Convergence of Random Variables

Ma Xuejun

School of Mathematical Sciences

Soochow University

https://xuejunma.github.io
http://web.suda.edu.cn/xuejunma/

2019.09.23-09.28

Outline

- Last lecture: review some basic probability concepts; introduce the statistics
- 4 types of convergence
- Relationship between different types of convergence
- Stochastic orders

Terms

- Converge in probability; Converge in L^p; converge in quadratic mean; almost sure converge; converge in distribution;
- O_p, o_p

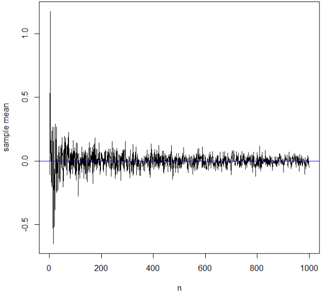
Note: May take 1-2 lectures for this topic.

Look into Sample mean

- Recall:

$$\text{Sample mean } \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

  Note: When $n \neq m$, X_n and X_m share the same expectation μ but have different distribution.
- Intuitively, when $n \to \infty$, $\bar{X}$ is very close to $\mu = E(X)$.



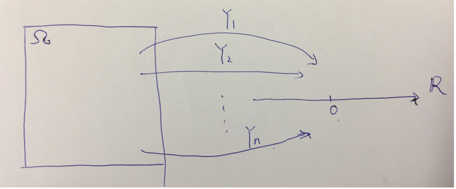
$N(0,1)$

```r
rm(list=ls())
n.vec <- seq(1, 10^3, 1)
n.len <- length(n.vec)
mean.full <- NULL
for(i in 1:n.len){
  mean.full[i] <- mean(rnorm(n.vec[i]))
}
plot(n.vec, mean.full, type="l",xlab = "n",
     ylab = "sample mean")
abline(h=0,lwd=1,col="blue")
```

If $x_1, x_2, \ldots, x_n, \ldots$ is an array of numbers, we know how to describe whether they convergence or not. But what if they are random variables? How to describe it?

Generalization

- Let $\{Y_n\}_{i=1}^{\infty} = Y_1, Y_2, \ldots, Y_n, \ldots$ denotes a sequence of random variables
- Problem: How to describe the limit of Y_n
- Consider 2 cases:
  - Case 1. $Y_i \sim F$ independently, $i = 1, 2, \ldots$
  - Case 2. $Z_1 = Z_2 = Z_3 = \ldots$, where $Z_1 \sim F$. Let $X \sim F$. Can we say $Y_i \to X$? Can we say $Z_i \to X$ How to differentiate these two cases?
- Recall: $Y_1, Y_2, \ldots, Y_n : \Omega \to R$. A sequence of functions



Convergence in Probability

Definition 5.5.1: Convergence in Probability

For a sequence of r.v.'s $\{X_n\}_{i=1}^{\infty} = X_1, X_2, \ldots, X_n, \ldots$, we say they converge in probability towards the r.v. X (i.e. $X_n \xrightarrow{p} X$) if for any $\varepsilon > 0$,

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

- The target X has the same sample space with all the X_i's
- $\{X_n\}$ are usually dependent, but not identically distribution.
- Practically, find the sequence of events $A_n = \{\omega \in \Omega, |X_n(\omega) - X(\omega)| \geq \varepsilon\}$ by obtaining $|X_n - X|$ as a new r.v., and check if $P(A_n) \to 0$ when $n \to \infty$.
- Interpretation: for any ε, the event that $|X_n - X|$ has probability smaller than δ when n is large enough. It concerns more about the probability measure and r.v., instead of the CDF only.

