Cauchy-Schwartz Inequality

- Let X, Y be r.v's w/ finite $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$,

Let $W = (X - cY)^2$, some constant $c \in \mathbb{R}$

Note: W is non-negative, thus $E[W] \geq 0$

$\Rightarrow E[(X-cY)^2] \geq 0$

$\Rightarrow E[X^2 - 2cXY + c^2Y^2] \geq 0$

$\Rightarrow E[X^2] - 2cE[XY] + c^2E[Y^2] \geq 0$

Let $c = \dfrac{E[XY]}{E[Y^2]}$

$\Rightarrow E[X^2] - 2\dfrac{E[XY]}{E[Y^2]}E[XY] + \dfrac{E[XY]^2}{E[Y^2]^{\cancel{2}}}\cancel{E[Y^2]} \geq 0$

$\Rightarrow E[X^2] - \dfrac{E[XY]^2}{E[Y^2]} \geq 0$

$\Rightarrow E[X^2]E[Y^2] \geq E[XY]^2$

$\Rightarrow \boxed{|E[XY]| \leq \sqrt{E[X^2]E[Y^2]}}$

If X, Y non-neg., $E[XY] \leq \sqrt{E[X^2]E[Y^2]}$

- WWTS $\mathrm{Corr}[X, Y] \in [-1, 1]$

Let $Z_X = \dfrac{X - \mu_X}{\sigma_X}$, $Z_Y = \dfrac{Y - \mu_Y}{\sigma_Y}$.

$\Rightarrow E[Z_X] = E[Z_Y] = 0$

$\Rightarrow SE[Z_X] = SE[Z_Y] = 1$

$\Rightarrow E[Z_X^2] = E[Z_Y^2] = 1$

$$\left| E[Z_X Z_Y] \right| \leq \sqrt{E[Z_X^2] E[Z_Y^2]} = 1$$

$\Rightarrow E[Z_X Z_Y] \in [-1, 1]$, thus

- $$\mathrm{Corr}[X, Y] = \frac{E[XY] - \mu_X \mu_Y}{\sigma_X \sigma_Y}$$

$$= \frac{E[(\sigma_X Z_X + \mu_X)(\sigma_Y Z_Y + \mu_Y)] - \mu_X \mu_Y}{\sigma_X \sigma_Y}$$

$$= \frac{\sigma_X \sigma_Y E[Z_X Z_Y] + \cancel{\mu_X \mu_Y} - \cancel{\mu_X \mu_Y}}{\sigma_X \sigma_Y}$$

$$= \frac{\cancel{\sigma_X \sigma_Y} E[Z_X Z_Y]}{\cancel{\sigma_X \sigma_Y}}$$

$= E[Z_X Z_Y] \in [-1, 1]$ from Cauchy-Schwartz

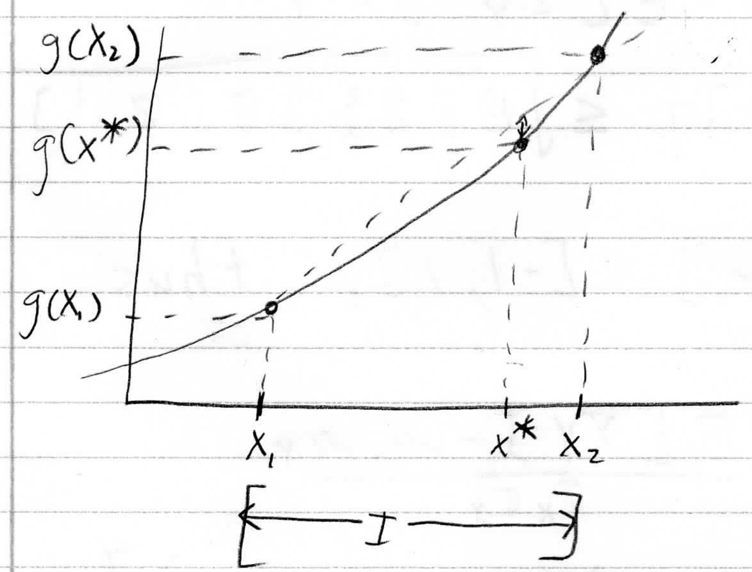
$\square$

- **Def**: g is a <u>convex function</u> on an interval $I \subset \mathbb{R}$ if $\forall \{x_1, x_2, \ldots\} \subset I$ & $\forall \{w_1, w_2, \ldots\}$ s.t. $\sum w_i = 1$ & $w_i > 0 \ \forall i$

  - Can consider the w_i's the "weights" for the x_i's.

$$g(w_1 x_1 + w_2 x_2 + \ldots) \leq w_1 g(x_1) + w_2 g(x_2) + \ldots$$

or $$g\left(\sum w_i x_i\right) \leq \sum w_i g(x_i)$$



Thm: $g''(x) \geq 0 \ \forall x \in I$
$\Longrightarrow g$ convex on I

- g is convex on support of discrete r.v X.

