{B}) \leq \mathbb{P}(X \in \overline{B})$$

Since $\mathbb{P}(X \in \mathring{B}) = \mathbb{P}(X \in \overline{B})$

and

$$\underline{\lim}_n \mathbb{P}(X_n \in \mathring{B}) \leq \underline{\lim}_n \mathbb{P}(X_n \in B)$$
$$\leq \overline{\lim}_n \mathbb{P}(X_n \in B) \leq \overline{\lim}_n \mathbb{P}(X_n \in \overline{B})$$

(5) $\Rightarrow$ (1) : Easy. $\square$

Prohorov's Theorem

(1) $X_n \to_w X$, then $X_n = O_p(1)$.

(2) If $X_n = O_p(1)$, then $\exists X_{n_k}, X$ s.t. $X_{n_k} \to_w X$, $X = O_p(1)$.

Pf: (1) Given $M > 0$,

$$\overline{\lim_n} \mathbb{P}(|X_n| \geq M) \leq \mathbb{P}(|X| \geq M),$$

let M be large enough.

(2) (Helly Selection) Let $Q = \{q_i\}$ be the set of rational. Note that

$\{F_{X_n}(q_1)\}$ is bounded

$\Rightarrow \exists \{n_{k_1}\}$ s.t. $F_{X_{n_{k_1}}}(q_1) \to \tilde{F}(q_1)$

$\Rightarrow \{F_{X_n}(q_2)\}$ is bounded

$\Rightarrow \exists \{n_{k_2}\} \subset \{n_{k_1}\}$ s.t.

$$F_{X_{n_{k_2}}} \to \check{F}(q_2)$$

$\Rightarrow \cdots \Rightarrow \exists \{n_{k_j}\} \subset \{n_{k_{j-1}}\}$ s.t.

$$F_{X_{n_j}}(q_j) \to \check{F}(q_j)$$

$\Rightarrow \exists \{n_k\}$ s.t.

$$F_{X_{n_k}}(q) \to \check{F}(q), \quad \forall q \in Q$$

and $\check{F} \uparrow$. Let

$$F(x) = \inf_{q > x} \check{F}(q).$$

choose $q, q' \in Q$ for $\forall$ continuous point of

F s.t. $q < x < q'$

$\Rightarrow \quad F(q) \leq F(x) \leq F(q')$

$\Rightarrow$

$$F_{n_k}(q) \leq F_{n_k}(x) \leq F_{n_k}(q')$$

$\Rightarrow \quad F(q) \leq \underline{\lim}_n F_{n_k}(x) \leq \overline{\lim}_n F_{n_k}(x) \leq F(q')$

Let $q, q' \to x$:

$\Rightarrow \quad F_{n_k}(x) \to F(x)$

Since $X_{n_k} = O_p(1)$, $\forall \varepsilon > 0$, $\exists M$

$\mathbb{P}(|X_{n_k}| > M) < \varepsilon$, then $\exists M$

s.t.

$$\mathbb{P}(|X| > M) < \varepsilon$$

$\Rightarrow \quad X = O_p(1)$ $\square$

Remark:

① If $X_n = O_p(1)$, if we want to show $X_n \xrightarrow{w} X$, just need to show $\forall X_{n_k}$, $\exists X_{n_{k_j}} \to Y =_d X$.

②
$$O_p(1)\, o_p(1) = o_p(1)$$
$$(1 + o_p(1))^{-1} = O_p(1)$$
$$o_p(O_p(1)) = o_p(1)$$

Pro

① If $X_n \to_w X$, $X_n - Y_n = o_p(1)$,
then $Y_n \to_w X$.

