7-Mar-2024

EE503: EXTRA SESSION ON STOCHASTIC CONVERGENCE

- Stochastic Convergence (UCMOPED) & Properties
- Markov's Inequality, Chebyshev's Inequality, Variance-Bias Decomposition
- Law of Large Numbers (LLN)
- Problems

Convergence: ($\forall\exists\forall$)

$$\lim_{n\to\infty} a_n = a \Leftrightarrow a_n \to a \Leftrightarrow \forall\, \varepsilon>0\ \exists\, n_0\in\mathbb{Z}^+ : \forall\, n\geq n_0,\ |a_n - a| < \varepsilon$$

Stochastic Convergence: (UCMOPED)

Important!! $u \to e \to o \to p \to d$, $m \to p$

random

U	C	M	O	P	E	D
Uniform	Cauchy	Mean-Squared	with probability One	in Probability	everywhere	in Distribution

almost everywhere (O)

almost surely (O)

1. $X_n \xrightarrow{e} X$ ($\forall\forall\exists\forall$) $\Leftrightarrow$ $\boxed{\forall\, \omega\in\Omega : X_n(\omega) \to X(\omega)}$

$\Leftrightarrow \forall\, \omega\in\Omega\ \forall\, \varepsilon>0\ \exists\, n_0\in\mathbb{Z}^+ : \forall\, n\geq n_0,\ |X_n(\omega) - X(\omega)| < \varepsilon$

$\Leftrightarrow \forall\, \varepsilon>0\ \forall\, \omega\in\Omega\ \exists\, n_0\in\mathbb{Z}^+ : \forall\, n\geq n_0,\ |X_n(\omega) - X(\omega)| < \varepsilon$

$\Leftarrow \forall\, \varepsilon>0\ \exists\, n_0\in\mathbb{Z}^+ : \forall\, \omega\in\Omega\ \forall\, n\geq n_0,\ |X_n(\omega) - X(\omega)| < \varepsilon$

$\Leftrightarrow \forall\, \varepsilon>0\ \exists\, n_0\in\mathbb{Z}^+ : \forall\, n\geq n_0\ \forall\, \omega\in\Omega,\ |X_n(\omega) - X(\omega)| < \varepsilon$

$\Leftrightarrow \boxed{\forall\, \varepsilon>0\ \exists\, n_0\in\mathbb{Z}^+ : \forall\, n\geq n_0,\ |X_n(\omega) - X(\omega)| < \varepsilon\ \forall\, \omega\in\Omega}$

REMEMBER!

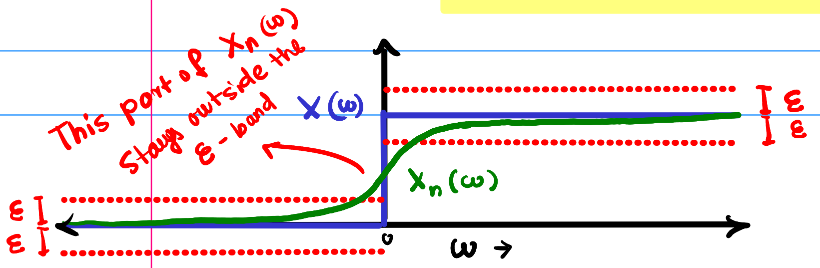
$\forall\forall \Leftrightarrow \forall\forall$

$\exists\exists \Leftrightarrow \exists\exists$

$\forall\exists \Rightarrow \exists\forall$ ($\cancel{\Leftarrow}$)

2. $\Leftrightarrow X_n \xrightarrow{u} X$ ($\forall\exists\forall\forall$)

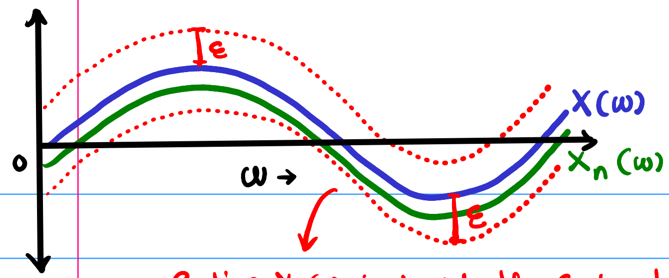
$\therefore\ \boxed{X_n \xrightarrow{u} X \Rightarrow X_n \xrightarrow{e} X\ (u \to e)}$