Example: Convergence in Probability

- Let X be a r.v. with prob 1 at 1, and $X_n \sim N\left(1, \frac{1}{n^2}\right)$. According to the property of normal distribution.,$X_n - X \sim N\left(0, \frac{1}{n^2}\right)$, so

$$P(|X_n - X| \geq \varepsilon) = P\left(|N(0, \frac{1}{n^2})|\right)$$
$$\leq \frac{1}{n^2 \varepsilon^2} \leq \delta, \;\; n \geq \frac{1}{\varepsilon\sqrt{\delta}}$$

So, $X_n \xrightarrow{p} X$.[1]

---

[1] Chebychev's inequality.

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\mathbf{Var}(X)}{\varepsilon^2}$$

Example: Convergence in Probability

- Let $X_n \sim Ber(0.5)$, and $X \sim Ber(0.5)$, X_n and X are independent. Note for any n,

$$
\begin{aligned}
&P(|X_n - X| \geq 1) \\
&= P\Big(\{X_n = 1, X = 0\} \cup \{X_n = 0, X = 1\}\Big) \\
&= P\Big(\{X_n = 1, X = 0\}\Big) + P\Big(\{X_n = 0, X = 1\}\Big) \\
&= \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2} \nrightarrow 0
\end{aligned}
$$

So, X_n does NOT converge to X in probability.

Convergence in L_r (rth mean)

Definition: Convergence in L_r

For a sequence of r.v.'s $\{X_n\}_{i=1}^{\infty} = X_1, X_2, \ldots, X_n, \ldots$, we say they converge in L_r towards the r.v. X (i.e. $X_n \xrightarrow{L^r} X$) if for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbf{E}\Big(|X_n - X|^r\Big) = 0.$$

where $\Big[\mathbf{E}\Big(|X_n - X|^r\Big)\Big]^{1/r}$ is the L^r distance between X_n and X

- The target X has the same sample space with all the X_i's
- When $r = 2$, converge in L^2 is also called converge in quadratic mean, i.e., $X_n \xrightarrow{qm} X$. The convergence in quadratic mean is generally used.
- To show L^r convergence, just figure out an upper bound of $\mathbf{E}\Big(|X_n - X|^r\Big)$, and show this upper bound goes to 0.

Example: Convergence in L_2

- Recall the pervious example when X has a point mass at 1, and $X_n \sim N\left(1, \frac{1}{n^2}\right)$. According to the property of normal distribution.,$X_n - X \sim N\left(0, \frac{1}{n^2}\right)$, so

$$\mathbf{E}\left(|X_n - X|^2\right) = \left(\mathbf{E}(X_n - X)\right)^2 + \mathbf{Var}(X_n - X)$$
$$= 0 + \frac{1}{n^2} = \frac{1}{n^2} \to 0.$$

Hence, $X_n \xrightarrow{L^2} X$

Properties: Convergence in L_2

- According to the deviation, if $Var(X_n - X) \to 0$, and $E(X_n - X) \to 0$, then there is

$$E\left(|X_n - X|^2\right) = \left(E(X_n - X)\right)^2 + Var(X_n - X) \to 0$$

Property 1

if $Var(X_n - X) \to 0$, and $E(X_n - X) \to 0$, then $X_n \xrightarrow{L^2} X$.

Properties: Convergence in L_2

Property 2

Let $0 < s < r < \infty$ if $X_n \xrightarrow{L_r} X$, then $X_n \xrightarrow{L_s} X$.

- Recall that with Holder inequality, there is

$$E(|YZ|) \leq E(|Y|^k)^{1/k} E(|Z|^l)^{1/l}$$

for $1 < k, l < \infty$ with $\frac{1}{k} + \frac{1}{l} = 1$.

- Let $Y = 1$, $Z = |X_n - X|^r, l = r/s$, and $k = 1/(1 - s/r) > 1$. Then

$$\begin{aligned} E\Big(|X_n - X|^s\Big) &= E\Big(|X_n - X|^s \times 1\Big) \\ &\leq \Big[E\Big(|X_n - X|^r\Big)\Big]^{s/r} \times 1^{1/k} \\ &= \Big[E\Big(|X_n - X|^r\Big)\Big]^{s/r} \to 0 \end{aligned}$$

Properties: Convergence in L_2

Property 3

Let $0 < r < \infty$ if $X_n \xrightarrow{L^r} X$, then $X_n \xrightarrow{p} X$.

