2023년 5월 16일 화요일 13:21

CM23-230516Tue

Announcements

1. Quiz #2 was taken. However, one student hasn't yet taken it. No discussion on the quiz, please.
2. Handouts, annotated handouts are available at http://cisl.postech.ac.kr/class/eece574/schedule.htm . Passwords are ...
3. Lecture videos are available at cisl.postech YouTube channel. Clickable time stamps, please.
4. HWs will be assigned. Check the PLMS.

Corrections; Clarifications; Common mistakes,misunderstandings, misconceptions, myth; Frequently Asked Questions; Conventional wisdom

1. Convergence in measure requires the convergences of uncoutably infinite number of sequences parameterized by epsilon. $\forall \varepsilon > 0,\ m(\{s \in \Omega : |X_n(s) - X(s)| > \varepsilon\}) \to 0$ as $n \to \infty$

Overview w/ Chapter numbers from the main textbook by Papoulis, (2)-(7) from Peebles

Pa 9.1 Definitions
Pe 6 Random Process--Temporal Characteristics
Pa 7.4 Stochastic Convergence and Limit Theorems
Pa 9.2 Systems with Stochastic Inputs
Pa 9.3 The Power Spectrum
Pe 7 Random Process--Spectral Characteristics
Pe 8 Linear Systems with Random Inputs
Pe 9 Optimal Linear Systems

$(X_n(s_0))_{n=1}^{\infty} \to X(s_0)$

$X : \Omega_1 \to R_1 \quad Y : \Omega_2 \to R_2$

$\mathcal{L}(x_1(t)) = \frac{1}{s+a} \quad \mathrm{Re}\{s+a\} < 0$

$\mathcal{L}(x_2(t)) = \frac{1}{s+a} \quad \mathrm{Re}\{s+a\} > 0$

(i) $\Omega_1 = \Omega_2 = \Omega$

(ii) $R_1, R_2 \subset R$

(iii) X & Y are real-valued fts

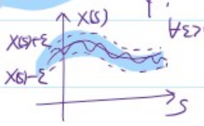
$\Rightarrow R_1, R_2 \subset \mathbb{R}$

$\Rightarrow R_1, R_2 \subset \mathbb{C}$

Review

1. two real- or complex-valued functions
   a. Equality
   b. Approximate Equality
      - Construct a sequence and show it converges.
2. 5 modes of convergence of a sequence of real functions
   a. uniform
   b. everywhere
   c. almost everythere
   d. in the pth mean
   e. in measure

Def. $(X_n(s))_{n=1}^{\infty}$ X. Y. $\forall \varepsilon > 0$



Def. $X = Y$ everywhere on Ω

if $X(s) = Y(s),\ \forall s \in \Omega$.

if $\{s \in \Omega : X(s) = Y(s)\} = \Omega$.

if $\{s \in \Omega : X(s) \neq Y(s)\} = \phi$

if $\forall \varepsilon > 0,\ \{s \in \Omega : |X(s) - Y(s)| \le \varepsilon\} = \Omega$

if $\forall \varepsilon > 0,\ \{s \in \Omega : |X(s) - Y(s)| > \varepsilon\} = \phi$

Def. $X = Y$ almost everywhere on Ω.

if $m(\{s \in \Omega : X(s) \neq Y(s)\}) = 0 \leftarrow$

~~(if $m(\{s \in \Omega : X(s) = Y(s)\}) = m(\Omega)$)~~

if $\int_{\Omega} |X(s) - Y(s)|^p ds = 0$. $\quad p \ge 1$

if $\forall \varepsilon > 0,\ m(\{s \in \Omega : |X(s) - Y(s)| > \varepsilon\}) = 0$

Preview

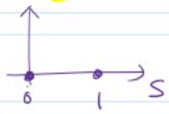
1. two random variables
   a. Equality
   b. Approximate Equality
      - Construct a sequence and show it converges.
2. 4 modes of convergence
   a. almost sure
   b. in probability (measure)
   c. in the pth mean
   d. in distribution
      i. CDF
      ii. characteristic function
3. LLN
   a. strong
   b. weak
4. CLT

(e) Def. $X_n \to X$ everywhere

if $\{s \in \Omega : X_n(s) \not\to X(s)\} = \phi$.

