Convergence in Probability to a Constant

"X_n coverges to constant c in probability"

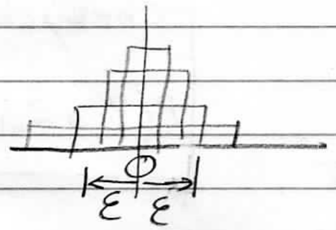
$X_n \xrightarrow{P} c$ means $\forall \varepsilon > 0 \; \lim_{n\to\infty} P(|X_n - c| \geq \varepsilon) = 0$

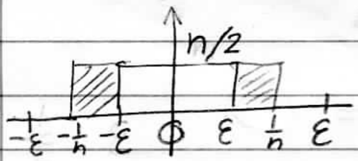
Example

Consider $X_n \sim \overset{\text{uniform}}{U}(-\frac{1}{n}, \frac{1}{n}) = \frac{n}{2}\mathbb{1}_{x\in[\frac{1}{n},\frac{1}{n}]}$

Prove $X_n \xrightarrow{P} 0$

lim means $\lim_{n\to\infty}$



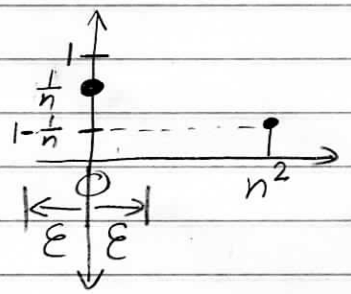


$$\begin{aligned}
&\lim P(|X_n - 0| \geq \varepsilon) \\
&= \lim P(|X_n| \geq \varepsilon) \\
&= \lim \left(P(X_n \leq -\varepsilon) + P(X_n \geq \varepsilon)\right) \\
&= \lim \left((-\varepsilon - -\tfrac{1}{n})\tfrac{n}{2}\mathbb{1}_{-\varepsilon > -n} + (\tfrac{1}{n} - \varepsilon)\tfrac{n}{2}\mathbb{1}_{\varepsilon < \frac{1}{n}}\right) \\
&= \lim 2(\tfrac{1}{n} - \varepsilon)\tfrac{n}{2}\mathbb{1}_{\varepsilon < \frac{1}{n}} \\
&= \lim (1 - \varepsilon n)\mathbb{1}_{\varepsilon < \frac{1}{n}} \\
&= 0
\end{aligned}$$

Example

Consider $X_n \sim \begin{cases} n^2 & \text{w.p. } \frac{1}{n} \\ 0 & \text{w.p. } 1 - \frac{1}{n} \end{cases}$

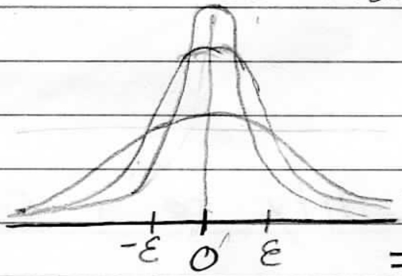
Prove $X_n \xrightarrow{P} 0$



$$\begin{aligned}
&\lim P(|X_n - 0| \geq \varepsilon) \\
&= \lim P(|X_n| \geq \varepsilon) \\
&= \lim P(X_n \geq \varepsilon) \\
&= \lim \tfrac{1}{n}\mathbb{1}_{\varepsilon < n^2} \\
&= \lim \tfrac{1}{n}(1) \quad \leftarrow \text{always 1 as } n\to\infty \\
&= 0
\end{aligned}$$

Example

Consider $X_n \sim N(\underset{\mu}{0}, \underset{\sigma^2}{\tfrac{1}{n}})$ ← Normal



$$\lim(|X_n - 0| \geq \varepsilon)$$
$$= \lim P(|X_n| \geq \varepsilon)$$
$$= \lim(P(X_n \leq -\varepsilon) + P(X_n \geq \varepsilon))$$
$$= 2 \lim P(X_n \leq -\varepsilon)$$
$$= 2 \lim\left(\int_{-\infty}^{\varepsilon} \frac{1}{\sqrt{2\pi \frac{1}{n}}} e^{-\frac{1}{2\frac{1}{n}}x^2} dx\right)$$

too difficult to do this way...