Proof:

$$P\Big(|X_n - X| \geq \varepsilon\Big) = P\Big(|X_n - X|^r \geq \varepsilon^r\Big)$$
$$\leq \frac{E\Big(|X_n - X|^r\Big)}{\varepsilon^r} \to 0$$

Markov's Inequality : non-negative r.v.

$$P(x \geq a) \leq \frac{E(X)}{a}$$

Markov's Inequality

Markov's(Chebyshev's) Inequality

- If g is strictly increasing and positive on $(0, \infty)$, $g(x) = g(-x)$.
- X is a r.v. such that $E[g(X)] < \infty$, then for each $a > 0$

$$P(|X| \geq a) \leq \frac{E[g(X)]}{g(a)}$$

Proof:

$$\begin{aligned} E[g(X)] &\geq E[g(X) I_{\{g(X) \geq g(a)\}}] \\ &\geq g(a) E[I_{\{g(X) \geq g(a)\}}] \\ &= g(a) E[I_{\{|X| \geq a\}}] \\ &= g(a) P(|X| \geq a) \end{aligned}$$

Some special cases: Markov's Inequality

$$g(x) = |x| \Longrightarrow P(|X| \geq a) \leq \frac{E|X|}{a}$$

$$g(x) = x^p \Longrightarrow P(|X| \geq a) \leq \frac{E|g(X^p)|}{a^p}$$

$$g(x) = x^2 \Longrightarrow P(|X - EX| \geq a) \leq \frac{Var(X)}{a^2}$$

$$g(x) = e^{t|x|} \Longrightarrow P(|X| \geq a) \leq \frac{E\left[e^{t|X|}\right]}{e^{ta}}$$

for some constant $t \geq 0$

Almost Sure Convergence

Definition 5.5.6

For a sequence of r.v.'s $\{X_n\}_{i=1}^{\infty} = X_1, X_2, \ldots, X_n, \ldots$, we say they almost sure convergence to r.v. X (i.e. $X_n \overset{a.s.}{\to} X$) if for any $\varepsilon > 0$,

$$P\Big(\lim_{n\to\infty} X_n(\omega) = X(\omega)\Big) = 1 \text{ or } P\Big(\lim_{n\to\infty} |X_n(\omega) - X(\omega)| < \varepsilon\Big) = 1$$

- The target X has the same sample space with all the X_i's.
- $\{X_n\}$ and X are usually dependent
- Practically, to show the a.s. convergence,
  - For each outcome ω, find the sequence $X_1(\omega), X_2(\omega), \ldots$ (sequence of real numbers) and the real number $X(\omega)$. Figure out whether $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ is true or not.
  - Let the event $A = \{\omega, \lim_{n\to\infty} X_n(\omega) = X(\omega)\}$.
  - Check if $P(A) = 1$
- Interpretation: for almost all the outcomes ω !, when n is large enough, $|X_n(\omega) - X(\omega)| \leq \varepsilon$ for any $\varepsilon > 0$.

Example 1: Almost Sure Convergence

- Let the sample space $\Omega = [0,1]$, with a probability measure that is uniform on this space, i.e. $P([a,b]) = b - a$ for any $0 \leq a \leq b \leq 1$.
- Let

$$X_n(\omega) = \begin{cases} 1, & 0 \leq \omega < \dfrac{n+1}{2n} \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad X(\omega) = \begin{cases} 1, & 0 \leq \omega < \dfrac{1}{2} \\ 0, & \text{otherwise} \end{cases}$$

For each $\omega \in [0,1]$.