(a.e.) Def. $X_n \to X$ almost everywhere on Ω.

if $m(\{s \in \Omega : X_n(s) \not\to X(s)\}) = 0$



$\Omega = [0, 1]$ $\quad X_n \to X$ for $s \in (0, 1)$

$X_n \not\to X$ for $s \in \{0, 1\}$

(pth mean) Def. $X_n \to X$ in the p-th mean $(p \ge 1)$

if $\int_{\Omega} |X_n(s) - X(s)|^p ds \to 0$ as $n \to \infty$.

(in measure)

Def. $X_n \to X$ in measure

if $\forall \varepsilon > 0,\ m(\{s \in \Omega : |X_n(s) - X(s)| > \varepsilon\}) \to 0$

$y_n(\varepsilon)$

uniform convergence

$\Downarrow$

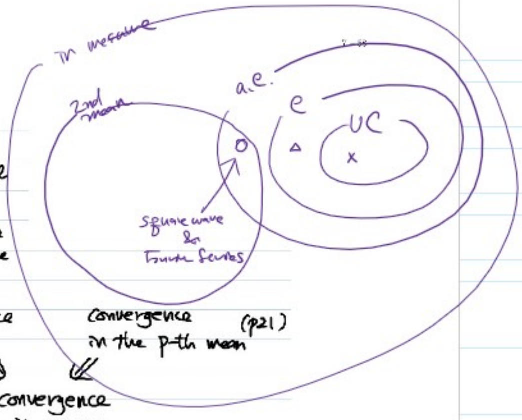
convergence everywhere

$\Downarrow$

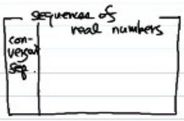
convergence a.e. $\qquad$ convergence in the p-th mean $(p \geq 1)$

$\searrow \qquad \swarrow$

convergence in measure



a sequence of real numbers $(x_n)_{n \geq 1}$ $\quad x_n \in \mathbb{R}$



limit

$\limsup_{n \to \infty} x_n$

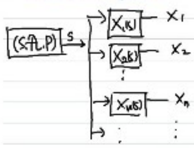
$\liminf_{n \to \infty} x_n$

a sequence of real functions

a sequence of real r. variables

→ Convergence of a sequence of random variables

Ch.7 random vector / random sequence



$(X_n)_{n \geq 1}$

$(X_n)_n$

(X_n)

X_n

$\{X_n\}$

$\{X_n\}_n$

$\{X_n\}_{n \geq 1}$

Q. a sequence of real r. variables converges to ?

Equality vs. Approximate Equality of Two Random Variables

- Apparently,
  - (i) the equality of two random variables is nothing but the equality of two functions because a random variable is a function and
  - (ii) the convergence of a sequence of random variables is nothing but the convergence of a sequence of functions.

Def. Two real RVs equal almost surely.

X & Y are equal w.p. 1

if $P(\{s \in S : X(s) = Y(s)\}) = 1$

$\equiv$ if $P(\{s \in S : X(s) \neq Y(s)\}) = 0$.

$\equiv$ if $\int_{s \in S} |X(s) - Y(s)|^p \, dP(s) = 0$ for $p \geq 1$.

$(\equiv E\{|X - Y|^p\} = 0)$

$\equiv$ if $\forall \varepsilon > 0$, $P(\{s \in S : |X(s) - Y(s)| > \varepsilon\}) = 0$.

- However,
  - the Borel measure needs to be changed to the probability measure of each probability space, and $P(S) = 1 < \infty$,

  (0) equal everywhere is not used, the uniform convergence is not used,

  (1) almost everywhere equality/convergence is called almost sure equality/convergence,

  (2) the almost everywhere equality/convergence is called the equality/convergence with probability one

  (3) the convergence in measure is called the convergence in probability, and

  (4) there is a new mode of convergence called convergence in distribution.
- In summary, almost surely = with probability 1 is used for equality, and there are 4 modes of convergence.

