Convergence Concepts

(1) Relations:

① **Thm.** For $r \geq 1$:

$$\left.\begin{array}{l} X_n \longrightarrow X \text{ a.s} \\ X_n \longrightarrow X \text{ in } L^r \end{array}\right\} \Rightarrow X_n \xrightarrow[p]{} X \Rightarrow X_n \xrightarrow[d]{} X.$$

For $r > s > 0$. $X_n \longrightarrow X$ in $L^r \Rightarrow X_n \longrightarrow X$ in L^s.

Remark: Some counterexamples:

i) $X_n \xrightarrow[p]{} X \not\Rightarrow X_n \longrightarrow X$ a.s or in L^r:

$p(X_n = 1) = 1/n$. $p(X_n = 0) = 1 - \frac{1}{n}$.

ii) $X_n \xrightarrow[d]{} X \not\Rightarrow X_n \xrightarrow[p]{} X$:

$X_n = -X \sim N(0,1)$

iii) $X_n \longrightarrow X$ a.s $\not\Leftrightarrow X_n \longrightarrow X$ in L^r.

$\not\Rightarrow$: $p(X_n = 0) = 1 - \frac{1}{n^2}$. $p(X_n = n^3) = \frac{1}{n^2}$

$\not\Leftarrow$: $p(X_n = 0) = 1 - \frac{1}{n}$. $p(X_n = 1) = \frac{1}{n}$

② **Partial Converse:**

i) **Thm.** $X_n \xrightarrow[d]{} c \Leftrightarrow X_n \xrightarrow[p]{} c$. for const. c.

Pf: Written in d.f. for $p(|X_n - c| \geq \varepsilon)$.

ii) <u>Monotone r.v's:</u>

<u>Thm.</u> $\{X_n\}$. mono r.v's. $X_n \xrightarrow{p} X \Rightarrow X_n \to X$ a.s.

<u>Pf:</u> <u>Lemma.</u> $X_n \le Y_n \le Z_n$. $X_n, Z_n \to Y$. a.s.

Then $Y_n \to Y$. a.s.

<u>Pf:</u> $p(\{w \mid Y_n(w) \to Y(w)\}) \ge p(\{w \mid X_n(w) \to X(w)\} \cap \{w \mid Z_n \to Z\})$

$\ge 1 - p(\{\cdots\}^c) - p(\{\cdots\}^c) = 1$.

$\Rightarrow$ WLOG. $X_n \ge X_{n+1}$. $\forall n \in \mathbb{Z}^+$. Since $\exists \{X_{n_k}\} \subset \{X_n\}$. st.

$X_{n_k} \to X$. a.s. Then $\forall n$. $\exists k$. $X_{n_k} \ge X_n \ge X_{n_{k+1}}$.

iii) <u>Converge completely:</u>

<u>Def:</u> $X_n \to X$ completely if $\forall \varepsilon > 0$. $\sum p(|X_n - X| > \varepsilon) < \infty$.

<u>Thm.</u> $X_n \to X$ completely $\Rightarrow X_n \to X$. a.s.

<u>Cor.</u> $\sum E(|X_n - X|^r) < \infty \Rightarrow X_n \to X$. a.s. $(r > 0)$.

iv) <u>Converge in another space:</u>

<u>Thm.</u> (Skorokhod's Representation)

If $X_n \xrightarrow{d} X$. Then $\exists$ r.v's Y. $\{Y_n\}$. on space

$((0,1), \mathcal{B}_{(0,1)}, m|_{(0,1)})$. m is Lebesgue measure. st.

$X_n \sim Y_n$. $X \sim Y$. $Y_n \to Y$. a.s.

<u>Pf:</u> Denote F_n. F are d.f of X_n. X.

Set $Y_n(t) = F_n^{-1}(t)$. $Y(t) = F^{-1}(t)$.

$\therefore Y_n \sim X_n$. $Y \sim X$.

Fix $\varepsilon > 0$, $t' \in (0,1)$. $\exists\, x \in C(F)$. $Y(t') < x < Y(t') + \varepsilon$

$\therefore F(x) \geq t' > t$. $\because F_n \to F$ at x. $\therefore \exists N$. $n > N$, st.