② If $X_n \to_w X$, $Y_n \to_p c$, then
$(X_n, Y_n) \to_w (X, c)$

Pf: ① $\forall$ bound Lipschitz function
f, $|f| < M$, $|f(x) - f(y)| \leq L|x-y|$

$$\begin{aligned}
&\mathbb{E} f(Y_n) \\
&= \mathbb{E} f(X_n + Y_n - X_n) \\
&= \mathbb{P}(|Y_n - X_n| > \varepsilon)) M + \\
&\quad \mathbb{E} \mathbb{1}(|Y_n - X_n| \leq \varepsilon) f(X_n) + L\varepsilon \\
&= \mathbb{E} f(X_n) + 2\mathbb{P}(|Y_n - X_n| > \varepsilon) M \\
&\qquad + L\varepsilon
\end{aligned}$$

$\Rightarrow \mathbb{E} f(Y_n) \to \mathbb{E} f(X_n)$.

② $(X_n, Y_n) - (X_n, c) = o_p(1)$
and $\forall$ bound Lipschitz function
f, $\mathbb{E} f(X_n, c) \to \mathbb{E} f(X, c)$
$\Rightarrow (X_n, c) \to_w (X, c)$ $\square$

Remark: (Slutsky)

① $X_n + Y_n \to_w X + c$

② $X_n Y_n \to_w cX$

③ $X_n / Y_n \to_w X / c$, $c \neq 0$

if $X_n \to_w X$, $Y_n \to_w C$.

Polya's Theorem

If $X_n \to_w X$, and F_X is continuous,
then $\|F_{X_n} - F_X\|_\infty = \sup_x |F_{X_n}(x) - F_X(x)| \to 0$.

Pf: Let $F^{-1}(\frac{i}{k}) = x_i$, then $\exists\, i$ s.t.

$$F_X(x_{i-1}) \le F_X(x) \le F_X(x_i)$$

then

$$F_X(x) - F_{X_n}(x) \le F_X(x_i) - F_{X_n}(x_{i-1})$$
$$= F_X(x_{i-1}) - F_{X_n}(x_{i-1}) + \frac{1}{k}$$

$$F_X(x) - F_{X_n}(x) \ge F_X(x_{i-1}) - F_{X_n}(x_i)$$
$$= F_X(x_i) - F_{X_n}(x_i) - \frac{1}{k}$$

$\Rightarrow$ $\sup_x |F_X(x) - F_{X_n}(x)|$
$$\le \sup_i |F_X(x_i) - F_{X_n}(x_i)| + \frac{1}{k}$$

$\square$

Characteristic Functions

Definition

$$\psi_X(t) = \mathbb{E}\, e^{itX}$$
$$= \mathbb{E}\cos(tX) + i\,\mathbb{E}\sin(tX)$$

Remark: $\psi_X \Leftrightarrow F_X$

Pro

① $\psi(0) = 1$ and $|\psi(t)| \leq 1$

② ψ is continuous at $t=0$.

Pf: $|\psi(t) - 1| \leq \mathbb{E}\,|e^{itX} - 1| \to 0$ $\square$

③ $\psi(t)$ is uniformly continuous.

Pf: $|\psi(t) - \psi(g)|$

$$= |\mathbb{E}\, e^{itX}(e^{i(g-t)X} - 1)|$$

$$\leq \mathbb{E}\,|e^{i|g-t|X} - 1| \to 0 \quad \square$$

④ $X \perp Y \Leftrightarrow \psi_{X,Y}(t,g) = \psi_X(t)\psi_Y(g)$

Pf: "$\Rightarrow$" $\psi_{X,Y}(t,g)$

$$= \mathbb{E}\, e^{i(t^TX + g^TY)}$$

$$= \mathbb{E}\, e^{it^TX} \cdot \mathbb{E}\, e^{ig^TY}$$

"$\Leftarrow$": Define $P = F_X \cdot F_Y$, then

$$\psi_P(t,g) = \psi_{X,Y}(t,g), \quad \forall t, g$$

$$\Rightarrow \quad F_{X,Y} = F_X \cdot F_Y. \qquad \square$$

Theorem

If $\mathbb{E}|X|^k < \infty$, then $\varphi_X^{(k)}(t)$ exists and is a uniformly continuous function s.t.