$$X(\omega) = \begin{cases} 0 & ; \ \omega<0 \\ 0.5 & ; \ \omega=0 \\ 1 & ; \ \omega>0 \end{cases}$$

$$X_n(\omega) = \frac{1}{1+e^{-n\omega}}$$

$X_n \xrightarrow{e} X$ but $X_n \not\xrightarrow{u} X$



Entire $X_n(\omega)$ is inside the ε-band

$X(\omega) = \sin\omega$

$X_n(\omega) = \sin\omega - 1/n$

$X_n \xrightarrow{u} X$

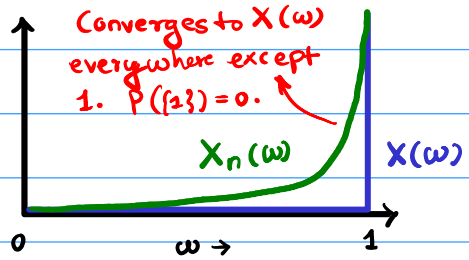
3. $X_n \xrightarrow{o} X \Leftrightarrow \boxed{P(X_n \to X) = 1} \Leftrightarrow P(\{\omega \in \Omega : X_n(\omega) \to X(\omega)\}) = 1$

(P(lim)) $\Leftrightarrow P(\{\omega \in \Omega : \forall \varepsilon > 0\ \exists n_0 \in \mathbb{Z}^+ : \forall n \geq n_0, |X_n(\omega) - X(\omega)| < \varepsilon\}) = 1$

$X_n \xrightarrow{e} X \Rightarrow P(X_n \to X) = P(\{\omega \in \Omega : X_n(\omega) \to X(\omega)\}) = P(\Omega) = 1$

$\Rightarrow X_n \xrightarrow{o} X$

$\therefore \boxed{X_n \xrightarrow{e} X \Rightarrow X_n \xrightarrow{o} X \quad (o \to e)}$



$P([0, \omega)) = \omega, \quad 0 < \omega < 1.$

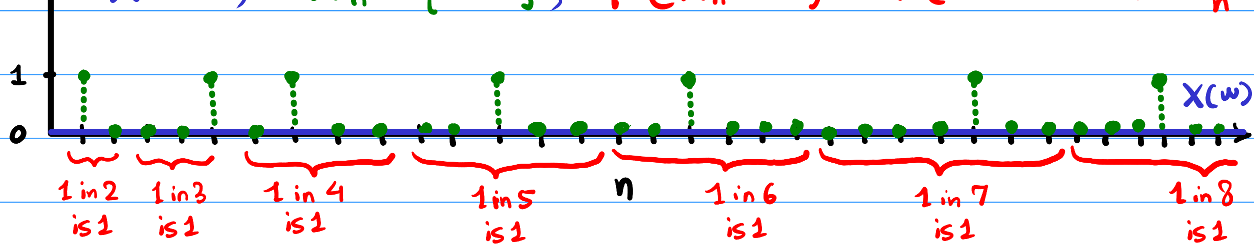
$X(\omega) = \begin{cases} 0, & 0 \leq \omega < 1 \\ 1, & \omega = 1 \end{cases}$ $\quad X_n(\omega) = \omega^n$

$X_n \xrightarrow{o} X$ but $X_n \cancel{\xrightarrow{e}} X$

4. $X_n \xrightarrow{p} X \Leftrightarrow \boxed{\forall \varepsilon > 0, \lim_{n \to \infty} P(|X_n - X| > \varepsilon) = 0}$

(lim P)

$X = 0, \quad X_n \in \{0, 1\}, \quad P(X_n = 1) = P(|X_n - X| > \varepsilon) = \frac{1}{n} \to 0$



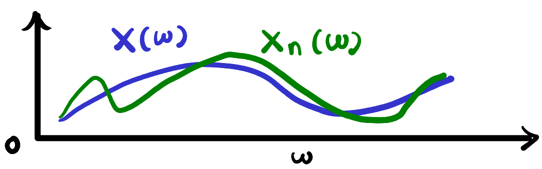
$X_n \xrightarrow{p} X$ but $X_n \cancel{\xrightarrow{e}} X$

$\boxed{X_n \xrightarrow{o} X \Rightarrow X_n \xrightarrow{p} X \quad (o \to p)}$

5. $X_n \xrightarrow{m} X \Leftrightarrow \boxed{\lim_{n \to \infty} E[(X_n - X)^2] = 0}$