$F_n(x) > t$. $\therefore Y_n(t) < x < Y(t') + \varepsilon$.

$\therefore \overline{\lim_n} Y_n \leq Y(t')$. Choose $t \in C(F)$. Let $t' \to t$.

$\therefore \overline{\lim} Y_n \leq Y$. on $C(F)$. Similarly. $\underline{\lim} Y_n \geq Y$.

$\therefore P(\{t \mid Y_n(t) \not\to Y(t)\}) \leq P((0,1)/C(F)) = 0$.

(2) <u>Convergence in Moments:</u>

① <u>Pratt's Lemma.</u>

i) $X_n \leq Y_n \leq Z_n$. a.s.

ii) $X_n \to X$ a.s. $Y_n \to Y$. a.s. $Z_n \to Z$. a.s.

iii) $E(X_n) \to E(X)$. $E(Z_n) \to E(Z)$. $X, Z \in L^1$.

$\Rightarrow E(Y_n) \to E(Y)$.

<u>Pf</u>: Apply Fatou's Lemma on $Z_n - Y_n$. $Y_n - X_n$.

② <u>Dominated Converge:</u>

<u>Lemma</u>. $X_n \xrightarrow[P]{} X$, $P(|X_n| \leq Y) = 1$. $\Rightarrow P(|X| \leq Y) = 1$.

<u>Pf</u>: $\forall \delta > 0$. $P(|X| > Y + \delta) =$

$P(\square, |X_n| \leq Y) + P(\square, |X_n| > Y)$

$\leq P(|X| > |X_n| + \delta, |X_n| \leq Y)$

$\leq P(|X - X_n| > \delta) \to 0$.

Lemma. $Y \in L^1$. $P(A_n) \to 0$. $\Rightarrow E_{A_n}(|Y|) \to 0$.

Pf: By MCT: $E(|Y| I_{\{|Y|>n\}}) \to 0$.

Truncation: $E_{A_n}(|Y|) = E_{A_n}(|Y| I_{\{|Y|\geq N\}} + |Y| I_{\{|Y|<N\}})$

$\leq \varepsilon + N P(A_n) \to 0$

Thm. If $X_n \xrightarrow[p]{} X$. $|X_n| \leq Y$. a.s. $\forall n$. $Y \in L^r$. $r>0$.

Then $X_n \to X$ in L^r. (So: $E(X_n^r) \to E(X^r)$).

Pf: By Lemma. $|X_n - X| \leq 2|Y|$. a.s.

$E(|X_n - X|^r) = E(|X_n - X|^r (I_{\{|X_n - X| \geq \varepsilon\}} + I_{\{|X_n - X| < \varepsilon\}}))$.

$\leq 2^r E(|Y| I_{A_n}) + \varepsilon^r$. (Apply Lemma.).

Remark: $X_n \to X$. a.s. $X \in L^r$. $\not\Rightarrow$ $X_n \to X$ in L^r.

e.g. $P(X_n = 2^n) = \frac{1}{2^n}$. $P(X=0) = 1 - \frac{1}{2^n}$. $X_n \to 0$. a.s.

Actually. $|X_n| \leq X + \varepsilon$. a.s for large n is wrong!

Cor. If $|X_n| \leq c$. a.s. Then $\forall r>0$. we have:

$X_n \to X$ in L^r $\Leftrightarrow$ $X_n \xrightarrow[p]{} X$.

Cor. $X_n \xrightarrow[p]{} 0 \Leftrightarrow E(\frac{|X_n|}{1+|X_n|}) \to 0$.

Pf: $P(|X_n| \geq \varepsilon) = P(\frac{|X_n|}{1+|X_n|} \geq \frac{\varepsilon}{1+\varepsilon})$

$\therefore \frac{|X_n|}{1+|X_n|} \xrightarrow[p]{} 0 \Leftrightarrow X_n \xrightarrow[p]{} 0$.

③ Uniformly Integrable:

i) Definitions:

· Because the dominated condition: $|X_n| \le Y \in L^r$ is strong. We will introduce a weaker condition: uniform integrability.