- Def. (almost sure equality) Given a probability space $(S, \mathcal{A}, P)$, two random variables X and Y are equal almost surely (a.s.) if

$$P(\{s \in S : X(s) = Y(s)\}) = 1,$$

  - ... iff $P(\{s \in S : X(s) \neq Y(s)\}) = 0$, i.e., $X \neq Y$ only on a set of probability zero.
  - $X = Y$ a.s. is alternatively called $X = Y$ with probability one (w.p. 1).
- Lemma. Two random variables X and Y are equal almost surely (a.s.) iff

$$\mathbf{E}\{|X - Y|^p\} = 0$$

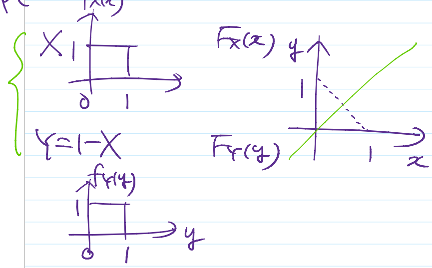
for some $p \geq 1$.

- Lemma. Given a probability space $(S, \mathcal{A}, P)$, two random variables X and Y are equal almost surely (a.s.) iff

$$\forall \varepsilon > 0, P(\{s \in S : |X(s) - Y(s)| > \varepsilon\}) = 0.$$

  - $\forall \varepsilon > 0, P(|X - Y| > \varepsilon) = 0$
  - $\forall \varepsilon > 0, P(|X - Y| \leq \varepsilon) = 1$
- Lemma. $X = Y$ a.s. $\Longrightarrow F_X(x) = F_Y(x)$. However, the converse is not true in general.

$P(X=Y)=0.$



□ $X=Y$ vs $\{X=Y\}$

(i) $X(\omega) = Y(\omega), \forall \omega \in \Omega$ equal surely

(ii) $P(\{\omega \in \Omega : X(\omega) \neq Y(\omega)\}) = 0$, " almost surely

(iii) $E[|X-Y|^p] = 0$. $\left(\Leftrightarrow \int_\Omega |X(\omega)-Y(\omega)|^p \, dP(\omega) = 0\right)$
$p \geq 1$.

(iv) $\forall \varepsilon > 0 \quad P(|X-Y| > \varepsilon) = 0$

Convergence of a Random Sequence

- Four modes of the convergence of a sequence of random variables
  - Given a probability space $(S, \mathcal{A}, P)$, a sequence $(X_n)_{n\geq 1}$ of random variables $X_1(s), X_2(s), \ldots$, and the limit random variable $X(s)$, we have the following four modes of convergence.
    1. Def. $X_n \to X$ almost surely (a.s.) w.p.1
    $$P(\{s : \lim_{n\to\infty} X_n(s) = X(s)\}) = 1 \iff P(\{s \in S : X_n(s) \not\to X(s)\}) = 0$$
    2. Def. $X_n \to X$ in the pth mean ($p \geq 1$)
    $$\lim_{n\to\infty} \mathbf{E}[|X_n - X|^p] = 0$$
    If $p=2$, then $\boxed{\underset{n\to\infty}{\mathrm{l.i.m.}}\, X_n = X}$
    $E\{|X_n - X|^2\} \to 0$ as $n \to \infty$
    3. Def. $X_n \to X$ in probability
    $$\forall \varepsilon > 0, \lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$$
    4. Def. $X_n \to X$ in distribution/law (weak convergence)
    $$\lim_{n\to\infty} F_{X_n}(x) = F_X(x), \text{ for all continuous points}$$
  - (convergence a.s.) $X_n \to X$ w.p. 1
  - (convergence in the pth mean) $X_n \to X$ in L^p, $X_n \xrightarrow{L^p} X$
  - (convergence in probability) $X_n \xrightarrow{\mathcal{P}} X$; convergence in measure, convergence in probability measure

$\lim_{n\to\infty} \underbrace{\Phi_{X_n}(\omega)}_{\triangleq E\{e^{j\omega X_n}\}} = \Phi_X(\omega) \triangleq E\{e^{j\omega X}\}, \ \forall \omega.$

  - (convergence in distribution) $X_n \xrightarrow{\mathcal{D}} X$
- Properties
  - a.s. convergence is often proved by using the Borel-Cantelli lemma.
  - a.s. $\Rightarrow$ in probability
  - $L^p \Rightarrow$ in probability
  - in probability $\Rightarrow$ in distribution
  - in distribution $\Leftrightarrow \lim_{n\to\infty} \Phi_{X_n}(\omega) = \Phi_X(\omega)$