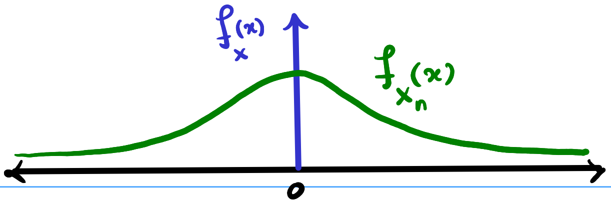
(MSE→0)

Regression:



n = # of epochs

Training error = $MSE(X_n, X) \to 0$



$X = 0 \qquad X_n \sim \text{Cauchy}(0, 1/n)$

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

$E[(X_n - X)^2] = V[X_n] \to \infty$

$X_n \xrightarrow{p} X$ but $X_n \not\xrightarrow{m} X$

$$X_n \xrightarrow{m} X \Rightarrow E[(X_n - X)^2] \to 0$$

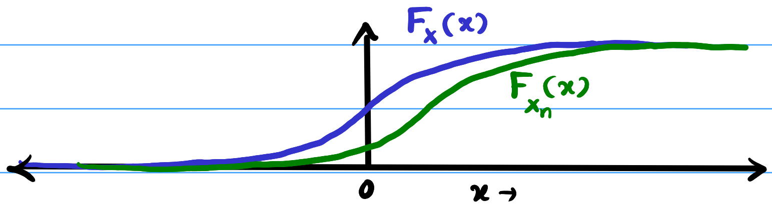
$$P(|X_n - X| > \varepsilon) = P((X_n - X)^2 > \varepsilon^2) \underset{MI}{\leq} \frac{E[(X_n - X)^2]}{\varepsilon^2} \to 0$$

$\therefore X_n \xrightarrow{p} X$

$$\therefore \boxed{X_n \xrightarrow{m} X \Rightarrow X_n \xrightarrow{p} X \quad (m \to p)}$$

6. $X_n \xrightarrow{d} X \iff$ $F_{X_n}(x) \to F_X(x)$ at all points of continuity of F_X

Eg: poisson law. $b(n,p) \xrightarrow{d} p(\lambda)$ if $\lambda = np$



$X \sim \mathcal{N}(0, 1)$

$X_n \sim \mathcal{N}(1/n, 1)$

$X_n - X \sim \mathcal{N}(1/n, 2)$

$X_n \xrightarrow{d} X$ but $X_n \not\xrightarrow{p} X$

$$\boxed{X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{d} X \quad (p \to d)}$$

7. Cauchy: Replace X with X_m & $\lim\limits_{n\to\infty}$ with $\lim\limits_{m\to\infty} \lim\limits_{n\to\infty}$

Example: $\{X_n\}$ is cauchy in mean-squared sense

$$\iff \lim_{n\to\infty} \lim_{m\to\infty} E[(X_n - X_m)^2] = 0.$$

X_n is cauchy in mean-squared sense $\Rightarrow X_n \xrightarrow{m} X$

BUT **NOT** for other modes of convergence.

(X_n is cauchy in p $\not\Rightarrow$ $X_n \xrightarrow{p} X$) unless complete space

IMPORTANT!

mode of convergence	e	p	m	d
a_n	$X_n(\omega)$	$P(\lvert X_n - X\rvert > \varepsilon)$	$E[(X_n - X)^2]$	$F_{X_n}(x)$
$a_n \to a$ a	$X(\omega)$	0	0	$F_X(x)$

Properties of Stochastic Convergence

1. $X_n \xrightarrow{a} X$ & $Y_n \xrightarrow{a} Y \Rightarrow X_n + Y_n \xrightarrow{a} X+Y$ & $X_nY_n \xrightarrow{a} XY$
2. $X_n \xrightarrow{P} X$ & $Y_n \xrightarrow{P} Y \Rightarrow X_n + Y_n \xrightarrow{P} X+Y$ & $X_nY_n \xrightarrow{P} XY$
3. $X_n \xrightarrow{m} X$ & $Y_n \xrightarrow{m} Y \Rightarrow X_n + Y_n \xrightarrow{m} X+Y$
4. $X_n \xrightarrow{a} X$ & g is continuous $\Rightarrow g(X_n) \xrightarrow{a} g(X)$
5. $X_n \xrightarrow{P} X$ & g is continuous $\Rightarrow g(X_n) \xrightarrow{P} g(X)$
6. $X_n \xrightarrow{d} X$ & g is continuous $\Rightarrow g(X_n) \xrightarrow{d} g(X)$