Def: $\{Y_i\}_{i\in I}$ r.v's on $(\Omega, \mathcal{A}, P)$ is u.i. iff

$$\lim_{c\to\infty} \sup_{i\in I} E(|Y_i| I_{[|Y_i|\ge c]}) = 0$$

Remark: (a) It motivates by $X \in L^1 \Leftrightarrow$

$E(|X| I_{[|X|>k]}) \to 0 \ (k\to\infty)$.

(b) It can imply:

$\sup_{i\in I} E(|X_i|) \le M < \infty$. But

converse doesn't hold!

Thm. (Criteria)

$\varphi \ge 0$. st. $\varphi(x)/x \to \infty \ (x\to\infty)$. If $\forall i \in I$.

$E(\varphi(|X_i|)) \le C < \infty$. Then $\{X_i\}_{i\in I}$ is u.i.

Pf: Denote $\varepsilon_m = \sup\{x/\varphi(x) \mid x \ge m\} \to 0 \ (m\to\infty)$

Then $E(|X_i| I_{[|X_i|\ge m]}) \le \varepsilon_m E(\varphi(|X_i|) I_{[|X_i|\ge m]})$

$\le C\varepsilon_m \to 0 \ (m\to\infty)$

Remark: e.g. $\varphi = x^p$. $p>1$. $(x\log x)^+$.

Lemma. (Absolute continuity)

If $X \in L^1$. Then $Q(A) = E_A(X)$ is absolutely conti. i.e. $\forall \varepsilon > 0.\ \exists \delta > 0.\ \forall A \in \mathcal{A}.\ P(A) < \delta \Rightarrow Q(A) < \varepsilon$.

Pf: Truncate X: $I_{\{|X| \le m\}}.\ I_{\{|X| > m\}}$.

Thm. (Equivalence Def for u.i)

r.v's $\{Y_i\}$ on $(\Omega, \mathcal{A}, P)$ is u.i. $\Leftrightarrow$ (a) $\sup_i E(|Y_i|) < \infty$

(b) $\forall \varepsilon > 0.\ \exists \delta > 0.$ s.t. $\forall A \in \mathcal{A}$. $P(A) < \delta \Rightarrow \sup_i E_A |Y_i| < \varepsilon$.

$(P(A_n) \to 0 \Rightarrow \sup_i E_{A_n} |Y_i| \to 0)$

Pf: ($\Rightarrow$) Trivial. By truncation.

($\Leftarrow$) Denote $M = \sup_{i \in I} E(|Y_i|) < \infty$. For $\forall \varepsilon > 0.\ \exists \delta > 0$.

$P(|Y_i| \ge c) \le M/c$. choose $c > M/\delta$.

$\therefore \sup_{i \in I} E_{\{|Y_i| > c\}} (|Y_i|) < \varepsilon$.

ii) Thm. (a) $\{X_i\}$ u.i $\Leftrightarrow \{|X_i|\}$. u.i $\Leftrightarrow \{X_i^+\}. \{X_i^-\}$. u.i.

(b) $|X_i| \le |Y_i|$. $\{Y_i\}$ u.i $\Rightarrow \{X_i\}$ u.i.

(c) $\{X_i\}. \{Y_i\}$ u.i $\Rightarrow \{X_i + Y_i\}$ u.i.

(d) If $|X_i| \le Y \in L^1$. a.s. $\forall i$. Then $\{X_i\}$ u.i.

Pf. (a). (b) are trivial.

(c) $|X_i + Y_i| I_{\{|X_i+Y_i| \geq c\}} \leq |X_i| I_{\{|X_i| \geq \frac{c}{2}\}} + |Y_i| I_{\{|Y_i| \geq \frac{c}{2}\}}$.

(d) $\sup_{i \in I} E(|X_i| I_{\{|X_i| \geq c\}}) \leq E(|Y| I_{\{|Y| \geq c\}}) \to 0$

Cor. If $\exists Y \in L^1$. $P(|Y_n| \geq \eta) \leq P(|Y| \geq \eta)$. $\forall \eta > 0$

Then $\{Y_n\}$ is u.i.