$$\varphi_X^{(k)}(0) = i^k \, \mathbb{E} X^k$$

Pf: $\forall t, k \in \mathbb{R}$

$$\varphi_X(t) - \varphi(g) \,/\, (t-g)$$

$$= \mathbb{E} \frac{(e^{i(t-g)X} - 1) \cdot e^{igX}}{t-g}$$

$$= \mathbb{E}(iX e^{igX} + o_p(1)), \quad g \to t$$

$$\to i \, \mathbb{E} X e^{itX}$$

$$\Rightarrow \varphi_X'(t) = i \, \mathbb{E} X e^{itX}$$

$$\Rightarrow \varphi_X^{(k)}(t) = i^k \, \mathbb{E} X^k e^{itX}$$

And

$$|\varphi_X^{(k)}(g) - \varphi_X^{(k)}(t)|$$

$$= |\mathbb{E} X^k e^{i(t-g)X} - 1| \to 0, \quad t \to k \qquad \square$$

Lemma

$$\mathbb{P}(|X| > \frac{2}{a}) \leq \frac{1}{a}\int_{-a}^{a}(1-\varphi_X(t))dt$$

Pf:

$$\frac{1}{a}\int_{-a}^{a}(1-\varphi_X(t))dt$$

$$= \frac{1}{a}\int_{-a}^{a}(1-\mathbb{E}e^{itX})dt$$

$$= \frac{1}{a}\mathbb{E}\int_{-a}^{a}(1-e^{itX})dt$$

$$= \mathbb{E}\,2 - \frac{1}{a}\int_{-a}^{a}\cos(tX)dt$$

$$= \mathbb{E}\left(2 - \frac{2\sin(aX)}{aX}\right)$$

$$\geq \mathbb{E}\left(2 - \frac{2\sin(aX)}{aX}\right)\mathbb{1}(|aX| \geq 2)$$

$$\geq \mathbb{P}(|aX| \geq 2)$$

$\square$

Levy continuous theorem

① $X_n \to_w X \Rightarrow \varphi_{X_n}(t) \to \varphi_X(t), \forall t$

② If $\varphi_{X_n}(t) \to \varphi(t)$, φ is continuous at 0, then $\exists\, X$ s.t.

$$X_n \to_w X, \quad \varphi_X = \varphi$$

Pf: ① $\varphi_{X_n}(t) = \mathbb{E}\, e^{itX_n} \to \mathbb{E} e^{itX}$.

② Claim that $X_n = O_p(1)$,

since

$$\mathbb{P}(|X_n| > M) \leq 2M \int_{-\frac{2}{M}}^{\frac{2}{M}} (1 - \varphi_{X_n}(t))$$

$$= 2M \int_{-\frac{2}{M}}^{\frac{2}{M}} (1 - \varphi(t)) + o(1)$$

,

then $X_n = O_p(1)$

$\Rightarrow \forall \{X_{n_k}\}, \exists \{X_{n_{k_j}}\}, X$ s.t. $X_{n_{k_j}} \to_w X$

$\Rightarrow \varphi_{X_{n_{k_j}}}(t) \to \varphi_X(t) = \varphi(t)$

$\Rightarrow X_n \to_w X$ $\square$

Remark: (Cramer-Wold device)

$$X_n \to_w X \text{ iff } c^T X_n \to_w c^T X, \forall c$$

Weak Law of Large Number

If $X_1, \cdots, X_n$ is iid, and $\mu = \mathbb{E}X_1$ exists, then

$$\bar{X}_n := \frac{X_1 + \cdots + X_n}{n} = \mu + o_p(1).$$

Pf: $\varphi_{\bar{X}_n}(t) = \mathbb{E} e^{it \frac{X_1 + \cdots + X_n}{n}}$

$$= \left(\mathbb{E} e^{i \frac{t}{n} X_1}\right)^n$$

$$= \left(1 + \frac{i\mu t}{n} + o\left(\frac{t}{n}\right)\right)^n$$

$$= e^{n(i\mu t/n + o(\frac{1}{n}))}$$

$$\to e^{i\mu t}$$

$$\Rightarrow \quad \bar{X}_n \to_w \mu$$

$\square$

Remark: $X_1 + \cdots + X_n = O_p(n)$.