Try using Chebyshev's Inequality

$$\boxed{\begin{array}{l}\text{Chebyshev's Inequality} \\ P(Y-\mu \geq a) \leq \dfrac{\sigma_Y^2}{a^2}\end{array}}$$

$$P(|X_n| \geq \varepsilon) \leq \lim \frac{\sigma_n^2}{\varepsilon^2}$$
$$\leq \lim \frac{1}{n\varepsilon^2}$$
$$\leq \frac{1}{\varepsilon^2} \lim \frac{1}{n}$$
$$P(|X_n| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}(0) = 0$$

so $P(|X_n| \geq \varepsilon) = 0$

Consider $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim}$ with mean μ (all have variance σ^2)

Consider $\overline{X}_1 = \frac{X_1}{1}$

$$\overline{X}_2 = \frac{X_1 + X_2}{2}$$

$$\overline{X}_3 = \frac{X_1 + X_2 + X_3}{3}$$

$$\vdots$$

$$\overline{X}_n = \frac{X_1 + \ldots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$E[\overline{X}_n] = \mu$$

$$Var[\overline{X}_n] = \frac{\sigma^2}{n}$$

Show $\overline{X}_n \xrightarrow{P} \mu$

$$\lim P(|\overline{X}_n - \mu| \geq \varepsilon) \leq \lim \frac{\sigma_n^2}{\varepsilon^2}$$

$$\leq \frac{1}{\varepsilon^2} \lim \frac{\sigma^2}{n}$$

$$\leq \frac{1}{\varepsilon^2}(0) = 0$$

$$\lim P(|\overline{X}_n - \mu| \geq \varepsilon) = 0$$

So $\overline{X}_n \xrightarrow{P} \mu$

This is the "Weak Law of Large Numbers"

Convergence in L^r to a constant

stronger than convergence in probability

$r \geq 1$: $X_n \xrightarrow{L^r} c$ means $\lim_{n\to\infty} E[(|X_n - c|)^r] = 0$

2 popular values of r:

If $r=1$, $X_n \xrightarrow{L^1} c$ is called "convergence in mean"

$$\lim E[|X_n - c|] = 0$$

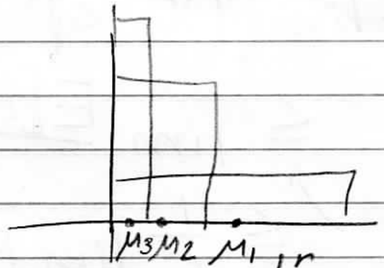
If $r=2$, $X_n \xrightarrow{L^2} c$ is called "mean square convergence"

$$\lim E[(X_n - c)^2] = 0$$

Example

Consider $X_n \sim \overset{\text{uniform}}{U}(0, \frac{1}{n}) = n \mathbb{1}_{x\in[0,\frac{1}{n}]}$

Prove $X_n \xrightarrow{L^r} 0$ for all $r \in \{1, 2, \dots\}$



Examine $X_n \xrightarrow{L^1} 0$

$\lim E[X_n] = 0$ OK

Prove $X_n \xrightarrow{L^r} 0$

Proof:

$$\lim E[(|X_n - 0|)^r] = \lim E[(X_n)^r]$$

$$= \lim \int_{\mathbb{R}} x^r \mathbb{1}_{x\in[0,\frac{1}{n}]}\, dx = \lim_{n\to\infty} \int_0^{\frac{1}{n}} x^r\, dr$$

$$= \lim_{n\to\infty} \left[\frac{x^{r+1}}{r+1}\right]_0^{\frac{1}{n}} = \lim_{n\to\infty} \frac{1}{r+1}\left(\frac{1}{n^{r+1}} - 0\right)$$