$X_n \xrightarrow{d} X \iff E[f(X_n)] \to E[f(X)]$ as $n \to \infty$
for every bounded continuous ft.

a.s. $\Rightarrow$ in prob. $\Leftarrow$ in L^p sense / in the pth mean ($p \geq 1$)

$\Downarrow$

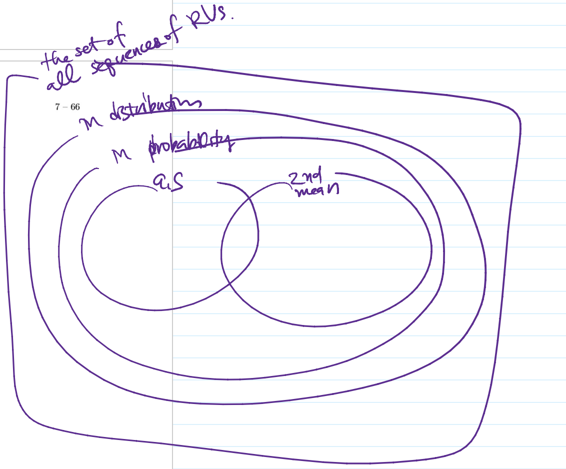
in distribution

Asymptotic 점근적 analysis

4 modes of convergence

$\boxed{\text{rate of convergence.}}$

big O, little o notations



- Proofs
  - a.s. $\Rightarrow$ in probability

" $\forall \varepsilon > 0,\ P(|X_n - X| > \varepsilon) \to 0$ "

$0 \le P(|X_n - X| > \varepsilon) \le P\left(\bigcup_{m \ge n} \{|X_m - X| > \varepsilon\}\right)$

Recall $\lim_{n\to\infty} P\left(\bigcup_{m\ge n}\{\ \}\right) = P\left(\bigcap_{n=1}^{\infty}\bigcup_{m\ge n}\{\ \}\right)$

$\limsup_{n\to\infty} C_n = \bigcap_{n=1}^{\infty}\bigcup_{m\ge n} C_m$

$\liminf_{n\to\infty} C_n = \bigcup_{n=1}^{\infty}\bigcap_{m\ge n} C_m$

$= P(\{\omega \in \Omega : X_n(\omega) \nrightarrow X(\omega)\}) = 0 \quad \because X_n \to X$ a.s.

where we used the Borel-Cantelli lemma.

  - in the 2nd mean $\Rightarrow$ in probability

$P(|X_n - X| > \varepsilon) = P(|X_n - X|^2 > \varepsilon^2) \le \frac{E\{|X_n - X|^2\}}{\varepsilon^2} \to 0$

  - in probability $\Rightarrow$ in distribution

$F_{X_n}(x) = P(X_n \le x)$

(i) upper bound

$P(X_n \le x) = P(\cancel{X_n \le x}, X \le x+\varepsilon) + Pr(X_n \le x, X > x+\varepsilon)$

$\le P(X \le x+\varepsilon) + Pr(|X_n - X| > \varepsilon)$

$= F_X(x+\varepsilon) + P(|X_n - X| > \varepsilon)$

$X > x+\varepsilon$
$+)\ -X_n \ge -x$
$X - X_n > \varepsilon$

(ii) lower bound

$F_X(x-\varepsilon) = Pr(X \le x-\varepsilon)$

$= Pr(X \le x-\varepsilon, X_n \le x) + Pr(X \le x-\varepsilon, X_n > x)$

$\le Pr(X_n \le x) + Pr(|X_n - X| > \varepsilon)$

$F_X(x-\varepsilon) \le F_{X_n}(x) + P(|X_n - X| > \varepsilon)$

$-P(|X_n - X| > \varepsilon) + F_X(x-\varepsilon) \le F_{X_n}(x)$

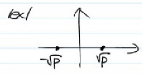
$n \to \infty$
$\varepsilon \downarrow 0$

$F_X(x) \le \lim_{n\to\infty} F_{X_n}(x) \le F_X(x)$

- Which mode of convergence is simplest to handle?