Central Limit Theorem

If $X_1, \cdots, X_n$ iid and $\mathbb{E} X_1^2 < \infty$, then

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \to_w N(0,1),$$

where $\mu = \mathbb{E} X$, $\sigma = \sqrt{\operatorname{Var} X}$.

Pf: let $\mu = 0$, $\sigma = 1$,

$$\varphi_{\sqrt{n}\bar{X}_n}(t) = \left(\mathbb{E} e^{i \frac{t}{\sqrt{n}} X_1}\right)^n$$

$$= \left(1 - \frac{1}{2}\frac{t^2}{n} + o\left(\frac{1}{n}\right)\right)^n$$

$$\to e^{-\frac{1}{2}t^2}$$

$\Rightarrow \sqrt{n}\bar{X}_n \to_w N(0,1)$ $\square$

Remark: ① $X_1 + \cdots + X_n = O_p(\sqrt{n})$.

② $X \sim N(0,1)$,

$$\sup_t \left| \mathbb{P}\left(\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \geq t\right) - \mathbb{P}(X \geq t) \right| \to 0.$$

Lindeberg - Feller CLT

If $X_{n,1}, \cdots, X_{n,k_n}$ are independent, and
① $\mathbb{E} X_{n,i} = 0$,
② $\sum_{i=1}^{k_n} \mathbb{E} X_{n,i}^2 = 1$, $k_n \to \infty$.

and $\sum_{i=1}^{k_n} \mathbb{E} X_{n,i}^2 \mathbb{1}(|X_{n,i}| > \varepsilon) \to 0$, $\forall \varepsilon > 0$

then $\sum_{i=1}^{k_n} X_{n,i} \to_w N(0, 1)$

Pf: $\left|\sum_{i=1}^{k_n} \log \varphi_{X_{n,i}}(t) - (\varphi_{X_{n,i}}(t) - 1)\right|$

$\leq \sum_{i=1}^{k_n} |\phi_i(t)|^2$, $\phi_i(t) := \varphi_{X_{n,i}}(t) - 1$

($\ln(1+x) - x \leq x^2$, if x is small enough)

$\leq \max_i |\phi_{i,n}(t)| \sum_{i=1}^{k_n} |\phi_{i,n}(t)|$.

Note

$$|\phi_{i,n}(t)| = \left|\mathbb{E} e^{i t \sqrt{\mathbb{E} X_{n,i}^2} \frac{X_{n,i}}{\sqrt{\mathbb{E} X_{n,i}^2}}} - 1\right|$$

$$= \left|1 - \frac{t^2}{2} \mathbb{E} X_{n,i}^2 + o(t^2 \mathbb{E} X_{n,i}^2)\right|$$

$\le \frac{t^2}{2}\mathbb{E}X_{n,i}^2$, if $\mathbb{E}X_{n,i}^2$ is small enough.

$\Rightarrow |\phi_{i,n}(t)| = o(1)$

$$\sum_{i=1}^{k_n} |\phi_{i,n}(t)| \le \frac{t^2}{2}$$

$$\Rightarrow \sum_{i=1}^{k_n} \log \varphi_{X_{n,i}}(t) = \sum_{i=1}^{k_n} (\varphi_{X_{n,i}}(t) - 1) + o(1)$$

Note that
$\varphi_{X_{n,i}}(t) - 1 = -\varphi^{(2)}_{X_{n,i}}(C_{i,t})\frac{t^2}{2}$ (where $0 \le C_{i,t} \le t$)

$$= -\frac{t^2}{2}\mathbb{E}X_{n,i}^2 + \underbrace{\left(\varphi^{(2)}_{X_{n,i}}(0) - \varphi^{(2)}_{X_{n,i}}(C_{i,t})\right)}_{\mathbb{E}X_{n,i}^2(1 - e^{iC_{i,t}X_{n,i}})}\frac{t^2}{2}$$

