lecture (Last)

Convergence in Probablity to a Constant: $X_n \xrightarrow{P} c$, This means:
$\forall \varepsilon > 0$, $\lim P(|X_n - c| \geq \varepsilon) = 0$

Easy way to Prove them:

Theorem: If X_n has Expectation μ for all n and Variance Sequence which is finite for all, then $\lim \sigma_n^2 = 0 \Rightarrow X_n \xrightarrow{P} \mu$

Proof: Recall Chebyshev's Inequality:

$$P(|X_n - \mu| \geq \varepsilon) \leq \frac{\sigma_n^2}{\varepsilon^2}$$

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

Because Probablities are between 0 and 1,
If you know the Probablity ≤ 0. That means Probablity is 0

$$\Rightarrow \lim P(|X_n - \mu| \geq \varepsilon) = 0 \Rightarrow X_n \xrightarrow{P} \mu$$

① **Example:** $X_n \sim U(-1/n, \frac{1}{n})$, Prove $X_n \xrightarrow{P} 0$

Soln: $E[X_n] = 0 = \mu \ \forall n$, $\sigma_n^2 = \dfrac{(1/n - -1/n)^2}{12} = \dfrac{4}{12n^2} = \dfrac{1}{3n^2}$

$\lim \sigma_n^2 = \lim \frac{1}{3n^2} = 0$, by theorem, $\Rightarrow X_n \xrightarrow{P} 0$ (Proved)

② **Example:** $X_n \sim N(0, \frac{1}{n})$, Prove $X_n \xrightarrow{P} 0$

$E[X_n] = 0 = \mu \ \forall n$, $\sigma^2 = 1/n$, $\lim \sigma_n^2 = \lim 1/n = 0 \Rightarrow X_n \xrightarrow{P} 0$ (Proved)

* let, $X_1, X_2 \cdots X_n$ be iid with mean μ and Var $\sigma^2 < \infty$,

$\bar{X}_n = \frac{1}{n} \sum_i X_i$, $E[\bar{X}_n] = \mu \ \forall n$, $Var[\bar{X}_n] = \frac{\sigma^2}{n}$

We wanna Prove, $\bar{X}_n \xrightarrow{P} \mu$

$$\lim Var \bar{X}_n = \lim \frac{\sigma^2}{n} = 0 \Rightarrow \bar{X}_n \xrightarrow{P} \mu$$

This is Very famous theorem, called "weak" weak Law of large numbers (W.L.N)

Beoz, I assumed finite σ^2 of the $X_1, X_2 \cdots$ one's You dont need it See HW

Beoz Convergence in Probablity is actually a weak type of Convergence. It turns out you can Prove "almost sure" conv...

"weak" Weak law of Large numbers

bcoz I assumed finite σ^2 of the $X_1, X_2 \dots$ r.v's
You dont need it (See HW)

Bcoz convergence in Probablity is actually a weak type of convergence.
It turns out you can Prove "almost Sure" convergence.
(But we wont discuss that)

As the "number" of Samples increases, the average cannot "escape" from the mean

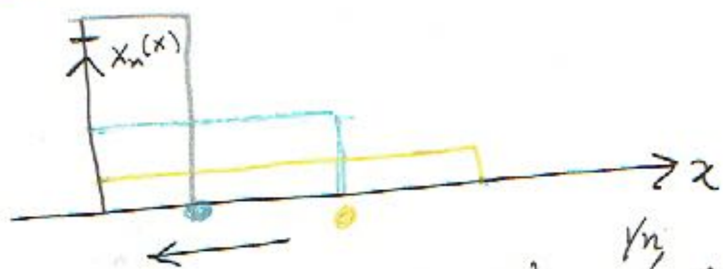
The last type of convergence we'll study is called "convergence in law" "convergence in L^r norm" to a constant where $r \geq 1$ denoted: $X_n \xrightarrow{L^r} c$ which means by defination.

$$\lim E\left[|X_n - c|^r\right] = 0$$

e.g. $r=1$, $\lim E\left[|X_n - c|\right] = 0$ "Convergence in mean"

$r=2$, $\lim E\left[(X_n - c)^2\right] = 0$ "mean Square Convergence"

e.g: $X_n \sim U(0, 1/n)$



$\boxed{\text{Prove}}$ $X_n \xrightarrow{L^r} 0$,

$$\lim E\left[|X_n - 0|^r\right] = \lim E\left[X_n^r\right] = \lim \int x^r n \, \mathbb{1}_{x \in [0, \frac{1}{n}]} \, dx = \lim n \int_0^{1/n} x^r \, dx$$

$$= \frac{1}{r+1} \lim n \frac{1}{n^{r+1}}$$

$$= \frac{1}{r+1} \lim \frac{1}{n^r} = 0$$

HW: Convergence in Probablity is Stronger than convergence in distribution which convergence is stronger? Law or Probablity?