Continuity Theorem

Proof (2): **Lemma:** $a+b>c \Rightarrow a > c/2$ or $b > c/2$

Proof: Prove the contra positive: $a \le c/2$ & $b \le c/2 \Rightarrow a+b \le c/2$

$a \le c/2$ & $b \le c/2 \Rightarrow a+b \le c/2 + c/2 = c$ QED

$X_n \xrightarrow{P} X \Rightarrow \lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0 \quad \forall \varepsilon > 0$

$Y_n \xrightarrow{P} Y \Rightarrow \lim_{n\to\infty} P(|Y_n - Y| > \varepsilon) = 0 \quad \forall \varepsilon > 0$

Pick $\varepsilon > 0$.

$A = \{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon/2\}$

$B = \{\omega \in \Omega : |Y_n(\omega) - Y(\omega)| > \varepsilon/2\}$

$C = \{\omega \in \Omega : |X_n(\omega) + Y_n(\omega) - X(\omega) - Y(\omega)| > \varepsilon\}$

$\omega \in C$

$\Rightarrow |X_n(\omega) - X(\omega)| + |Y_n(\omega) - Y(\omega)| \overset{TI}{\geq} |X_n(\omega) + Y_n(\omega) - X(\omega) - Y(\omega)| > \varepsilon$

$\Rightarrow |X_n(\omega) - X(\omega)| > \varepsilon/2$ or $|Y_n(\omega) - Y(\omega)| > \varepsilon/2$

$\Rightarrow \omega \in A$ or $\omega \in B \Rightarrow \omega \in A \cup B$

$\therefore C \subset A \cup B$

$\therefore P(C) \le P(A \cup B) \overset{BI}{\le} P(A) + P(B)$

$\therefore P(|X_n + Y_n - X - Y| > \varepsilon) \le P(|X_n - X| > \varepsilon) + P(|Y_n - Y| > \varepsilon)$

$\therefore 0 \le \lim_{n\to\infty} P(|X_n + Y_n - X - Y| > \varepsilon) \le \lim_{n\to\infty} P(|X_n - X| > \varepsilon) + \lim_{n\to\infty} P(|Y_n - Y| > \varepsilon)$

$= 0 + 0 = 0$

$\therefore \lim_{n\to\infty} P(|X_n + Y_n - X - Y| > \varepsilon) = 0 \quad \therefore X_n + Y_n \xrightarrow{P} X+Y$ QED

Markov's Inequality: If $X > 0$ & $E[X] < \infty$ then

$$\boxed{P(X > \varepsilon) \leq \frac{E[X]}{\varepsilon}}$$

Proof:

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_{-\infty}^{\varepsilon} x f_X(x)\,dx + \int_{\varepsilon}^{\infty} x f_X(x)\,dx$$

$$\geq \int_{\varepsilon}^{\infty} x f_X(x)\,dx \geq \int_{\varepsilon}^{\infty} \varepsilon f_X(x)\,dx = \varepsilon \int_{\varepsilon}^{\infty} f_X(x)\,dx$$

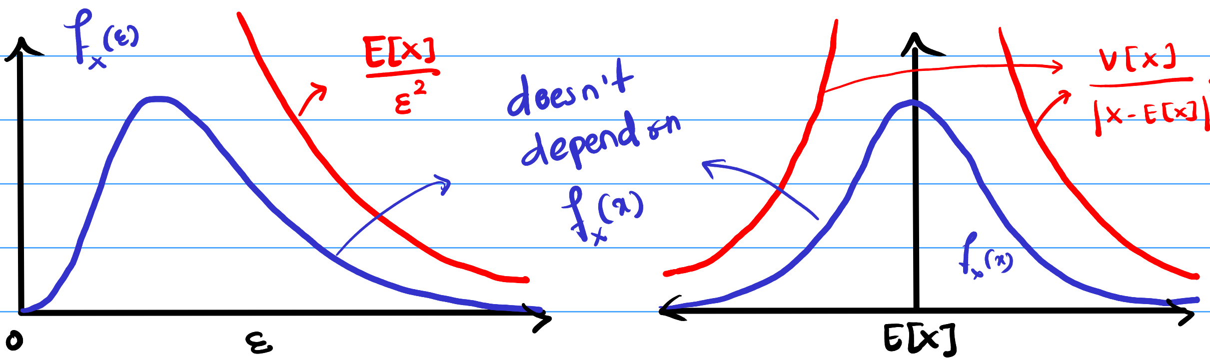
$$= \varepsilon P(X > \varepsilon)$$

$$\therefore \quad P(X > \varepsilon) \leq \frac{E[X]}{\varepsilon} \qquad \text{QED}$$