Among

$$\left|\sum_{i=1}^{k_n} \mathbb{E}\, X_{n,i}^2 (1-e^{iC_{i,t}X_{n,i}})\right|$$

$$\le 2\sum_{i=1}^{k_n} \mathbb{E}\, X_{n,i}^2\, \mathbb{1}(|X_{n,i}|>\varepsilon)$$

$$+ \sum_{i=1}^{k_n} \mathbb{E}\, X_{n,i}^2\, |1-e^{iC_{i,t}X_{n,i}}|\, \mathbb{1}(|X_{n,i}|\le\varepsilon)$$

$$\le \sum_{i=1}^{k_n} t\,\mathbb{E}\, X_{n,i}^3\, \mathbb{1}(|X_{n,i}|\le\varepsilon)$$

$$\le t\varepsilon \sum_{i=1}^{k_n} \mathbb{E}X_{n,i}^2 \le t\varepsilon$$

$$\Rightarrow \log\varphi_{\sum_{i=1}^{k_n} X_{n,i}}(t) = -\frac{t^2}{2} + o(1)$$

$\square$

Remark:

① $X_1, \cdots, X_n$ is independent, and

$\triangle$ $\mathbb{E} X_i = \mu_i$

$\triangle$ $\sum_{i=1}^{n} \operatorname{Var} X_i := \sigma_n^2 < \infty$

$\mathbb{E}\, X_i^2 \mathbb{1}(|X_i| \geq \sigma_n \varepsilon) / \sigma_n^2 < \infty$,

$\forall \varepsilon > 0$. Then:

$$\frac{\sum_{i=1}^{n} X_i - \mu_i}{\sigma_n} \to_w N(0, 1)$$

② (Lyapunov) If $\exists\, \delta > 0$,

s.t.

$$\frac{\sum_{i=1}^{n} \mathbb{E}\, |X_i|^{2+\delta}}{\sigma_n^{2+\delta}} \to 0,$$

then

$$\frac{\sum_{i=1}^{n} X_i - \mu_i}{\sigma_n} \to_w N(0,1),$$

where $X_1, \cdots, X_n$ is independent.

Stein's Method

Let $Z \sim N(0,1)$, define Stein's equation:

$$f'(w) - wf(w) = \mathbb{1}(w \le x) - \Phi(x),$$

where $\Phi(x) = \mathbb{P}(Z \le x)$.

Indeed, $\forall x$, $\exists!$ bounded f to achieve stein's equation.

If $X \sim N(0,1)$, then

$\mathbb{E} f'(X) - Xf(X) = 0$ for some bound, differentiable functions.

Actually, if f is bounded, absolutely continuous, $|f'| < \infty$, then

$$\mathbb{E} f'(X) = \mathbb{E} X f(X),$$

which is called the Stein's identity.

Now, let $X_1, \dots, X_n$ iid s.t. $\mathbb{E} X_i = 0$, $\sum_{i=1}^{n} \mathbb{E} X_i^2 = 1$, and $\gamma = \mathbb{E}|X_1|^3 < \infty$, $S_n = \sum_{i=1}^{n} X_i$, then

Berry - Esseen

$$\sup_t | \mathbb{P}(S_n \le t) - \Phi(t) | \lesssim n\gamma$$

Pf: $\mathbb{P}(S_n \le t) - \Phi(t)$

$$= \mathbb{E}\left(\mathbb{1}(S_n \le t) - \Phi(t)\right), \exists f_t$$

$$= \mathbb{E}\, f_t'(S_n) - S_n f_t(S_n) \quad (1)$$

Among,

$$\mathbb{E}\, S_n f(S_n) = n\, \mathbb{E}\, X_n f(S_n)$$

$$= n\, \mathbb{E}\, X_n \left(f(S_n) - f(S_{n-1})\right)$$

$$= n\, \mathbb{E}\, X_n \int_0^{X_n} f'(t + S_{n-1})\, dt$$

$$= n\, \mathbb{E}_{S_{n-1}} \underline{\mathbb{E}_{X_n} X_n \int_0^{X_n} f'(t + S_{n-1})\, dt}$$

