Lecture 11: Convergence modes and stochastic orders

$c = (c_1, ..., c_k) \in \mathcal{R}^k$, $\|c\|_r = (\sum_{j=1}^k |c_j|^r)^{1/r}$, $r > 0$.
If $r \geq 1$, then $\|c\|_r$ is the L_r-distance between 0 and c.
When $r = 2$, $\|c\| = \|c\|_2 = \sqrt{c^\tau c}$.

Definition 1.8. Let $X, X_1, X_2, \ldots$ be random k-vectors defined on a probability space.
(i) We say that the sequence $\{X_n\}$ converges to X almost surely (a.s.) and write $X_n \to_{a.s.} X$ if and only if $\lim_{n\to\infty} X_n = X$ a.s.
(ii) We say that $\{X_n\}$ converges to X in probability and write $X_n \to_p X$ if and only if, for every fixed $\epsilon > 0$,

$$\lim_{n\to\infty} P(\|X_n - X\| > \epsilon) = 0.$$

(iii) We say that $\{X_n\}$ converges to X in L_r (or in rth moment) and write $X_n \to_{L_r} X$ if and only if

$$\lim_{n\to\infty} E\|X_n - X\|_r^r = 0,$$

where $r > 0$ is a fixed constant.
(iv) Let F, F_n, $n = 1, 2, ...$, be c.d.f.'s on $\mathcal{R}^k$ and P, P_n, $n = 1, ...$, be their corresponding probability measures. We say that $\{F_n\}$ converges to F weakly (or $\{P_n\}$ converges to P weakly) and write $F_n \to_w F$ (or $P_n \to_w P$) if and only if, for each continuity point x of F,

$$\lim_{n\to\infty} F_n(x) = F(x).$$

We say that $\{X_n\}$ converges to X in distribution (or in law) and write $X_n \to_d X$ if and only if $F_{X_n} \to_w F_X$.

$\to_{a.s.}$, $\to_p$, $\to_{L_r}$: How close is between X_n and X as $n \to \infty$?

$F_{X_n} \to_w F_X$: X_n and X may not be close (they may be on different spaces)

Example 1.26. Let $\theta_n = 1 + n^{-1}$ and X_n be a random variable having the exponential distribution $E(0, \theta_n)$ (Table 1.2), $n = 1, 2, ...$ Let X be a random variable having the exponential distribution $E(0, 1)$. For any $x > 0$, as $n \to \infty$,

$$F_{X_n}(x) = 1 - e^{-x/\theta_n} \to 1 - e^{-x} = F_X(x)$$

Since $F_{X_n}(x) \equiv 0 \equiv F_X(x)$ for $x \leq 0$, we have shown that $X_n \to_d X$.

$X_n \to_p X$?
Need further information about the random variables X and X_n.
We consider two cases in which different answers can be obtained.
First, suppose that $X_n \equiv \theta_n X$ (then X_n has the given c.d.f.).
$X_n - X = (\theta_n - 1)X = n^{-1}X$, which has the c.d.f. $(1 - e^{-nx})I_{[0,\infty)}(x)$.

$$P(|X_n - X| \geq \epsilon) = e^{-n\epsilon} \to 0$$

for any $\epsilon > 0$. (In fact, by Theorem 1.8(v), $X_n \to_{a.s.} X$)
Since $E|X_n - X|^p = n^{-p}EX^p < \infty$ for any $p > 0$, $X_n \to_{L_p} X$ for any $p > 0$.

Next, suppose that X_n and X are independent random variables.
Since p.d.f.'s for X_n and $-X$ are $\theta_n^{-1}e^{-x/\theta_n}I_{(0,\infty)}(x)$ and $e^x I_{(-\infty,0)}(x)$, respectively, we have

$$P\left(|X_n - X| \le \epsilon\right) = \int_{-\epsilon}^{\epsilon}\int \theta_n^{-1}e^{-x/\theta_n}e^{y-x}I_{(0,\infty)}(x)I_{(-\infty,x)}(y)dxdy,$$

which converges to (by the dominated convergence theorem)

$$\int_{-\epsilon}^{\epsilon}\int e^{-x}e^{y-x}I_{(0,\infty)}(x)I_{(-\infty,x)}(y)dxdy = 1 - e^{-\epsilon}.$$

Thus, $P\left(|X_n - X| \ge \epsilon\right) \to e^{-\epsilon} > 0$ for any $\epsilon > 0$ and, therefore, $X_n \to_p X$ does not hold.

Proposition 1.16 (Pólya's theorem). If $F_n \to_w F$ and F is continuous on $\mathcal{R}^k$, then

$$\lim_{n\to\infty} \sup_{x\in\mathcal{R}^k} |F_n(x) - F(x)| = 0.$$

Lemma 1.4. For random k-vectors $X, X_1, X_2, \ldots$ on a probability space, $X_n \to_{a.s.} X$ if and only if for every $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(\bigcup_{m=n}^{\infty}\{\|X_m - X\| > \epsilon\}\right) = 0. \tag{1}$$

Proof. Let $A_j = \cup_{n=1}^{\infty}\cap_{m=n}^{\infty}\{\|X_m - X\| \le j^{-1}\}$, $j = 1, 2, ...$
Then

$$\bigcap_{j=1}^{\infty} A_j = \{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}$$