Chebyshev's Inequality: If $V[X] < \infty$ then

$$\boxed{P(|X - E[X]| > \varepsilon) \leq \frac{V[X]}{\varepsilon^2}}$$

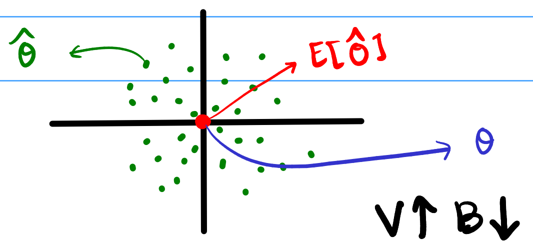
This has to be the mean of the random variable in front of it. (X)



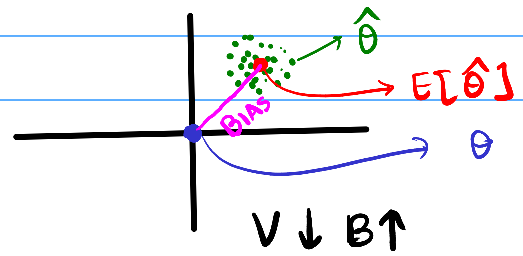
MARKOV'S INEQUALITY CHEBYSHEV'S INEQUALITY

Variance - Bias Decomposition:

$$\boxed{\underbrace{E[(\hat{\theta} - \theta)^2]}_{\text{MSE}} = \underbrace{V[\hat{\theta}]}_{\text{Variance}} + \underbrace{(E[\hat{\theta}] - \theta)^2}_{\text{Bias}^2}}$$



$V\uparrow B\downarrow$



$V\downarrow B\uparrow$

Sampling Statistics: If $X_1, \cdots, X_n$ iid $E[X] = \mu_x$ & $V[X] = \sigma_x^2$

$$E[\bar{X}_n] = \mu_x$$
$$V[\bar{X}_n] = \sigma_x^2/n$$

where $\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$

Law of Large Numbers: (LLN)

If $X_1, \cdots, X_n$ iid & $\sigma_x^2 < \infty$ then

$$\bar{X}_n \xrightarrow{mopd} \mu_x$$

1. $\bar{X}_n \xrightarrow{o} \mu_x$ Strong LLN (SLLN)
2. $\bar{X}_n \xrightarrow{p} \mu_x$ Weak LLN (WLLN)
3. $\bar{X}_n \xrightarrow{m} \mu_x$ Mean-squared LLN (MSLLN)

Proof (2): Pick $\varepsilon > 0$.

$$0 \leq \lim_{n\to\infty} P(|\bar{X}_n - \mu_x| > \varepsilon) = \lim_{n\to\infty} P(|\bar{X}_n - E[\bar{X}_n]| > \varepsilon)$$
$$\underset{CI}{\leq} \lim_{n\to\infty} \frac{V[\bar{X}_n]}{\varepsilon^2} = \lim_{n\to\infty} \frac{\sigma_x^2}{n\varepsilon^2} = 0.$$

$$\therefore \lim_{n\to\infty} P(|\bar{X}_n - \mu_x| > \varepsilon) = 0$$

$$\therefore \bar{X}_n \xrightarrow{p} \mu_x$$

QED.

Q: The random sequence $X_1, X_2, X_3, \ldots$ consists of similarly distributed gamma random variables $X_n \sim \Gamma(\frac{n^2}{2}, 2)$. Define a new similarly distributed sequence $Y_1, Y_2, Y_3, \ldots$ by $Y_n = X_n/n^3$. Where does the random sequence $Y_1, Y_2, Y_3, \ldots$ converge in distribution?