Let $w_i = p(x_i) \ \forall x_i \in supp[X]$.

$$g\left(\sum_{x \in supp[X]} p(x)\, x\right) \leq \sum_{x \in supp[X]} p(x)\, g(x)$$

$$\Longrightarrow \boxed{g\left(E[X]\right) \leq E\left[g(X)\right]}$$

"Jensen's Inequality"

- For g concave on $supp[X]$,

$$g\left(E[X]\right) \geq E\left[g(X)\right]$$

true for continuous r.v's too, but proof is more complicated.

ex. Let $g(x) = x^2$. By Jensen's Inequality,

$$\underbrace{E[X]^2}_{\mu^2} \leq \underbrace{E[X^2]}_{\mu^2+\sigma^2}, \quad \sigma^2 = 0 \text{ for degenerate}$$

Derivation of Young's Inequality

Let $a, b > 0$ & $p, q > 0$ s.t $\frac{1}{p} + \frac{1}{q} = 1$.

Consider $X \sim \begin{cases} a^p & wp \ 1/p \\ b^q & wp \ 1/q \end{cases}$

$$\Rightarrow E[X] = \frac{a^p}{p} + \frac{b^q}{q}.$$

Let $g(t) = -\ln(t)$ (convex)

$g(X) \sim \begin{cases} -p\ln(a) & wp \ 1/p \\ -q\ln(b) & wp \ 1/q \end{cases}$

$$\Rightarrow E[g(X)] = -\ln(a) - \ln(b) = -\ln(ab)$$

By Jensen, $-\ln\left(\frac{a^p}{p} + \frac{b^q}{q}\right) \leq -\ln(ab)$

$$\Rightarrow \ln\left(\frac{a^p}{p} + \frac{b^q}{q}\right) \geq \ln(ab)$$

$$\Rightarrow \boxed{ab \leq \frac{a^p}{p} + \frac{b^q}{q}} \quad \text{"Young's Inequality"}$$

- Let $a = X$, $b = Y$, non-neg r.v's

$$\Rightarrow \quad XY \le \frac{X^p}{p} + \frac{Y^q}{q}.$$

$$\Rightarrow \boxed{E[XY] \le \frac{E[X^p]}{p} + \frac{E[X^q]}{q}}$$

- Let $a = \dfrac{X}{E[X^p]^{1/p}}$, $b = \dfrac{Y}{E[Y^q]^{1/q}}$,

$$\Rightarrow \frac{E[XY]}{E[X^p]^{1/p}E[Y^q]^{1/q}} \le \frac{E\left[\left(\frac{X}{E[X^p]^{1/p}}\right)^p\right]}{p} + \frac{E\left[\left(\frac{Y}{E[Y^q]^{1/q}}\right)^q\right]}{q}$$

$$= \frac{\frac{E[X^p]}{E[X^p]}}{p} + \frac{\frac{E[Y^q]}{E[Y^q]}}{q}$$

$$= \frac{1}{p} + \frac{1}{q} = 1$$

$$* \Rightarrow \boxed{E[XY] \le E[X^p]^{1/p} E[Y^q]^{1/q}}$$

"Holder's Inequality"

- Let $0 < r < s$, $p = \frac{s}{r}$, $q = \frac{p}{p-1}$

$$q = \frac{p}{p-1} = \frac{s/r}{s/r - 1} = \frac{s}{s-r} \Rightarrow \frac{1}{p} + \frac{1}{q} = \frac{r}{s} + \frac{s-r}{s} = 1$$

Let $X = |V|^r$, $Y = 1$

Using Holder's Inequality,

$$E[|V|^r] \leq E\left[(|V|^r)^{s/r}\right]^{r/s}$$
$$= E[|V|^s]^{r/s}$$

$\Rightarrow E[|V|^s]$ is finite

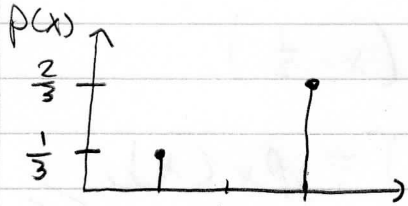
$\Rightarrow E[|V|^r]$ is finite.