By Proposition 1.1(iii),

$$P(A_j) = \lim_{n\to\infty} P\left(\bigcap_{m=n}^{\infty}\{\|X_m - X\| \le j^{-1}\}\right) = 1 - \lim_{n\to\infty} P\left(\bigcup_{m=n}^{\infty}\{\|X_m - X\| > j^{-1}\}\right)$$

(1) holds for every $\epsilon > 0$ if and only if $P(A_j) = 1$ for every j, i.e., $P(\cap_{j=1}^{\infty}A_j) = 1$

$$P(A_j) \ge P(\bigcap_{j=1}^{\infty} A_j) = 1 - P(\bigcup_{j=1}^{\infty} A_j^c) \ge 1 - \sum_{j=1}^{\infty} P(A_j^c)$$

Lemma 1.5. (Borel-Cantelli lemma). Let A_n be a sequence of events in a probability space and $\limsup_n A_n = \cap_{n=1}^{\infty}\cup_{m=n}^{\infty} A_m$.
(i) If $\sum_{n=1}^{\infty} P(A_n) < \infty$, then $P(\limsup_n A_n) = 0$.
(ii) If $A_1, A_2, ...$ are pairwise independent and $\sum_{n=1}^{\infty} P(A_n) = \infty$, then $P(\limsup_n A_n) = 1$.
Proof. (i) By Proposition 1.1,

$$P\left(\limsup_{n\to\infty} A_n\right) = P\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m\right) = \lim_{n\to\infty} P\left(\bigcup_{m=n}^{\infty} A_m\right) \le \lim_{n\to\infty}\sum_{m=n}^{\infty} P(A_n) = 0$$

if $\sum_{n=1}^{\infty} P(A_n) < \infty$.

(ii) We prove the case of independent A_n's.

$$P\left(\limsup_{n\to\infty} A_n\right) = \lim_{n\to\infty} P\left(\bigcup_{m=n}^{\infty} A_m\right) = 1 - \lim_{n\to\infty} P\left(\bigcap_{m=n}^{\infty} A_m^c\right) = 1 - \lim_{n\to\infty} \prod_{m=n}^{\infty} P(A_m^c)$$

$$\prod_{m=n}^{n+k} P(A_m^c) = \prod_{m=n}^{n+k} [1 - P(A_m)] \le \prod_{m=n}^{n+k} \exp\{-P(A_m)\} = \exp\left\{-\sum_{m=n}^{n+k} P(A_m)\right\}$$

$(1 - t \le e^{-t} = \exp\{t\})$. Letting $k \to \infty$,

$$\prod_{m=n}^{\infty} P(A_m^c) = \lim_{k\to\infty} \prod_{m=n}^{n+k} P(A_m^c) \le \exp\left\{-\sum_{m=n}^{\infty} P(A_m)\right\} = 0.$$

See Chung (1974, pp. 76-78) for the pairwise independence A_n's.

The notion of $O(\cdot)$, $o(\cdot)$, and stochastic $O(\cdot)$ and $o(\cdot)$

In calculus, two sequences of real numbers, $\{a_n\}$ and $\{b_n\}$, satisfy $a_n = O(b_n)$ if and only if $|a_n| \le c|b_n|$ for all n and a constant c
$a_n = o(b_n)$ if and only if $a_n/b_n \to 0$ as $n \to \infty$

Definition 1.9. Let $X_1, X_2, \ldots$ be random vectors and $Y_1, Y_2, \ldots$ be random variables defined on a common probability space.
(i) $X_n = O(Y_n)$ a.s. if and only if $P(\|X_n\| = O(|Y_n|)) = 1$.
(ii) $X_n = o(Y_n)$ a.s. if and only if $X_n/Y_n \to_{a.s.} 0$.
(iii) $X_n = O_p(Y_n)$ if and only if, for any $\epsilon > 0$, there is a constant $C_\epsilon > 0$ such that $\sup_n P(\|X_n\| \ge C_\epsilon |Y_n|) < \epsilon$.
(iv) $X_n = o_p(Y_n)$ if and only if $X_n/Y_n \to_p 0$.

Since $a_n = O(1)$ means that $\{a_n\}$ is bounded, $\{X_n\}$ is said to be bounded in probability if $X_n = O_p(1)$.

$X_n = o_p(Y_n)$ implies $X_n = O_p(Y_n)$
$X_n = O_p(Y_n)$ and $Y_n = O_p(Z_n)$ implies $X_n = O_p(Z_n)$
$X_n = O_p(Y_n)$ does not imply $Y_n = O_p(X_n)$
If $X_n = O_p(Z_n)$, then $X_n Y_n = O_p(Y_n Z_n)$.
If $X_n = O_p(Z_n)$ and $Y_n = O_p(Z_n)$, then $X_n + Y_n = O_p(Z_n)$.
The same conclusion can be obtained if $O_p(\cdot)$ and $o_p(\cdot)$ are replaced by $O(\cdot)$ a.s. and $o(\cdot)$ a.s., respectively.
If $X_n \to_d X$ for a random variable X, then $X_n = O_p(1)$
If $E|X_n| = O(a_n)$, then $X_n = O_p(a_n)$, where $a_n \in (0, \infty)$.
If $X_n \to_{a.s.} X$, then $\sup_n |X_n| = O_p(1)$.