F: Convergence in distribution

M: $X_n \xrightarrow{d} X \Leftrightarrow F_{X_n}(x) \to F_X(x)$ at all points of continuity of F_X

also $m - p - d$

A: $X_n \sim \gamma(n^2/2, 2) \Rightarrow E[X_n] = n^2, \quad V[X_n] = 2n^2$

$$E[Y_n] = \frac{n^2}{n^3} = \frac{1}{n} \longrightarrow \boxed{0} \qquad V[Y_n] = \frac{2n^2}{n^6} \longrightarrow \boxed{0}$$

Note: If $V[Y_n] \to 0$ then always check for convergence at $\lim_{n\to\infty} E[Y_n]$

Method 1: p & p → d.

$$0 \le \lim_{n\to\infty} P(|Y_n - 0| > \varepsilon) = \lim_{n\to\infty} P(Y_n > \varepsilon) \underset{MI}{\le} \lim_{n\to\infty} \frac{E[Y_n]}{\varepsilon}$$

$$= \lim_{n\to\infty} \frac{1}{n\varepsilon} = 0$$

$$\therefore \lim_{n\to\infty} P(|Y_n - 0| > \varepsilon) = 0 \quad \therefore Y_n \xrightarrow{p} 0 \Rightarrow Y_n \xrightarrow{d} 0 \quad (m \to p \to d)$$

Method 2: m & m → p → d

$$\lim_{n\to\infty} E[(Y_n - 0)^2] = \lim_{n\to\infty} E[Y_n^2] = \lim_{n\to\infty} V[Y_n] + \lim_{n\to\infty} (E[Y_n])^2$$

$$= \lim_{n\to\infty} \frac{2}{n^4} + \lim_{n\to\infty} \frac{1}{n^2} = 0 + 0 = 0$$

$$\therefore Y_n \xrightarrow{m} 0 \Rightarrow Y_n \xrightarrow{p} 0 \Rightarrow Y_n \xrightarrow{d} 0 \quad (m \to p \to d)$$

C: 0.

The random sequence $X_1, X_2, \ldots$ consists of similarly distributed binomial random variables with $X_n \sim \text{Binomial}\left(n, \frac{1}{n}\right)$. Define a new sequence of random variables $Y_n = \frac{X_n}{\sqrt{n}}$. Where if anywhere does the random sequence $Y_1, Y_2, \ldots$ converge **in probability**?

I: Convergence in probability

M: $X_n \xrightarrow{p} X \Leftrightarrow \forall \varepsilon > 0, \lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$

also m → p → d

A: $X_n \sim b(n, 1/n) \Rightarrow E[X_n] = n\frac{1}{n} = 1$ & $V[X_n] = n\frac{1}{n}\left(1 - \frac{1}{n}\right)$
$= 1 - 1/n$

$$Y_n = \frac{X_n}{\sqrt{n}} \Rightarrow E[Y_n] = E\left[\frac{X_n}{\sqrt{n}}\right] = \frac{1}{\sqrt{n}} E[X_n] = \frac{1}{\sqrt{n}} \cdot 1 = \frac{1}{\sqrt{n}} \longrightarrow 0$$

$$\& \ V[Y_n] = V\left[\frac{X_n}{\sqrt{n}}\right] = \frac{1}{n} V[X_n] = \frac{1}{n}\left(1 - \frac{1}{n}\right) = \frac{1}{n} - \frac{1}{n^2} \longrightarrow 0$$

Check for convergence at 0.

Method 1 : p

$$0 \le \lim_{n\to\infty} P(|Y_n - 0| > \varepsilon) = \lim_{n\to\infty} P(Y_n > \varepsilon) \overset{MI}{\le} \lim_{n\to\infty} \frac{E[Y_n]}{\varepsilon}$$

$$= \lim_{n\to\infty} \frac{1}{\sqrt{n}\,\varepsilon} = 0.$$

$$\therefore \lim_{n\to\infty} P(|Y_n - 0| > \varepsilon) = 0$$

$$\therefore Y_n \xrightarrow{p} 0 .$$

Method 2 : m & m → p

$$\lim_{n\to\infty} E[(Y_n - 0)^2] = \lim_{n\to\infty} E[Y_n^2] = \lim_{n\to\infty} V[Y_n] + \lim_{n\to\infty} (E[Y_n])^2$$

$$= \lim_{n\to\infty}\left(\frac{1}{n} - \frac{1}{n^2}\right) + \lim_{n\to\infty} \frac{1}{n^2} = 0 + 0 = 0.$$

$$\therefore \quad Y_n \xrightarrow{m} 0 \Rightarrow Y_n \xrightarrow{p} 0 . \qquad (m \to p \to d)$$