- If $\omega \in [0, 1/2)$, then $X_n(\omega) = 1 = X(\omega)$.
- If $\omega = 1/2$, then $X_n(\omega) = 1 \nrightarrow X(\omega) = 0$.
- If $\omega \in (1/2, 1]$, then $X_n(\omega) = 0 = X(\omega)$, when $\frac{n+1}{2n} < \omega$, which is equivalent with $n \geq \frac{1}{2\omega - 1}$.

So, $A = [0, 1/2) \cup (1/2, 1]$. Check $P(A) = 1$?

Example 5.5.7: Almost Sure Convergence

- Let the sample space $\Omega = [0,1]$, with a probability measure that is uniform on this space, i.e. $P([a,b]) = b - a$ for any $0 \le a \le b \le 1$.
- Define r.v.

$$X_n(\omega) = \omega + \omega^n \text{ and } X(\omega) = \omega$$

For each $\omega \in [0,1]$.

- If $\omega \in [0,1)$, $\omega^n \to 0$, then $X_n(\omega) \to \omega = X(\omega)$.
- If $\omega = 1$, then $X_n(\omega) = 2 \nrightarrow X(\omega) = 1$ for every n

So, $A = [0,1)$. Check $P(A) = 1$?

Almost Sure Convergence

- Comparison between almost sure convergence and converge in probability
  - Convergence in probability: for each n, consider $P(|X_n(\omega) - X(\omega)| > \varepsilon)$, and check the limit of this probability
  - Almost sure convergence: for each ω, check the limit $\lim_{n\to\infty} X_n(\omega)$, and find the probability of the set that the limit does not equal to $X(\omega)$

Almost Sure Convergence

- Can we express it as the limit of probability?

Theorem: Almost Sure Convergence

The following statements are equivalent:

1. $X_n \overset{a.s.}{\to} X$
2. $\forall \varepsilon > 0, P\Big(\bigcap_{k\geq n}\{|X_k - X| < \varepsilon\}\Big) \to 1$
3. $\forall \varepsilon > 0, P\Big(\bigcup_{k\geq n}\{|X_k - X| \geq \varepsilon\}\Big) \to 0$
4. $\forall \varepsilon > 0,$
$$\lim_{n\to\infty} P\Big(\sup_{k\geq n} |X_k - X| > \varepsilon\Big) = 0$$

Here, we consider the set $\cup_{k\geq n}\{|X_k - X| > \varepsilon\}$

Property 1: Almost Sure Convergence

Property 1

If $X_n \stackrel{a.s.}{\rightarrow} X$, then $X_n \stackrel{p}{\rightarrow} X$.

Proof: for for any $\varepsilon > 0$,

$$\begin{aligned} 0 \geq & \lim_{n\to\infty} P\Big(|X_n - X| \geq \varepsilon\Big) \\ & \lim_{n\to\infty} P\Big(\sup_{k\geq n} |X_k - X| \geq \varepsilon\Big) \\ & = 0 \end{aligned}$$

Hence, $\lim_{n\to\infty} P\Big(|X_n - X| \geq \varepsilon\Big) = 0$, which implies $X_n \stackrel{p}{\rightarrow} X$.

Convergence in Distribution

Definition 5.5.9

Let $\{X_n\}_{i=1}^{\infty} = X_1, X_2, \ldots, X_n, \ldots$ be a sequence of r.v.'s with CDF $F_1, \ldots, F_n, \ldots$, and X be r.v. with CDF F. we say they converges in distribution to r.v. X (i.e. $X_n \xrightarrow{d} X$) if

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

at every point at which F is continuous.

- $\{X_n\}$ and X can be dependent or independent
- Convergence:
  - If X is discrete, the convergence stands at points F does not jump
  - If X is cont., the convergence stands at every point
- Convergence in distribution is really the CDFs that converge, not the r.v. Hence it quite different from conv. in prob. or alm. sure conv.

Property 1: Convergence in Distribution

Property 1

If $X_n \xrightarrow{p} X$, then $X_n \xrightarrow{d} X$.