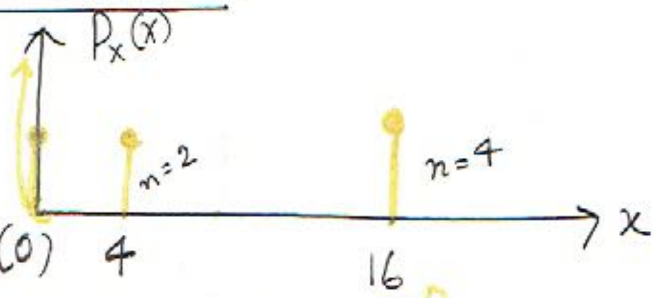
Soln: Proof for $X_n \xrightarrow{L^r} c \Rightarrow X_n \xrightarrow{P} c$

$$\lim P(|X_n - c| \geq \varepsilon) = \text{limit}\, P(|X_n - c|^r \geq \varepsilon^r) \underset{\text{Markov's}}{\leq} \lim \frac{E\left[|X_n - c|^r\right]}{\varepsilon^r} = 0$$

$\boxed{\text{The Converse is not TRUE}}$

$X_n \xrightarrow{P} C \not\Rightarrow X_n \xrightarrow{L^1} C$, Counter example:

e.g.: $X_n \sim \begin{cases} n^2 & \text{wp } 1/n \\ 0 & \text{wp } 1-\frac{1}{n} \end{cases}$



It clear that,

$$X_n \xrightarrow{P} 0$$

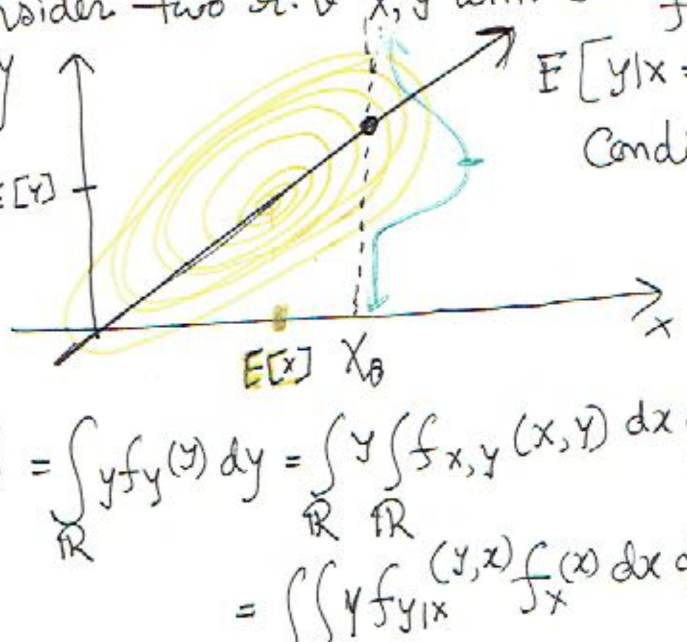
$$\lim E\left[|X_n - 0|^1\right] = \lim E\left[X_n^1\right] = \lim \sum_{x \in \{0, n^2\}} x P_X(x) = \lim 0(1 - 1/n) + n^2 \cdot 1/n$$

$$= \lim n^2/n = \lim n = \infty \neq 0$$

$$\Rightarrow X_n \xrightarrow{L^1}\!\!\!\!\!\!/\;\; 0$$

Law of Iterated Expectation

Consider two r.v X, Y with Jdf $f_{X,Y}(x,y)$



$E[Y|X=x] = E[Y|X]$, this is called the Conditional Expectation function (CEF)

$$E[Y] = \int_{\mathbb{R}} y f_Y(y)\, dy = \int_{\mathbb{R}} y \int_{\mathbb{R}} f_{X,Y}(x,y)\, dx\, dy$$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} y f_{Y|X}(y,x) f_X(x)\, dx\, dy$$

Now will Switch the order of integration

$$= \int_{\mathbb{R}} f_X(x) \underbrace{\int_{\mathbb{R}} y f_{Y|X}(y,x)\, dy}_{E[Y|X]}\, dx = \int E[Y|X] f_X(x)\, dx = E_X\left[E_Y[Y|X]\right]$$

Low of Total Variance:

