Lecture 9: Convergence modes and stochastic orders

Notation

c = (c_1, ..., c_k) ∈ ℜ^k, ∥c∥_r = (∑_{j=1}^{k} |c_j|^r)^{1/r}, r > 0.
If r ≥ 1, then ∥c∥_r is the L_r-distance between 0 and c.
When r = 2, ∥c∥ = ∥c∥_2 = √{c^T c}.

Definition 1.8 (Convergence modes)
Let X, X_1, X_2, ... 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 →_{a.s.} X iff \lim_{n→∞} X_n = X a.s.

(ii) We say that \{X_n\} converges to X in probability and write X_n →_p X iff, for every fixed ε > 0,

\[ \lim_{n→∞} P(∥X_n - X∥ > ε) = 0. \]
Lecture 9: Convergence modes and stochastic orders

Notation

\[ c = (c_1, ..., c_k) \in \mathbb{R}^k, \| c \|_r = (\sum_{j=1}^{k} |c_j|^r)^{1/r}, \quad r > 0. \]

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 (Convergence modes)

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 \( X_n \to_{a.s.} X \) iff \( \lim_{n \to \infty} X_n = X \) a.s.

(ii) We say that \( \{X_n\} \) converges to \( X \) in probability and write \( X_n \to_p X \) iff, for every fixed \( \varepsilon > 0 \),

\[ \lim_{n \to \infty} P(\|X_n - X\| > \varepsilon) = 0. \]
Definition 1.8 (continued)

(iii) We say that \( \{X_n\} \) converges to \( X \) in \( L_r \) (or in \( r \)th moment) and write \( X_n \xrightarrow{L_r} X \) iff
\[
\lim_{n \to \infty} E\|X_n - X\|_r^r = 0,
\]
where \( r > 0 \) is a fixed constant.

(iv) Let \( F, F_n, n = 1, 2, \ldots, \) be c.d.f.'s on \( \mathbb{R}^k \) and \( P, P_n, n = 1, \ldots, \) be their corresponding probability measures.
We say that \( \{F_n\} \) converges to \( F \) weakly (or \( \{P_n\} \) converges to \( P \) weakly) and write \( F_n \xrightarrow{w} F \) (or \( P_n \xrightarrow{w} P \)) iff, 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 \( X_n \xrightarrow{d} X \) iff \( F_{X_n} \xrightarrow{w} F_X \).
Remarks

- $\rightarrow_{a.s.}, \rightarrow_p, \rightarrow_{L_r}$: How close is between $X_n$ and $X$ as $n \to \infty$?
- $F_{X_n} \rightarrow_w F_X$: $F_{X_n}$ is close to $F_X$ but $X_n$ 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), $n = 1, 2, \ldots$.
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

$$X_n \rightarrow_d X.$$
Remarks

- $\rightarrow a.s., \rightarrow_p, \rightarrow_{L_r}$: How close is between $X_n$ and $X$ as $n \rightarrow \infty$?
- $F_{X_n} \rightarrow_w F_X$: $F_{X_n}$ is close to $F_X$ but $X_n$ 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), $n = 1, 2, \ldots$.

Let $X$ be a random variable having the exponential distribution $E(0, 1)$.

For any $x > 0$, as $n \rightarrow \infty$,

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

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

$$X_n \rightarrow_d X.$$
Need further information about the random variables $X$ and $X_n$. We consider two cases in which different answers can be obtained.

### Case 1

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.

$$(1 - e^{-nx})I_{[0,\infty)}(x).$$

Then, $X_n \to_p X$ because, for any $\varepsilon > 0$,

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

(In fact, by Theorem 1.8(v), $X_n \to_{a.s.} X$)

Also, $X_n \to_{L_p} X$ for any $p > 0$, because

$$E|X_n - X|^p = n^{-p}EX^p \to 0$$
$X_n \to_p X$?

- Need further information about the random variables $X$ and $X_n$.
- We consider two cases in which different answers can be obtained.