Proof: Denote $F_n(x) = P(X_n \leq x)$ and $F(X) = P(X \leq x)$. First we have

$$\begin{aligned} F_n(x) &= P(X_n \leq x) \\ &= P(X_n \leq x, |X_n - X| \leq \varepsilon) + P(X_n \leq x, |X_n - X| > \varepsilon) \\ &\leq P(X \leq x - (X_n - X), |X_n - X| \leq \varepsilon) + P(|X_n - X| > \varepsilon) \\ &\leq P(X \leq x + \varepsilon) + P(|X_n - X| > \varepsilon) \\ &= F(x + \varepsilon) + P(|X_n - X| > \varepsilon) \end{aligned}$$

On the other hand,

$$\begin{aligned} F_n(x) &= 1 - P(X_n \geq x) \\ &= 1 - P(X_n \geq x, |X_n - X| \geq \varepsilon) - P(X_n \geq x, |X_n - X| \leq \varepsilon) \\ &\geq 1 - P(X \geq x - (X_n - X), |X_n - X| \leq \varepsilon) - P(|X_n - X| \leq \varepsilon) \\ &\geq 1 - P(X \leq x - \varepsilon) - P(|X_n - X| \leq \varepsilon) \\ &= F(x - \varepsilon) - P(|X_n - X| \leq \varepsilon) \end{aligned}$$

Combining the two, we have

$$F(x-\varepsilon) - P(|X_n - X| \leq \varepsilon) \leq F_n(x) \leq F(x+\varepsilon) + P(|X_n - X| \leq \varepsilon)$$

Letting $n \to \infty$ and since $X_n \xrightarrow{p} X$,

$$F(x-\varepsilon) \leq \lim \inf_{n\to\infty} \leq F_n(x) \leq \lim \sup_{n\to\infty} \leq F(x+\varepsilon)$$

Recall that F is continuous at x, which means $F(x-\varepsilon) \to F(x)$ and $F(x+\varepsilon) \to F(x)$ as $\varepsilon \to 0$. Hence,

$$F(x) \leq \lim \inf_{n\to\infty} \leq F_n(x) \leq \lim \sup_{n\to\infty} \leq F(x)$$

Theorem: Convergence in Distribution

Recall the characteristic function for $X \sim F$ is $\phi_X(t) = E(e^{\imath t})$. If $\phi_X(t) = \phi_Y(t)$ then X and Y have the same distribution.

Theorem: Convergence in Distribution

Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of r.v.'s with characteristic functions $\phi_{X_n}(t)$ and X be a r.v. with the characteristic function $\phi_X(t)$. Then,

$$X_n \xrightarrow{d} X \Longleftrightarrow \lim_{n\to\infty} \phi_{X_n}(X) = \phi_X(t)$$

Example: Suppose that $X_n \sim N(\mu + 1/n, \sigma^2 + 1/n)$, then

$$\phi_{X_n}(t) = \exp\{(\mu + 1/n^2)\imath t - t^2(\sigma^2 + 1/n)/2\} \to \exp\{\mu\imath t - t^2\sigma^2/2\}$$

Note that the limit is the characteristic function for $X \sim N(\mu, \sigma^2)$. So, $X_n \xrightarrow{d} X$. It is easier than the analysis on the CDF of X_n.

Relationship Between 4 Types of Convergence

Theorem

1. $$\left.\begin{array}{l} X_n \xrightarrow{a.s.} X \\ X_n \xrightarrow{L_r} X \end{array}\right\} \Rightarrow X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{d} X$$
2. If $0 < s < r < \infty$, $X_n \xrightarrow{L_r} X \Rightarrow X_n \xrightarrow{L_s} X$.
3. No other implications hold in general.