Law of Total Variance :

$$Var_y[y] = E_y[y^2] - E_y[y]^2$$

$$= E_X\left[E_X[Y^2|X]\right] - E_X\left[E_y[y|x]\right]^2$$

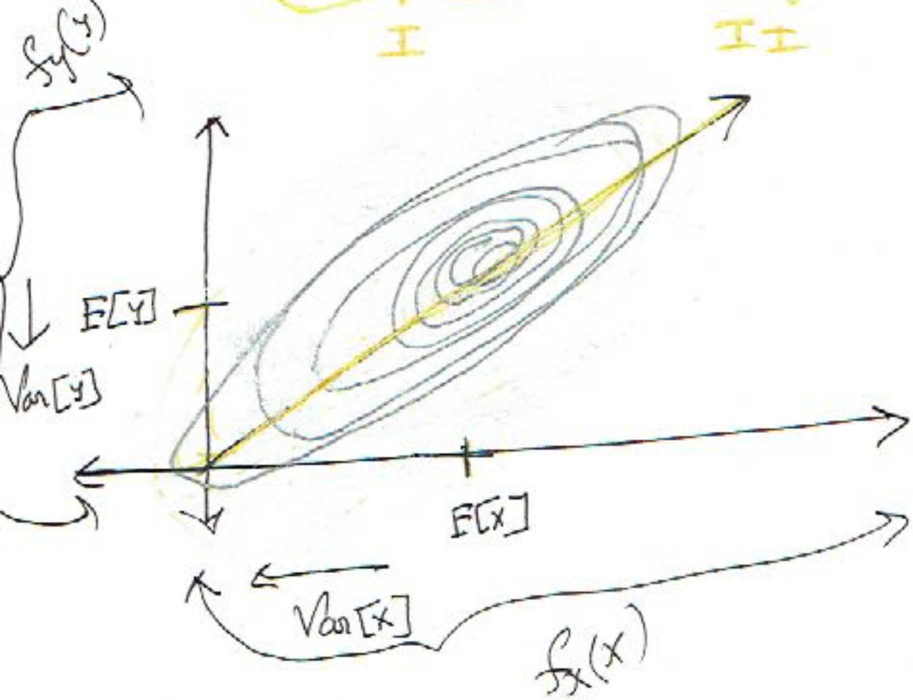
$$= E_X\left[Var_y[y|x] + E_y[y|x]^2\right] - E_X\left[E_y[y|x]\right]^2$$

$$= E_X\, Var_y[y|x] + E_X\left[E_y[y|x]^2\right] - E_X\left[E_y[y|x]\right]^2$$

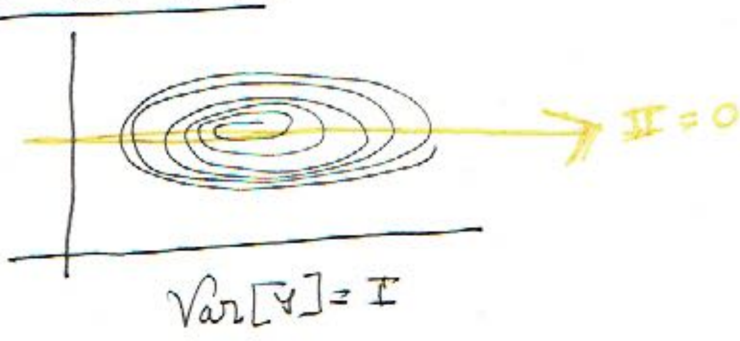
let, $C = E_y[y|x]$

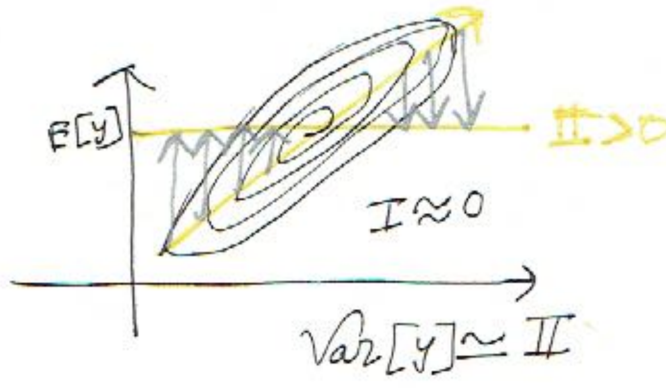
$$= E_X\left[Var_y[y|x]\right] + \underbrace{E_X[C^2] - E_X[C]^2}_{Var_X[C]}$$

$$= \underbrace{E_X\left[Var[y|x]\right]}_{\mathrm{I}} + \underbrace{Var_X\left[E_y[y|X]\right]}_{\mathrm{II}}, \quad \boxed{\text{decomposition formula}}$$



Two Situation :





DONE !