**Case 1**

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.

$$(1 - e^{-nx})I_{[0, \infty)}(x).$$

Then, $X_n \to_p X$ because, for any $\varepsilon > 0$,

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

(In fact, by Theorem 1.8(v), $X_n \to_{a.s.} X$)

Also, $X_n \to_{L^p} X$ for any $p > 0$, because

$$E|X_n - X|^p = n^{-p}EX^p \to 0$$
Case 2

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

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

which converges to (by the dominated convergence theorem)

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

Thus,

$$P(|X_n - X| \geq \varepsilon) \to e^{-\varepsilon} > 0$$

for any $\varepsilon > 0$ and, therefore, $X_n \xrightarrow{p} X$ does not hold.
Proposition 1.16 (Pólya’s theorem)

If \( F_n \xrightarrow{w} F \) and \( F \) is continuous on \( \mathbb{R}^k \), then

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

This proposition implies the following useful result:

If \( F_n \xrightarrow{w} F \) a continuous \( F \) and \( c_n \in \mathbb{R}^k \) with \( c_n \to c \), then

\[ F_n(c_n) \to F(c). \]

Lemma 1.4

For random \( k \)-vectors \( X, X_1, X_2, \ldots \) on a probability space, \( X_n \xrightarrow{a.s.} X \) iff for every \( \varepsilon > 0 \),

\[
\lim_{n \to \infty} P \left( \bigcup_{m=n}^{\infty} \left\{ \| X_m - X \| > \varepsilon \right\} \right) = 0.
\]
Proposition 1.16 (Pólya’s theorem)

If \( F_n \rightarrow_w F \) and \( F \) is continuous on \( \mathbb{R}^k \), then
\[
\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}^k} |F_n(x) - F(x)| = 0.
\]

This proposition implies the following useful result:
If \( F_n \rightarrow_w \) a continuous \( F \) and \( c_n \in \mathbb{R}^k \) with \( c_n \rightarrow c \), then
\[
F_n(c_n) \rightarrow F(c).
\]

Lemma 1.4

For random \( k \)-vectors \( X, X_1, X_2, \ldots \) on a probability space, \( X_n \rightarrow_{a.s.} X \) iff for every \( \varepsilon > 0 \),
\[
\lim_{n \rightarrow \infty} P \left( \bigcup_{m=n}^{\infty} \{ \| X_m - X \| > \varepsilon \} \right) = 0.
\]
Proposition 1.16 (Pólya’s theorem)

If \( F_n \rightarrow_w F \) and \( F \) is continuous on \( \mathbb{R}^k \), then

\[
\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}^k} |F_n(x) - F(x)| = 0.
\]

This proposition implies the following useful result:

If \( F_n \rightarrow_w \) a continuous \( F \) and \( c_n \in \mathbb{R}^k \) with \( c_n \rightarrow c \), then

\[ F_n(c_n) \rightarrow F(c). \]

Lemma 1.4

For random \( k \)-vectors \( X, X_1, X_2, \ldots \) on a probability space, \( X_n \rightarrow_{a.s.} X \) iff for every \( \varepsilon > 0 \),

\[
\lim_{n \rightarrow \infty} P \left( \bigcup_{m=n}^{\infty} \left\{ \|X_m - X\| > \varepsilon \right\} \right) = 0.
\]
Proof

It can be verified that

\[ \bigcap_{j=1}^{\infty} A_j = \{ \omega : \lim_{n \to \infty} X_n(\omega) = X(\omega) \}, \quad A_j = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{ \| X_m - X \| \leq j^{-1} \} \]