- (1). (a) If $X_n \overset{a.s.}{\to} X$, then $X_n \overset{p}{\to} X$. The converse may not hold. Let

$$P(X_n = 0) = 1 - \frac{1}{n}, \;\; P(X_n = 1) = \frac{1}{n}$$

and X_n's are independent. Since $P(|X_n - 0| > \varepsilon) = P(X_n = 1) = n^{-1} \to 0$, Then $X_n \overset{p}{\to} 0$. However, $X_n \overset{a.s.}{\nrightarrow} 0$ since for any $0 < \varepsilon < 1$, we have

$$\lim_{n\to\infty} P\Big(\bigcap_{k\geq n} \{|X_k - 0| < \varepsilon\}\Big) = \lim_{n\to\infty} P\Big(\lim_{r\to\infty} \bigcap_{k\geq n}^{r} \{|X_k| < \varepsilon\}\Big)$$

$$= \lim_{n\to\infty} \lim_{r\to\infty} P\Big(\bigcap_{k\geq n}^{r} \{|X_k| < \varepsilon\}\Big) = \lim_{n\to\infty} \lim_{r\to\infty} \prod_{k=n}^{r} \Big(1 - \frac{1}{k}\Big)$$

$$= \lim_{n\to\infty} \lim_{r\to\infty} \frac{n-1}{n} \frac{n}{n+1} \cdots \frac{r-1}{r} = \lim_{n\to\infty} \lim_{r\to\infty} \frac{n-1}{r} = 0$$

- (b) If $X_n \xrightarrow{L_r} X$, then $X_n \xrightarrow{p} X$. The converse may not hold.

$$P(X_n = 0) = 1 - \frac{1}{n}, \ \ P(X_n = n) = \frac{1}{n}$$

Then $X_n \xrightarrow{p} 0$ since

$$P(|X_n - 0| > \varepsilon) = P(X_n = n) = \frac{1}{n} \to 0.$$

But $EX_n = 1 \nrightarrow 0$.

- (b) If $X_n \xrightarrow{p} X$, then $X_n \xrightarrow{d} X$. The converse may not hold.

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

Then $X_n \xrightarrow{d} X$, but $X_n \overset{p}{\nrightarrow} X$ since

$$P(|X_n - X| > \varepsilon) = P(2|X| > \varepsilon) \nrightarrow 0.$$

- (2) If $0 < s < r < \infty$, $X_n \xrightarrow{L_r} X \Rightarrow X_n \xrightarrow{L_s} X$. The converse may not hold.

$$P(X_n = 0) = 1 - \frac{1}{n^2}, \;\; P(X_n = n) = \frac{1}{n^2}$$

Then $X_n \xrightarrow{L_1} X$ since

$$E|X_n - 0| = \frac{1}{n^2} \times n = \frac{1}{n} \to 0$$

. But $X_n \xrightarrow{L_2} X$ since

$$E|X_n - 0|^2 = \frac{1}{n^2} \times n^2 = 1 \nrightarrow 0$$

.

- (3). We now show that "a.s. convergence" and "mean convergence" do not imply each other.
  - Let $P(X_n = 0) = 1 - n^{-2}$ and $P(X_n = n^3) = n^{-2}$. Then $X_n \overset{a.s.}{\to} 0$, but $X_n \overset{L_1}{\nrightarrow} 0$. Since

$$\lim_{n\to\infty} P\Big(\bigcup_{k\geq n}\{|X_k - 0| \geq \varepsilon\}\Big) = \lim_{n\to\infty} P\Big(\lim_{r\to\infty}\bigcup_{k\geq n}^{r}\{|X_k| \geq \varepsilon\}\Big)$$

$$= \lim_{n\to\infty} \lim_{r\to\infty} P\Big(\bigcup_{k\geq n}^{r}\{|X_k| \geq \varepsilon\}\Big) = \lim_{n\to\infty}\lim_{r\to\infty}\sum_{k=n}^{r}\frac{1}{k^2}$$

$$= \lim_{n\to\infty}\lim_{r\to\infty}\frac{1}{n^2} + \frac{1}{(n+1)^2} + \cdots + \frac{1}{r^2} \to 0.$$