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

$r, s > 0, \quad s > r$

<u>Prove:</u> $X_n \xrightarrow{L^s} c \Longrightarrow X_n \xrightarrow{L^r} c$

implies

Proof: $\lim E[(|X_n - c|)^r] \leq \lim E[(|X_n - c|)^s]$

by Holder's inequality $\leq \left(\lim E[|X_n - c|^s]\right)^{r/s}$

implies $\Longrightarrow$

$$\lim E[(|X_n - c|)^r] \leq 0^{r/s} = 0$$

Note: $E[\underbrace{(|X_n - c|^r)}_{|X_n - c| \text{ non-neg. r.v.}}] \geq 0$

Prove:
For any r, $X_n \xrightarrow{L^r} c \Longrightarrow X_n \xrightarrow{P} c$
but not the converse

converse $\Leftarrow$

$X_n \xrightarrow{L^r} c$ means $\lim P(|X_n - c| \geq \varepsilon) = 0$

$$\lim P(|X_n - c| \geq \varepsilon) = \lim (|X_n - c|)^r \geq \varepsilon^r)$$

$$\leq \lim \frac{E[(|X_n - c|)^r]}{\varepsilon^r}$$

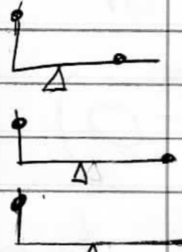
I know $\lim E[(|X - c|)^r] = 0$

by Markov's Inequality $\leq \frac{1}{\varepsilon^r} \lim E[|X_n - c|)^r]$

≤ 0

so $X_n \xrightarrow{P} 0$

for Converse:



consider $X_n \sim \begin{cases} n^2 & \text{w.p. } \frac{1}{n} \\ 0 & \text{w.p. } 1 - \frac{1}{n} \end{cases}$

$$E[(|X_n - 0|)^1] = E[X_n] = 0(1 - \tfrac{1}{n}) + n^2(\tfrac{1}{n})$$

$$= 0 + n^2 = n \leftarrow \text{not } 0$$

$\hookrightarrow$ so this is counterexample for converse

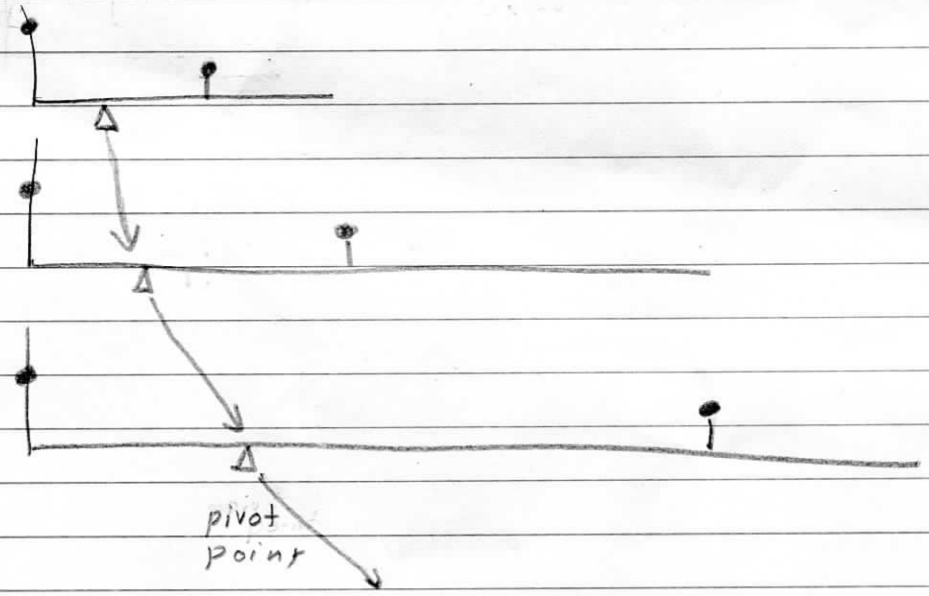
$$\lim E[(|X_n - 0|)^r] = \lim E[X_n^r]$$