$$= \mathbb{E}_{X_n} \int \big(\mathbb{1}(0 \le t \le X_n) + \mathbb{1}(X_n \le t \le 0)\big) \cdot X_n f'(t + S_{n-1})\, dt$$

$$= \int_{-\infty}^{\infty} \mathbb{E}_{X_n} X_n f'(t + S_{n-1}) \underline{\quad}\, dt,$$

then

$$\mathbb{E} S_n f(S_n) = \mathbb{E} \int_{-\infty}^{\infty} f'(t + S_{n-1})\, k(t)\, dt,$$

where

$$k(t) := n\, \mathbb{E} X_n \big(\mathbb{1}(0 \le t \le X_n) + \mathbb{1}(X_n \le t \le 0)\big),$$

and

$$\int_{-\infty}^{\infty} k(t)\, dt = n \int_{-\infty}^{0} \mathbb{E} X_n \mathbb{1}(X_n \le t)\, dt + n \int_{0}^{+\infty} \mathbb{E} X_n \mathbb{1}(X_n \ge t)\, dt$$

$$= n \int_{0}^{+\infty} \mathbb{E} |X_n| \mathbb{1}(|X_n| \ge t)\, dt$$

$= n\,\mathbb{E}\,X_n^2 = 1$

Let $T \sim K(t)$, $T \perp X_1, \cdots, X_n$

$\Rightarrow \mathbb{E}\, S_n f(S_n) = \mathbb{E} f'(S_{n-1}+T)$

(Stein identity)

Then $(1) = \mathbb{E} f_x'(S_n) - f_x'(S_{n-1}+T)$

$= \mathbb{E}\, S_n f_x(S_n) - (S_{n-1}+T) f_x(S_{n-1}+T)$

$+ \mathbb{E}(\mathbb{1}(S_n \le x) - \mathbb{1}(S_{n-1}+T \le x))$

$\le \mathbb{E}|X_n| + \mathbb{E}|T| +$

$\mathbb{P}(x - \max\{X_n, T\} \le S_{n-1} \le x - \min\{X_n, T\})$

where

$\mathbb{E}|T| = n\int_0^{+\infty} \mathbb{E}\; |X_n||t|\,\mathbb{1}(|X_n| \ge t)\,dt$