- Convergence in distribution and Probability of error

Ex)



$Y_n = \sqrt{P}\, b + N + \cancel{\sqrt{P_I} N_I}$

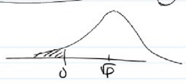
$b \in \{1, -1\}$

$N \sim N(0, \sigma_N^2)$

X_i i.i.d. $B(1, \frac{1}{2})$

$N_I^{(n)} = \frac{(2X_1 - 1) + (2X_2 - 1) + \cdots + (2X_n - 1)}{n} = \frac{2\sum X_i - n}{n}$

$P_b = Pr(Y_n < 0 \mid b = 1)$



$= E\left[1_{(-\infty, 0]}(Y) \mid b = 1\right]$

$\cong Pr(Y < 0 \mid b = 1)$ (X)

$\simeq Pr(\sqrt{P}\, b + N + \sqrt{P_I} N_I^{(n)} \mid b = 1)$

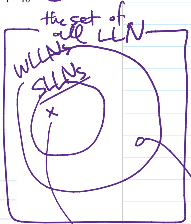
$N(0,1)$

$\frac{\sqrt{n} N_I^{(n)}}{\sqrt{n}} = \frac{1}{\sqrt{n}} \frac{(2X_1 - 1) + (2X_2 - 1) + \cdots + (2X_n - 1)}{\sqrt{n}}$

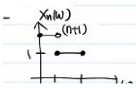
$= Q\left(\sqrt{\frac{P}{\sigma_N^2 + \frac{P_I}{n}}}\right)$

the LLN (X)
the Laws of Large numbers

- Examples
  - The strong law of large numbers (SLLN) is an almost sure convergence.
  - The weak law of large numbers (WLLN) is a convergence in measure.
  - The central limit theorem (CLT) is a convergence in distribution.



- Counter-Examples
  - a.s. $\nRightarrow$ in the 2nd mean



(i) $\{\omega \in \Omega : X_n(\omega) \to X(\omega)\} = (0, 1]$

$P((0,1]) = 1 \quad \therefore X_n \to X$ w.p. 1

$P > 2$

$(\sqrt{n})^P \frac{1}{n} = n^{\frac{P}{2} - 1}$

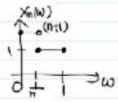
$\frac{P}{2} - 1 < 0$

$P < 2$

$(\sqrt{n})^2 \frac{1}{n} = \frac{n}{n} = 1$

a.s. $\not\Rightarrow$ in the 2nd mean

–



(i) $\{\omega\in\Omega : X_n(\omega)\to X(\omega)\} = (0,1]$

$P((0,1]) = 1$ $\quad\therefore\ X_n \to X$ w.p. 1

(ii) $E\{|X_n - X|^2\} = \int_0^1 |X_n(\omega) - X(\omega)|^2 d\omega$

$= \int_0^{\frac{1}{n}} (n)^2 \, d\omega = n \not\to 0$ as $n\to\infty$.

(iii) $P(\{s\in S : |X_n(s) - X(s)| > \varepsilon\})$

$= P(\{s\in S : 0 \le s \le \frac{1}{n}\})$

$= \frac{1}{n} \to 0,\ \forall\varepsilon > 0.$

in probability

a.s. convergence

$(\sqrt{n})^2 \frac{1}{n} = \frac{n}{n} = 1$

$\frac{p}{2} - 1 < 0$

$p < 2$

Def. Given a sequence X_n of real RVs

$E\{X_n\} = \mu$

Conditions on X_n's

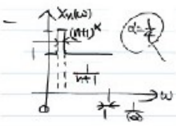
$M_n \triangleq \frac{X_1 + \cdots + X_n}{n}$

$M_n \to \mu$ in some sense

Then we have a LLN.

– a.s. $\not\Rightarrow$ in the pth mean ($p \ge 4$)
in the pth mean ($1 \le p < 4$) $\not\Rightarrow$ a.s.

–



$E\{|X_n - X|^p\} = \frac{1}{n+1}\left((n+1)^{\frac{1}{4}}\right)^p$

$\to 0$ as $n\to\infty$ $\quad 1 \le p < 4.$

$X_n \not\to X$ a.s.

- Cauchy Criterion

  Recall that

Def. $(x_n)_{n\ge1}$ is a Cauchy sequence if

$\forall\varepsilon>0,\ \forall N(\varepsilon)\in\mathbb{N},\ \forall m\in\mathbb{N},\ n \ge N(\varepsilon) \Rightarrow |x_{n+m} - x_n| < \varepsilon$

Theorem. $(x_n)_{n\ge1}$ $x_n\in\mathbb{R}$. $x\in\mathbb{R}$. Existence Theorem.