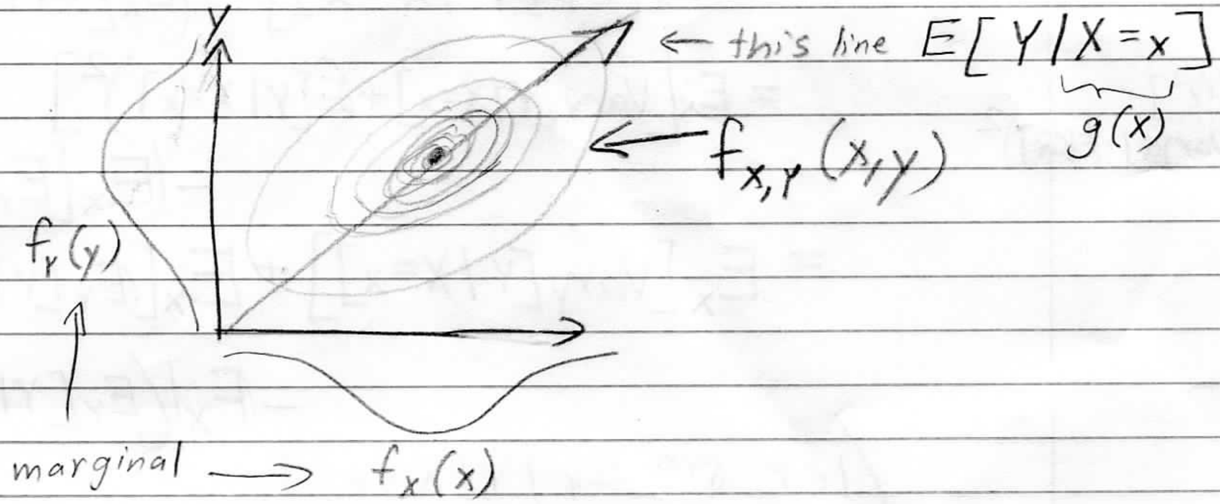
$$= \lim \left(0\left(1-\tfrac{1}{n}\right) + (n^2)^r \left(\tfrac{1}{n}\right) \right)$$

since know
$X_n \xrightarrow{P} 0$

$$= \lim n^{2r-1} = \infty \neq 0$$

$$\Rightarrow X_n \not\xrightarrow{L^r} 0$$





$$E[Y] = \int_{Supp[Y]} y\, f_Y(y)\, dy = \int_{Supp[Y]} y \left(\int_{Supp[X]} f_{XY}(x,y)\, dx \right) dy$$

$$= \int_{Supp[Y]} y \int_{Supp[X]} f_{Y|X}(x,y) f_X(x)\, dx\, dy$$

$$= \int_{Supp[X]} \int_{Supp[Y]} y\, f_{Y|X}(x,y)\, f_X(x)\, dy\, dx$$

$$= \int_{Supp[X]} f_X(x) \left(\int_{Supp[Y]} y\, f_{Y|X}(x,y)\, dy \right) dx$$

$$= \int_{Supp[X]} E_Y[Y|X=x]\, f_X(x)\, dx$$

$$\underbrace{E[Y] = E_X\left[E_Y[Y|X=x]\right]}_{\text{Law of Internal Expectation}}.$$

Law of Double Expectation?

$$\boxed{E[Y] = E_X\left[E_Y[Y|X=x]\right]}$$

$$\mathrm{Var}[Y] = E[Y^2] - (E[Y])^2$$

$$= E_X\left[E_Y[Y^2|X=x]\right] - \left(E_X\left[E_Y[Y|X=x]\right]\right)^2$$

$E[Q^2] = \mathrm{Var}[Q] + (E[Q])^2$

$$= E_X\left[\mathrm{Var}_Y[Y|X=x] + (E[Y|X=x])^2\right] - \left(E_X\left[E_Y[Y|X=x]\right]\right)^2$$

$$= E_X\left[\mathrm{Var}_Y[Y|X=x]\right] + E_X\left[(E_Y[Y|X=x])^2\right] - E_X\left[(E_Y[Y|X=x])^2\right]$$