Pf: $E(|Y_n| I_{\{|Y_n| \geq c\}}) = \int_0^c + \int_c^\infty P(|Y_n| \geq \eta) I \, d\eta$

$$= c P(|Y_n| \geq c) + \int_c^\infty P(|Y_n| \geq \eta) d\eta$$

$$\leq c P(|Y| \geq c) + \int_c^\infty P(|Y| \geq \eta) d\eta$$

$$= E(|Y| I_{\{|Y| \geq c\}}) \to 0 \ (c \to \infty)$$

Remark: Denote $Y \geq_{stoch} Y_n$. if $F_Y(\eta) \leq F_{Y_n}(\eta)$. $\forall \eta$.

iii) Converge in pr. + u.i $\Rightarrow$ Converge in L^r:

Thm. (Vitali's)

$X_n \xrightarrow{p} X$. $X_n \in L^r$. $\forall n$. Then followings are equivalent.

(a) $\{X_n^r\}$ u.i.

(b) $X_n \to X$ in L^r. $X \in L^r$.

(c) $E(|X_n|^r) \to E(|X|^r) < \infty$.

Pf: (a) $\Rightarrow$ (b):

1°) $X \in L^r$.

since exist $\{X_{n_k}\} \subseteq \{X_n\}$. $X_{n_k} \to X$. a.s.

By Fatou's: $E(|X|^r) \le \underline{\lim} E(|X_{n_k}|^r) \le \sup E(|X_n|) < \infty$.

2°) $\{|X_n - X|\}$. u.i.

Lemma. (Cr - Inequility)

$$|x+y|^r \le Cr(|x|^r + |y|^r). \quad Cr = \begin{cases} 1, & 0<r<1 \\ 2^{r-1}, & 1 \le r. \end{cases}$$

Pf: $0<r<1$: $\lambda^r + (1-\lambda)^r \ge \lambda + 1 - \lambda = 1$. $\lambda = |x|/|x|+|y|$.

$1 \le r$: by convexity of $|x|^r$.

$\Rightarrow |X_n - X|^r \le Cr(|X_n|^r + |X|^r)$. u.i.

3°) $X_n \to X$ in L^r:

$$E(|X_n - X|^r) = E(\square I_{\{|X_n-X| \ge c\}} + I_{\{\varepsilon \le |X_n - X| \le c\}} + I_{\{|X_n - X| \le \varepsilon\}})$$

$$\le c^r P(|X_n - X| \ge \varepsilon) + \varepsilon^r + E(|X_n - X|^r I_{\{|X_n - X| \ge c\}})$$

$$\to \varepsilon^r + E(|X_n - X|^r I_{\{|X_n-X| \ge c\}} \to \varepsilon^r.$$

(Directly. $E(|X_n - X|^r I_{\{|X_n - X| \ge \varepsilon\}}) \to 0$ $(n \to \infty)$. by 2^{nd}-def).

(b) $\Rightarrow$ (c): Cr - Inequility. Minkovsky - Inequility

(c) $\Rightarrow$ (a): Set $f_A(x) \in C_B$. st.

$$f_A(x) = \begin{cases} |x|^r. & |x| \le A \\ 0, & |x| \ge A+1. \end{cases} \quad \therefore \lim E(f_A(X_n)) = E(f_A(X)).$$

$$\underline{\lim} E(|X_n|^r I_{\{|X_n| \le A+1\}}) \ge \underline{\lim} E(f_A(X_n)).$$

$$\ge E(f_A(X)) \ge E(|X|^r I_{\{|X| \le A\}})$$

$\Rightarrow \overline{\lim} E(|X_n|^r I_{\{|X_n| \geq A+1\}}) \leq E(|X|^r I_{\{|X|>A\}})$

$\forall \varepsilon > 0.\ \exists A(\varepsilon).\ n_0(A(\varepsilon)).\ \sup_{n \geq n_0} E(|X_n|^r I_{\{|X_n| \geq A+1\}}) < \varepsilon.$

Cr. $X_n \to X$ in L^r. $r > 0$. $X \in L^r$. Then

(a) $E(|X_n|^r) \to E(|X|^r)$. (b) $E(X_n^r) \to E(X^r)$.

Pf. (b) is from (a). since $\{X_n^r\}$ is u.i.

$X_n^r \to X^r$ in L^1.

(a) By. Cr. Minkovsky Inequility. analogously!