$(x_n)_{n\ge1}$ is Cauchy $\Leftrightarrow$ x_n converges to $\boxed{\text{some}}\ x\in\mathbb{R}$.

$\exists x\in\mathbb{R}$.

– For a sequence of random variables,

(i) $(X_n)_{n\ge1}$ converges a.s. $\Leftrightarrow$ $P(\{\omega : \boxed{X_n(\omega) \text{ is a Cauchy seq}}\}) = 1$.

$\forall m\in\mathbb{N},\ |X_{n+m}(\omega) - X_n(\omega)| \to 0$ as $n\to\infty$

(ii) $(X_n)_{n\ge1}$ " in the mean-square sense.

$\Leftrightarrow \forall m\in\mathbb{N},\ E\{|X_{n+m} - X_n|^2\} \to 0$ as $n\to\infty$

(iii) $(X_n)_{n\ge1}$ converges in p $\Leftrightarrow$ $\forall\varepsilon>0,\ P(|X_{n+m} - X_n| > \varepsilon) \to 0,\ \forall m$

Law of Large Numbers (LLN)

- Simply put, a law of large numbers is about
  - the convergence of the arithmetic average of random variables
    to the statistical average
  - in some sense.
    - If the mode is the convergence a.s., then it is called a strong law of large numbers (SLLN).
    - If the mode is the convergence in probability, then it is called a weak law of large numbers (WLLN).
- To define the arithmetic average, we first need a sequence of random variables $(X_i)_{i=1}^{\infty}$.
  - There are many different versions of LLNs.

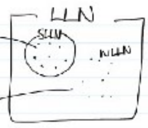
- Each LLN has a different set of restrictions on the random variables $X_1, X_2, \ldots$.
- Thus, you must be very careful in invoking a law of large numbers.
- However, they mostly have the common condition of $\mathbb{E}[X_i] = \mu_X$, a constant for all i.

- Thus, an LLN is to show

$$M_n \triangleq \frac{X_1 + X_2 + \cdots + X_n}{n} \longrightarrow \mu_X$$

in some sense under certain conditions on $X_1, X_2, \ldots$.

I SLLN, WLLN



ex/ X_i's are i.i.d. r.v.'s with $E\{X_i\} = \mu$, $\forall i$

ex/ X_i's are pairwise uncorrelated r.v.'s w/ $E\{X_i\} = \mu$, $Var\{X_i\} = \sigma^2$

sample mean / arithmetic average — statistical (ensemble) average

$$\boxed{\overline{X}_n} \triangleq \frac{X_1 + X_2 + \cdots + X_n}{n} \longrightarrow \mu \qquad \text{a.s. SLLN} \quad \text{P WLLN}$$

II CLT

X_i's i.i.d. r.v.'s w/ $E\{X_i\} = \mu$, $Var\{X_i\} = \sigma^2$

$$\underbrace{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{(X_i - \mu)}{\sigma}}_{\text{mean } 0,\ \text{variance } 1} \xrightarrow{d} X \sim N(0,1) \iff \frac{1}{\sqrt{n}} \Sigma (X_i - \mu) \to X \sim N(0, \sigma^2)$$

$$\not\Leftrightarrow \underbrace{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i}_{\text{mean } \sqrt{n}\mu,\ \text{variance } \sigma^2} \to \cancel{X \sim N(\mu, \sigma^2)}$$

$$= \sqrt{n}\,\overline{X}_n \qquad N(\sqrt{n}\mu, \sigma^2)$$

$$\overline{X}_n \qquad N(\mu, \tfrac{\sigma^2}{n})$$

- Theorem. (an SLLN) Given a sequence $(X_i)_{i=1}^{\infty}$ of independent and identically distributed random variables with mean μ_X and variance $\sigma_X^2 < \infty$, the arithmetic average converges almost surely to the statistical average μ_X, i.e.,

a RV.

$$M_n \triangleq \frac{X_1 + X_2 + \cdots + X_n}{n} \longrightarrow \mu_X \text{ w.p. } 1.$$

a.s.

- Proof. Omitted.

