Wednesday December 2^{nd} 2020 - Lecture 22

We'll continue with two more inequalities. Consider rv's X and Y with finite means and variances, $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$ and let $W = (X - cY)^2$ for some constant c. Note that W is non-negative by construction.

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

$$\Rightarrow E[X^2] - 2c\,E[XY] + c^2E[Y^2] \geq 0 \qquad \text{let } c = \frac{E[XY]}{E[Y^2]} \in \mathbb{R}$$

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

Multiply by $E[Y^2]$

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

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

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

if X, Y non neg $\Rightarrow E[XY] \leq \sqrt{E[X^2]E[Y^2]}$ these called the "Cauchy Schwartz inequalities" and they're famous

Recall
$$Cov[X,Y] = E[XY] - E[X]E[Y]$$

$$Corr[X,Y] = \frac{Cov[X,Y]}{SD[X]SD[Y]} \in [-1,1]$$

Correlation of X and Y a unitless metric. We now prove that its range is always -1 to 1:

Let $Z_X = \frac{X - \mu_X}{\sigma_X}$ and $Z_Y = \frac{Y - \mu_Y}{\sigma_Y} \Rightarrow E[Z_X] = E[Z_Y] = 0$

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

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

$|E[Z_X Z_Y]| \leq \sqrt{E[Z_X^2]E[Z_Y^2]} = \sqrt{(1)(1)} = 1 \Rightarrow E[Z_X Z_Y] \in [-1, 1]$

$$Corr[X,Y] = \frac{E[XY] - \mu_X \mu_Y}{\sigma_X \sigma_Y} = \frac{E[(\sigma_X Z_X + \mu_X)(\sigma_Y Z_Y + \mu_Y)] - \mu_X \mu_Y}{\sigma_X \sigma_Y}$$

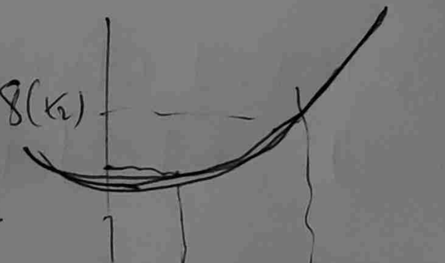
$$= \frac{\sigma_X \sigma_Y E[Z_X Z_Y] + \mu_X \sigma_Y E[Z_Y] + \mu_Y \sigma_X E[Z_X] + \mu_X \mu_Y - \mu_X \mu_Y}{\sigma_X \sigma_Y}$$

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

Definition: g is a "Convex function" on an Interval, I a connected subset of the reals if for all $\{x_1, x_2, \ldots\} \subset I$ and for all $w_1, w_2, \ldots \in (0,1)$ s.t the sum of the w_i's $= 1$

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

$\Rightarrow g(\sum w_i x_i) \leq \sum w_i g(x_i)$



Thm: if g is twice differentiable and $g''(x) \geq 0$ for all x in I, then g is Convex on I.

for all w_1, w_2
s.t $w_1 + w_2 = 1$
$w_1 x_1 + w_2 x_2$
$w_1 g(x_1) + w_2 g(x_2)$

Consider a discrete r.v X with PMF p

$E[X] = \sum_{x \in Supp[X]} p(x) x$ and $Supp[X] = \{x_1, x_2, \ldots\}$ Let $w_i = p(x_i)$

$\Rightarrow \sum w_i = 1$

and a convex function g. then just using the definition of convexity, we get the following inequality: $E[g(x)] = \sum p(x_i)\, g(x_i)$

$$\Rightarrow g(E[X]) \leq E[g(x)]$$

Jensens Inequality.

Types of convergence of random variables. we begin with reviewing convergence in distribution. Consider a sequence

$X_2 \ldots$ denoted X_n

$X_n \longrightarrow X$ is defined as $\lim_{n\to\infty} F_{X_n}(x) = F_x(x)\ \forall x$

let $X_n \sim \begin{cases} \frac{1}{n+1} & \text{wp } \frac{1}{3} \\ 1-\frac{1}{n+1} & \text{wp } \frac{2}{3} \end{cases}$ e.g. $X_3 \sim \begin{cases} \frac{1}{4} & \text{wp } \frac{1}{3} \\ \frac{3}{4} & \text{wp } \frac{2}{3} \end{cases}$

It seems plausible that PMF convergence and CDF convergence are equivalent. thm: $\operatorname{Supp}[X_n] \subset \mathbb{Z}$ and $\operatorname{Supp}[X] \subset \mathbb{Z}$ then they are equivalent.

Pf: CDF convergence $\Rightarrow$ PMF convergence (for discrete sequences). $p_{X_n}(x) = F_{X_n}(x+\frac{1}{2}) - F_{X_n}(x-\frac{1}{2})\ \forall x \in \mathbb{Z}$

$\lim_{n\to\infty} p_{X_n}(x) = \lim F_{X_n}(x+\frac{1}{2}) - \lim F_{X}(x-\frac{1}{2}) = F_x(x+\frac{1}{2}) - F_X(x-\frac{1}{2}) = p_X(x)$

Pf: PMF Convergence $\Rightarrow$ CDF Convergence

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

$$\lim F_{X_n}(x) = \lim \sum_{y=-\infty}^{x} P_{X_n}(y) = \sum_{y=-\infty}^{x} \lim P_{X_n}(x) = \sum_{y=-\infty}^{x} P_X(x) = P(X \le x) = F_X$$

$X_n \sim \text{Binom}(n, \frac{\lambda}{n})$ show $X_n \xrightarrow{d} X \sim \text{Poisson}(\lambda)$ Hw

---

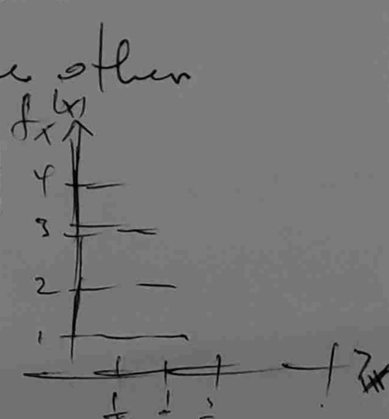
Let $c \in \mathbb{R}$ and let $X_n \xrightarrow{d} c$ be defined as ~~$X_n \xrightarrow{d} X_n$ degenerate~~

$X_n \xrightarrow{d} X \sim \text{Deg}(c)$

which means by definition of conv. in distr.

$$\forall x \ \lim F_{X_n}(x) = \begin{cases} 1 & \text{if } x \ge c \\ 0 & \text{o.t.} \end{cases}$$

For Continuous rv's is pdf Convergence equivalent to CDF Convergence unconditionally? No. PDF convergence $\Rightarrow$ CDF Convergence. ~~Convergence~~ here's a counterexample to the other direction: $X_n \sim U(0, \frac{1}{n}) = n \mathbb{1}_{x \in [0, \frac{1}{n}]}$



$\lim f_{X_n}(x) = \infty$ no PDF

Hw. show $\lim F_{X_n}(x) = \begin{cases} 1 & x \ge 0 \\ 0 & x < 0 \end{cases} \Rightarrow X \sim \text{deg}(0)$

Convergence in probability. But only to a constant c. Rv's can converge in probability to other rv's, but we just won't study it. A sequence of rv's X_1, X_2. (denoted X_n)