$= n\,\mathbb{E}\,\frac{|X_n|^3}{2} = \frac{n\gamma}{2}$

Since $\mathbb{E}|X_n|^3 \geq \mathbb{E}|X_n| \, \mathbb{E}|X_n|^2$

$$= \mathbb{E}|X_n| / n$$

$$\Rightarrow \mathbb{E}|X_n| + \mathbb{E}|T| \leq \mathbb{E}|T|$$

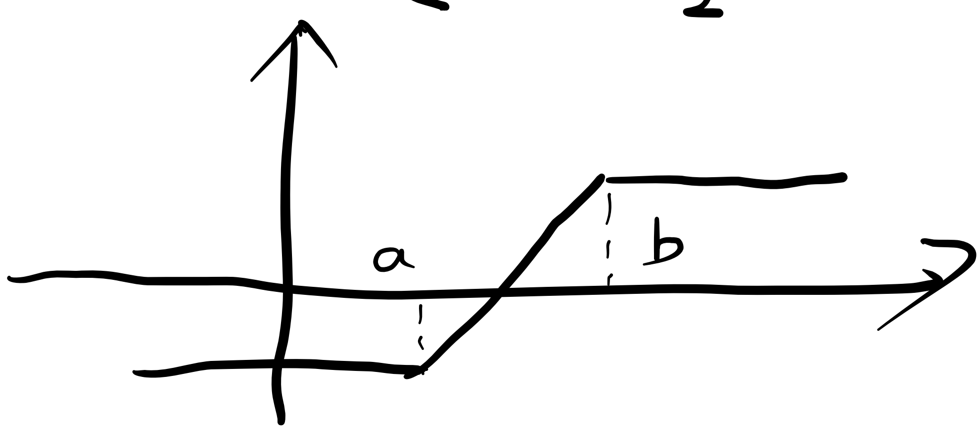
$$= n r$$

Note that $\forall$ reasonable f,

$$\mathbb{E} S_n f(S_n) = \mathbb{E} f'(S_{n-1} + T)$$

Let $f'(w) = \mathbb{1}(a \leq w \leq b)$

$$|f(w)| \leq \frac{b-a}{2}$$



Then $|\mathbb{E}\, S_n f(S_n)|$

$$\leq \frac{b-a}{2}\,\mathbb{E}|S_n|$$

$$\leq \frac{b-a}{2}\,(\mathbb{E}\, S_n^2)^{\frac{1}{2}}$$

$$\leq \frac{b-a}{2}$$

and $\mathbb{E}\, S_n f(S_n)$

$$= \mathbb{P}(a \leq S_{n-1} + T \leq b)$$

$$\geq \mathbb{P}(a-\delta \leq S_{n-1} \leq b+\delta,\ |T| \leq \delta)$$

$$\geq \mathbb{P}(a-\delta \leq S_{n-1} \leq b+\delta)\left(1 - \frac{\mathbb{E}|T|}{\delta}\right)$$

then (Stein's concentration)

$$\mathbb{P}(a-\delta \leq S_{n-1} \leq b+\delta) \leq \frac{b-a}{1-\frac{\mathbb{E}|T|}{\delta}}$$

$\Rightarrow (\delta = 2\mathbb{E}|T|)$

$\mathbb{P}(x - \max\{X_n, T\} \leq S_{n-1} \leq x - \min\{X_n, T\} \mid X_n, T)$

$\lesssim x - \min\{X_n, T\} - 2\mathbb{E}(T)$

$- (x - \max\{X_n, T\} + 2\mathbb{E}(T))$

$= |X_n - T| + 4\mathbb{E}|T|$

$\Rightarrow \mathbb{P}(x - \max\{X_n, T\} \leq S_{n-1} \leq x - \min\{X_n, T\})$

$\lesssim \mathbb{E}|T|$

$\Rightarrow$

$\sup_t |\mathbb{P}(S_n \leq t) - \Phi(t)| \lesssim \mathbb{E}|T|$

$= n\gamma$ $\square$

Remark:

① Since $\mathbb{E}|X_n| \leq \frac{1}{\sqrt{n}}$

$\Rightarrow |X_n| = O_p(\frac{1}{\sqrt{n}}) \Rightarrow |X_n|^3 = O_p((\frac{1}{\sqrt{n}})^3)$

$\Rightarrow \quad \mathbb{E}|X_n|^3 = O\left(\left(\frac{1}{\sqrt{n}}\right)^3\right)$

$\Rightarrow \quad n\gamma = O\left(\frac{1}{\sqrt{n}}\right)$

② If we just know $\sigma^2 = \mathbb{E}X_i^2$ and $\gamma = \mathbb{E}|X_i|^3 < \infty$, let

$$Y_i = \frac{X_i}{\sqrt{n}\,\sigma}$$

$$\Rightarrow \sup_t \left|\mathbb{P}\left(\sum_{i=1}^{n} Y_i \le t\right) - \Phi(t)\right|$$

$$\lesssim n\,\mathbb{E}|Y_n|^3$$

$$= n\,\mathbb{E}|X_n|^3 / (n\sqrt{n}\,\sigma^3)$$

$$= \frac{\gamma}{\sqrt{n}\,\sigma^3}.$$