Remark: $X_n \to X$ in $L^r \not\Rightarrow X_n$ or $X \in L^r$.

e.g. $X_n \sim X \sim$ Cauchy. $r = 1$.

iv) Converge in dist. + u.i $\Rightarrow$ Converge in moments:

Thm. $X_n \xrightarrow{d} X$. $\{X_n^r\}$ u.i. Then. we have:

(a) $E(|X|^r) < \infty$. (b). $\lim E(|X_n|^r) = E(|X|^r)$. (c) $\lim E(X_n^r) = E(X^r)$.

Pf: Apply Skorokhod's Representation.

Rmk: $X_n - X \nrightarrow Y_n - Y$. even if $X_n \sim Y_n$. $X \sim Y$.

$\therefore X_n \nrightarrow X$ in L^r. Commonly.

Besides X_n's may not be in the same prob. space.

(3) Closed Operations:

① Algebraic Operations:

Thm. $X_n \to X$. $Y_n \to Y \Rightarrow X_n \pm Y_n \to X \pm Y$.

It holds for converging a.s / in pr / in L^r.

Pf: By $\{|X_n + Y_n| \geq \varepsilon\} \subseteq \{|X_n| \geq \frac{\varepsilon}{2}\} \cup \{|Y_n| \geq \frac{\varepsilon}{2}\}$. Cr-Inequality.

Rmk: It fails when converge in dist: e.g. $X \sim N(0,1)$.

$X_n = X = -Y = Y_n$. (Actually. $X_1 \sim X_2$. $Y_1 \sim Y_2$. Then: $X_1 + Y_1 \nsim X_2 + Y_2$).

Generally, $X_n \xrightarrow{d} X$. $Y_n \xrightarrow{d} Y \not\Rightarrow g(X_n, Y_n) \xrightarrow{d} g(X, Y)$. for $g \in C(\mathbb{R}^2)$. except: $(X_n, Y_n) \xrightarrow{d} (X, Y)$.

Thm. $X_n \xrightarrow{d} X_\infty$. $Y_n \xrightarrow{d} Y_\infty$. X_n. Y_n indept. $\forall n \in \mathbb{Z}^+$. Then:

$\exists X \sim X_\infty$. $Y \sim Y_\infty$. $X_n \pm Y_n \xrightarrow{d} X \pm Y$.

Pf. $\varphi_{X_n+Y_n} = \varphi_{X_n} \varphi_{Y_n} \to \varphi_X \varphi_Y$. where we choose X. Y. are just random sample (size = 1). from X_∞. Y_∞.

$\therefore X$. Y indept. $\varphi_X \varphi_Y = \varphi_{X+Y}$. identical ch.f.

Rmk: X_∞. Y_∞ may not be indept. e.g. X_∞ isn't degenerated $Y_n = Y_\infty = X_\infty$. $\forall n$. even if Y_n indept with X_k. $\forall n.k$.

Also note that we only require X_n. Y_n are in the same probability space.

Thm. $X_n \to X$. $Y_n \to Y$. $\Rightarrow X_n Y_n \to XY$. holds

for convergence a.s. / in pr.

Pf: i) $\{X_n Y_n \not\to XY\} \subseteq \{X_n \not\to X\} \cup \{Y_n \not\to Y\}$.

ii) $P(|X_n Y_n - XY| \geq \varepsilon) \leq P(|X_n - X||Y_n| \geq \frac{\varepsilon}{2})$

$+ P(|Y_n - Y||X| \geq \frac{\varepsilon}{2})$.

For the former. separate: (latter is same)

$\Omega = \{|Y_n| = 0\} + \{0 < |Y_n| < M\} + \{M \leq |Y_n|\}$.

$\Rightarrow \leq P(|X_n - X| > \frac{\varepsilon}{2M}) + P(|Y_n| \geq M)$.

$\leq P(|X_n - X| > \frac{\varepsilon}{2M}) + P(|Y_n - Y| \geq \frac{M}{2}) + P(|Y| > \frac{M}{2})$

Firstly, fix large M. Then $n \to \infty$.

Remark: i) L^r doesn't holds: Choose $X_n = Y_n \in L^1 \cap (L^2)^c$.

$\therefore \|X_n\|_{L^2}^2 = \infty$.

dist doesn't hold: $X_n = Y_n = X = -Y \sim N(0,1)$.

ii) $X_n \to X$ in L^p. $Y_n \to Y$ in L^q. $1/p + 1/q = 1$.