- Theorem. (a WLLN) Given a sequence $(X_i)_{i=1}^{\infty}$ of pairwise uncorrelated random variables with same mean μ_X and same variance $\sigma_X^2 < \infty$, the arithmetic average converges in probability to the statistical average μ_X, i.e.,

$$M_n \triangleq \frac{X_1 + X_2 + \cdots + X_n}{n} \longrightarrow \mu_X \text{ in prob.}$$

- Proof. Use the Chebychev inequality.

$\boxed{\text{i.i.d. w/ a finite variance.}}$

a.s. $\Downarrow$ in the pth mean $\Downarrow$

in prob

$\Downarrow$

in distribution

- Theorem (a WLLN)
  Given a seq. of pairwise uncorrelated r.v.'s X_i w/ $E[X_i]=\mu_X$ & $Var[X_i]=\sigma_X^2$ $\forall i$,
  $$Y_n=\frac{X_1+X_2+\cdots+X_n}{n}\to\mu_X \text{ in } P.$$

Proof)
To show $\forall\varepsilon>0$, $P(\{s\in S: |Y_n(s)-\mu_X|>\varepsilon\})\to 0$
Note that
$$Pr(|Y_n-\mu_X|>\varepsilon)=Pr(|Y_n-\mu_X|^2>\varepsilon^2)\le\frac{E\{|Y_n-\mu_X|^2\}}{\varepsilon^2}=\frac{\sigma_X^2}{n\varepsilon^2}\to 0,\ \forall\varepsilon>0$$
where
$$E\{|Y_n-\mu_X|^2\}=E\left\{\left|\frac{X_1+X_2+\cdots+X_n}{n}-\mu_X\right|^2\right\}$$
$$=E\left\{\left|\frac{(X_1-\mu_X)+(X_2-\mu_X)+\cdots+(X_n-\mu_X)}{n}\right|^2\right\}=E\left\{\left|\frac{1}{n}\sum_{i=1}^n(X_i-\mu_X)\right|^2\right\}$$
$$=\frac{1}{n^2}E\left\{\left(\sum_{i=1}^n(X_i-\mu_X)\right)\left(\sum_{j=1}^n(X_j-\mu_X)\right)\right\}=\frac{1}{n^2}\cdot n\cdot\sigma_X^2=\frac{\sigma_X^2}{n}$$

Central Limit Theorem (CLT)

- Simply put, a central limit theorem is about
  - the convergence of $\sqrt{n}$ times the arithmetic average of zero-mean random variables
  - to a Gaussian random variable
  - in distribution.
- Again, to define the arithmetic average, we first need a sequence of random variables $(X_i)_{i=1}^{\infty}$.
  - There are many different versions of CLTs.
  - Each CLT has a different set of restrictions on the random variables $X_1, X_2, \ldots$.
  - Thus, you must be very careful in invoking a CLT.
- Although
  - the exact CLT is to justify a Gaussian approximation to $\sqrt{n}$ times the arithmetic average of zero-mean random variables,
  - it is often used to justify a Gaussian approximation to the arithmetic average of nonzero-mean random variables.

the set of all CLTs

~~A~~ Sequence of real RV converges in distribution to a standard Gaussian RV.

$X\sim N(0,1)$

$\cancel{Y=-X}$

CDF $F_{Z_n}(z)\to F_Z(z)=\int_z^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$

CF $\Phi_{Z_n}(\omega)\to\Phi_Z(\omega)=e^{-\frac{\omega^2}{2}}$

- Theorem. (a CLT) Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with mean μ_X and variance σ_2, and define
  $$Z_n \triangleq \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i-\mu_X}{\sigma}.$$
  Then, the characteristic function $\Phi_{Z_n}(\omega)$ of Z_n converges to that of a Gaussian random variable with mean 0 and variance 1, i.e.,
  $$\lim_{n\to\infty}\Phi_{Z_n}(\omega)=\exp\left(-\frac{\omega^2}{2}\right).$$
  - Proof. Omitted.
- The above CLT says
  $$\frac{X_1+X_2+\ldots X_n}{n}\approx Y\sim\mathcal{N}\left(\mu_X,\frac{\sigma^2}{n}\right).$$
  However, an LLN says
  $$\frac{X_1+X_2+\ldots X_n}{n}\approx\mu_X.$$
  Which approximation do you prefer?