By Proposition 1.1(iii, continuity),

\[ P(A_j) = \lim_{n \to \infty} P \left( \bigcap_{m=n}^{\infty} \{ \| X_m - X \| \leq j^{-1} \} \right) \]

\[ = 1 - \lim_{n \to \infty} P \left( \bigcup_{m=n}^{\infty} \{ \| X_m - X \| > j^{-1} \} \right) \]

\[ P(\bigcup_{m=n}^{\infty} \{ \| X_m - X \| > \varepsilon \}) \to 0 \text{ for every } \varepsilon > 0 \text{ iff } P(A_j) = 1 \text{ for every } j, \]

which is equivalent to \( P(\bigcap_{j=1}^{\infty} A_j) = 1 \) (i.e., \( X_n \to a.s. X \)), because

\[ P(A_j) \geq P \left( \bigcap_{j=1}^{\infty} A_j \right) = 1 - P \left( \bigcup_{j=1}^{\infty} A_j^c \right) \geq 1 - \sum_{j=1}^{\infty} P(A_j^c) \]
Lemma 1.5 (Borel-Cantelli lemma)

Let $A_n$ be a sequence of events in a probability space and

$$\limsup_{n} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m.$$ 

(i) If $\sum_{n=1}^{\infty} P(A_n) < \infty$, then $P(\limsup_{n} A_n) = 0$.

(ii) If $A_1, A_2, \ldots$ are pairwise independent and $\sum_{n=1}^{\infty} P(A_n) = \infty$, then $P(\limsup_{n} A_n) = 1$.

Proof of Lemma 1.5 (i)

By Proposition 1.1,

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

where the last equality follows from the condition

$$\sum_{n=1}^{\infty} P(A_n) < \infty.$$
Lemma 1.5 (Borel-Cantelli lemma)

Let $A_n$ be a sequence of events in a probability space and

$$\limsup_{n} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m.$$

(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 of Lemma 1.5 (i)

By Proposition 1.1,

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

where the last equality follows from the condition

$$\sum_{n=1}^{\infty} P(A_n) < \infty.$$
Proof of Lemma 1.5 (ii)

We prove the case of independent $A_n$’s. See Chung (1974, pp. 76-78) for the pairwise independence $A_n$’s.

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

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

\[
(1 - t \leq e^{-t} = \exp\{t\}).
\]

Letting $k \to \infty$,

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

Hence,

\[
\lim_{n \to \infty} P\left(\bigcap_{m=n}^{\infty} A_m^c\right) = \lim_{n \to \infty} \prod_{m=n}^{\infty} P(A_m^c) = 0.
\]
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

- $a_n = O(b_n)$ iff $|a_n| \leq c|b_n|$ for all $n$ and a constant $c$
- $a_n = o(b_n)$ iff $a_n/b_n \to 0$ 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. iff $P(\|X_n\| = O(|Y_n|)) = 1$.
(ii) $X_n = o(Y_n)$ a.s. iff $X_n/Y_n \to_{a.s.} 0$.
(iii) $X_n = O_p(Y_n)$ iff, for any $\epsilon > 0$, there is a constant $C_\epsilon > 0$ such that

$$\sup_n P(\|X_n\| \geq C_\epsilon |Y_n|) < \epsilon.$$ 

(iv) $X_n = o_p(Y_n)$ iff $X_n/Y_n \to_p 0$. 
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

- $a_n = O(b_n)$ iff $|a_n| \leq c|b_n|$ for all $n$ and a constant $c$
- $a_n = o(b_n)$ iff $a_n/b_n \to 0$ 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. iff $P(\|X_n\| = O(|Y_n|)) = 1$.
(ii) $X_n = o(Y_n)$ a.s. iff $X_n/Y_n \to a.s. 0$.
(iii) $X_n = O_p(Y_n)$ iff, for any $\varepsilon > 0$, there is a constant $C_\varepsilon > 0$ such that

$$\sup_n P(\|X_n\| \geq C_\varepsilon |Y_n|) < \varepsilon.$$ 

(iv) $X_n = o_p(Y_n)$ iff $X_n/Y_n \to p 0$. 

UW-Madison (Statistics)
Stat 709 Lecture 9
2016 10 / 11
Discussions and properties

- 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 $X_n \rightarrow_d X$ 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 $X_n \rightarrow_a.s. X$, then $\sup_n |X_n| = O_p(1)$.