However,

$$E|X_n - 0| = \frac{1}{n^2} \times n^3 \to \infty$$

  - $X_n \overset{L_1}{\to} 0$, but $X_n \overset{a.s.}{\nrightarrow} 0$

$$P(X_n = 0) = 1 - \frac{1}{n}, \;\; P(X_n = 1) = \frac{1}{n}$$

Properties of Convergence

- $X_n \to X$ and $Y_n \to Y$, then $X_n \pm Y_n \to X + Y$
  - $X_n \xrightarrow{a.s.} X, Y_n \xrightarrow{a.s.} Y$, then $X_n + Y_n \xrightarrow{a.s.} X + Y$,
  - $X_n \xrightarrow{L_r} X, Y_n \xrightarrow{L_r} Y$, then $X_n + Y_n \xrightarrow{L_r} X + Y$,
  - $X_n \xrightarrow{p} X, Y_n \xrightarrow{p} Y$, then $X_n + Y_n \xrightarrow{p} X + Y$,
  - $X_n \xrightarrow{d} X, Y_n \xrightarrow{d} Y$, it is not sure that $X_n + Y_n \xrightarrow{d} X + Y$
- Slutsky's Theorem Let $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} C$, then
  1. $X_n + Y_n \xrightarrow{d} X + C$
  2. $X_n Y_n \xrightarrow{d} CX$
  3. $X_n / Y_n \xrightarrow{d} X/C$ id $C \neq 0$
- The Continuous Mapping Theorem: if $g(\cdot)$ is a continuous function, then
  - $X_n \xrightarrow{a.s.} X$, then $g(X_n) \xrightarrow{a.s.} g(X)$,
  - $X_n \xrightarrow{p} X$, then $g(X_n) \xrightarrow{p} g(X)$,
  - $X_n \xrightarrow{d} X$, then $g(X_n) \xrightarrow{d} g(X)$,

Stochastic Orders

Recall:

- In mathematics, we use o and O notations to denote the order of terms
- $a_n = o(1)$ means $a_n \to 0$ when $n \to \infty$; $a_n = o(b_n)$ means that $a_n/b_n = o(1)$.
- $a_n = O(1)$ means $|a_n| \le C$ for some constant $C > 0$, for all large n; $a_n = O(b_n)$ mean $a_n/b_n = O(1)$.

Now we consider the probabilistic version:

Definition o_p

If $X_n \xrightarrow{p} 0$, i.e. $P(|X_n| \ge \varepsilon) \to 0$ for every $\varepsilon > 0$, then we say that $X_n = o_p(1)$

Definition O_p

We say that $X_n = O_p(1)$, or X_n is bounded in probability, if for any $\varepsilon > 0$, there exists $C_\varepsilon > 0$, such that

$$P(|X_n| > C_\varepsilon) \le \varepsilon.$$

Stochastic Orders

Generalisation: Consider a sequence $X_1, X_2 \dots$ of of r.v.'s and $a_1, a_2, \dots,$ a sequence of positive real numbers,

- For a r.v. X, $X_n \xrightarrow{p} X$ if only if $X_n - X = o_p(1)$
- $X_n = o_p(a_n)$ if only if $a_n^{-1} X_n = o_p(1)$. a_n is the rate.
- $X_n = O_p(a_n)$ if only if $a_n^{-1} X_n = O_p(1)$. a_n is the rate.

Examples:

- If $X_n \sim N(0, 1/n)$, then $X_n = o_p(1)$ and $X_n = O_p(1/\sqrt{n})$
- If $X_n = o_p(1)$, then $X_n = O_p(1)$

Properties:

- $O_p(1) o_p(1) = o_p(1), O_p(1) O_p(1) = O_p(1)$
- $O_p(1) + o_p(1) = O_p(1)$
- $O_p(a_n) o_p(b_n) = o_p(a_n b_n), O_p(a_n) O_p(b_n) = O_p(a_n b_n)$
- $\left(1 + o_p\right)^{-1} = O_p(1)$
- $o_p(O_p(1)) = o_p(1)$

Thank you