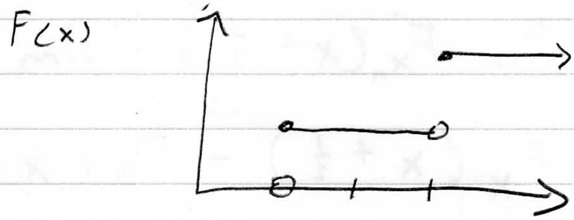
Convergence in Distribution

- $X_n \xrightarrow{d} X$ means $\lim_{n\to\infty} F_n(x) = F_X(x)$

- Consider $X_n \sim \begin{cases} \frac{1}{n+1} & \text{wp } \frac{1}{3} \\ 1 - \frac{1}{n+1} & \text{wp } \frac{2}{3} \end{cases}$

ex, $X_3 \sim \begin{cases} \frac{1}{4} & \text{wp } \frac{1}{3} \\ \frac{3}{4} & \text{wp } \frac{2}{3} \end{cases}$





$\lim p_n(x) = \text{Bern}\left(\frac{2}{3}\right)$

- Theorem: For $Supp[X_n] \subset \mathbb{Z}$, $Supp[X] \subset \mathbb{Z}$, then PMF Convergence $\Longleftrightarrow$ CDF Convergence, (i.e $X_n \xrightarrow{d} X$).

- Proof

$(\Leftarrow)$ $\forall x \in \mathbb{Z}$, $p_{X_n}(x) = F_{X_n}(x + \frac{1}{2}) - F_{X_n}(x - \frac{1}{2})$

$= P(X_n \in [x \pm \frac{1}{2}]) = P(X_n = x)$.

$\lim p_{X_n}(x) = \lim [F_{X_n}(x + \frac{1}{2}) - F_{X_n}(x - \frac{1}{2})]$

$= \lim F_{X_n}(x + \frac{1}{2}) - \lim F_{X_n}(x - \frac{1}{2})$

$= F_X(x + \frac{1}{2}) - F_X(x - \frac{1}{2}) = p_X(x)$. //

$(\Rightarrow)$ $\lim F_{X_n}(x) = \lim P(X_n \le x) = \lim \sum_{y=-\infty}^{x} p_{X_n}(x)$

$= \sum_{y=-\infty}^{x} \lim p_{X_n}(y) = \sum_{y=-\infty}^{x} p_X(y)$

$= P(X \le x) = F_X(x)$.

$\square$

- $X_n \sim Bin(n, \lambda/n)$

  $X_n \xrightarrow{d} X \implies X \sim Poisson(\lambda)$

- PDF Convergence $\Rightarrow$ CDF Convergence, but converse <u>does not</u> hold.

<u>Counter example</u>

ex. $X_n \sim U(0, 1/n) = n \mathbb{1}_{x \in [0, \frac{1}{n}]} = f_{X_n}(x)$

$X_n \xrightarrow{d} Deg(0)$

$\lim f_{X_n}(x) = \begin{cases} \infty & \text{if } x = 0 \\ 0 & \text{o.w} \end{cases}$ (Not a valid PD

$F_{X_n} = \begin{cases} 0 & \text{if } x < 0 \\ nx & \text{if } x \in [0, 1/n] \\ 1 & \text{if } x \geq 1/n \end{cases}$

$\Rightarrow F_X(X) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$

$\Rightarrow X \sim Deg(0)$

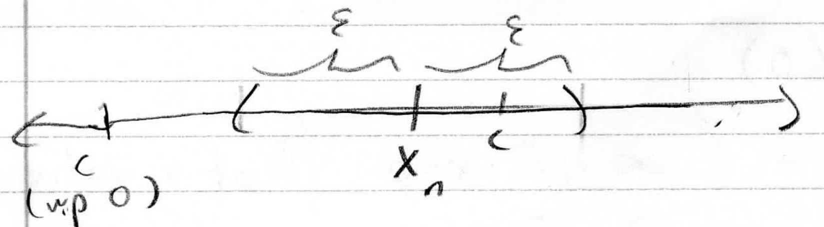
Convergence in Probability

Convergence to constants / degenerate distr. only.

We say $X_n \xrightarrow{P} c$, constant $c \in \mathbb{R}$ if

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

$$\iff \lim P(|X_n - c| \leq \varepsilon) = 1.$$



ex. Let $X_n \sim U(-\frac{1}{n}, \frac{1}{n}) = \frac{n}{x} \mathbb{1}_{x \in [-\frac{1}{n}, \frac{1}{n}]}$