$\Rightarrow X_n Y_n \to XY$ in L^1.

Pf: $X, Y \notin L^p, L^q$. It's trivial.

$X \in L^p$. $Y \in L^q$: $\exists N$. $n > N$. $X_n, Y_n \in L^p, L^q$.

$|X_n Y_n - XY| \leq |X_n||Y_n - Y| + |Y||X_n - X|$.

By Hölder Inequility.

① Transformations:

Thm. X_n, X are k-dimensional random vectors.

$g: \mathbb{R}^k \to \mathbb{R}'$. conti. Then. we have:

i) $X_n \to X$. a.s $\Rightarrow$ $g(X_n) \to g(X)$. a.s.

ii) $X_n \to_p X$. $\Rightarrow$ $g(X_n) \to_p g(X)$.

iii) $X_n \to_d X$ $\Rightarrow$ $g(X_n) \to_d g(X)$.

Pf: i) By conti. $\{X_n \to X\} \subset \{g(X_n) \to g(X)\}$.

ii) Fix $\varepsilon > 0$. Choose M large enough. st.

$p(|X| \geq M) < \varepsilon$. Then on $|X| \leq M + \varepsilon$:

$\exists \delta < \varepsilon$. St. $|g(x) - g(y)| < \varepsilon$ if $|x - y| < \delta$. $|x| \leq M$.

Then $\{|g(X_n) - g(X)| \geq \varepsilon, |X_n - X| < \delta\} \subseteq \{|X| > M\}$.

partion $p(|g(X_n) - g(X)| \geq \varepsilon) = p(\square, |X_n - X| < \delta) + p(\square, |X_n - X| \geq \delta)$

iii) By Skorokhod's Representation.

Cor. Extend to g is a.s. conti w.r.t P_X. i.e.

$p(\{w \mid g \text{ disconti at } X(w)\}) = 0$.

Pf: i) $\{X_n \to X\} \subset \{g(X_n) \to g(X)\} + \{g \text{ disconti at } X(w)\}$

ii) Similarly, separate $\{g \text{ conti at } X\} + \{g \text{ disconti}\}$

iii) For $\forall f \in C_B$. By Shorokhod Representation

since $p(\{w \mid f \circ g \text{ disconti at } X(w)\}) = 0$.

$\therefore f(g(X_n)) \to f(g(X))$. a.s.

Then Apply DCT. $\therefore g(X_n) \Rightarrow g(X)$.

(3) Slutsky Thm:

If $X_n \to_d X$. $Y_n \to_p c$. Then:

i) $X_n \pm Y_n \to_d X \pm c$

ii) $X_n Y_n \to_d cX$.

iii) $X_n / Y_n \to_d X/c$. for $c \neq 0$.

Pf: Consider $F_{X_n+Y_n}(x) = p(X_n+Y_n \leq x) = 1 - p(X_n+Y_n > x)$.

separate into : $\{|Y_n - c| \geq \varepsilon\} + \{|Y_n - c| < \varepsilon\}$.

Written into d.f. Then apply $\overline{\lim}$. $\underline{\lim}$.

Remark: The point is: if $X \sim Y$. c is const. Then

$X + c \sim Y + c$. (It's obvious by ch.f's)

(4) Cauchy Convergence:

$\{X_n\}$ converge in L^2. a.s. pr $\Leftrightarrow$ It's Cauchy in L^2. a.s. pr.

Pf. ($\Rightarrow$) is trivial.

($\Leftarrow$). Find $\{X_{n_k}\} \subseteq \{X_n\}$. $X_{n_k} \to X$. a.s.

i) L^2: $E(|X_n - X|^2) = E(\underline{\lim}_k |X_n - X_{n_k}|^2)$

$\leq \underline{\lim}_k E(|X_n - X_{n_k}|^2) \to 0$. (Fatou's)

ii) a.s: In Ω/N: $X_n(\omega) \to X(\omega) \in \mathbb{R}^1$. $\forall \omega$

iii) pr: $\{|X_n - X| \geq \varepsilon\} \subseteq \{|X_n - X_{n_k}| \geq \frac{\varepsilon}{2}\} \cup \{|X_{n_k} - X| \geq \frac{\varepsilon}{2}\}$