(let $Q = E_Y[Y|X=x]$)

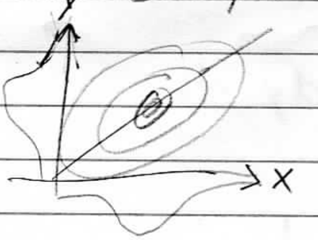
have $E_X[\mathrm{Var}_Y[Y|X=x]] + E_X[Q^2] - (E_X[Q])^2$

get

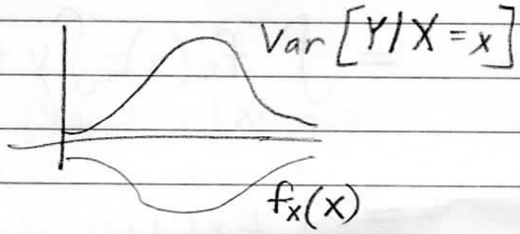
Law of Total Variance

$$\mathrm{Var}[Y] = \underbrace{E_X\left[\mathrm{Var}_Y[Y|X=x]\right]}_{\mathrm{I}} + \underbrace{\mathrm{Var}_X\left[E_Y[Y|X=x]\right]}_{\mathrm{II}}$$

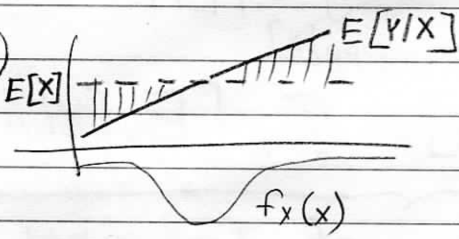
so variance can be decomposed into components I and II

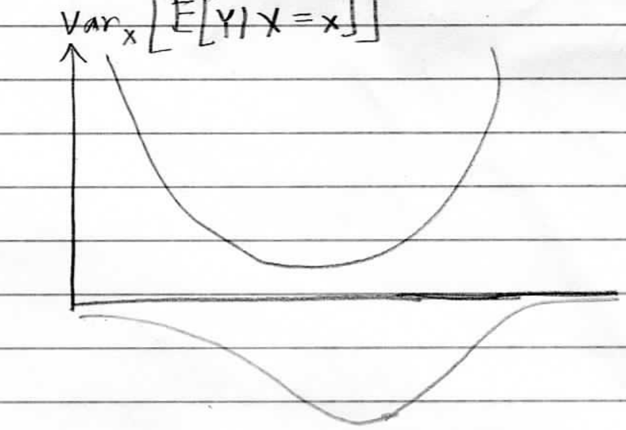


I

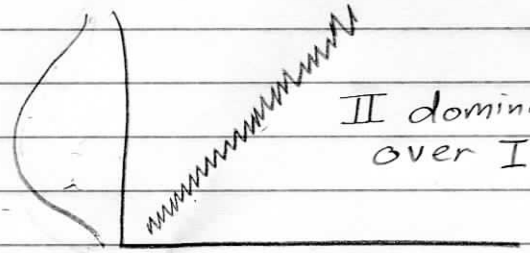


II

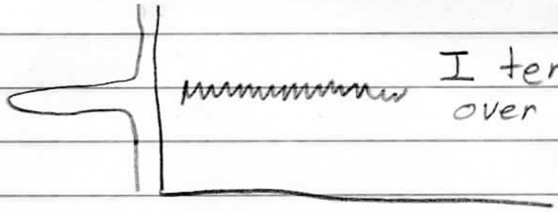




also



II dominates over I term



I term dominates over II term