Lecture 22

Consider rv's X and Y with finite means and variances, $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$ and let $W = (X - cY)^2$ where c is a real constant.

Note: W is nonnegative

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

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

Choose $c = \frac{E[XY]}{E[Y^2]} \in \mathbb{R}$ $\rightarrow E[X^2] - 2\frac{E[XY]}{E[Y^2]}E[XY] + \frac{E[XY]^2}{E[Y^2]^2}E[Y^2] \geq 0$

Multiply by $E[Y^2]$ $\rightarrow E[X^2]E[Y^2] - 2E[XY]^2 + E[XY] \geq 0$

$\rightarrow E[XY]^2 \leq E[X^2]E[Y^2]$

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

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

These are rel. famous; they're called the Cauchy - Schwartz inequalities. We'll use it to prove a basic fact useful in stat.

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

$Corr[X,Y] := \frac{Cov[X,Y]}{SD[X]SD[Y]}$ "correlation btw X and Y"

let $Z_X = \frac{X-\mu_X}{\sigma_X}$ and $Z_Y = \frac{Y-\mu_Y}{\sigma_Y}$ $\Rightarrow E[Z_X] = E[Z_Y]$,
$SD[Z_X] = SD[Z_Y]$
$= E[Z_X^2] = E[Z_Y^2]$

$|E[Z_XZ_Y]| \leq \sqrt{E[Z_X^2]E[Z_Y^2]} = \sqrt{1 \cdot 1} = 1 \rightarrow E[Z_XZ_Y] \in [-1,1]$

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

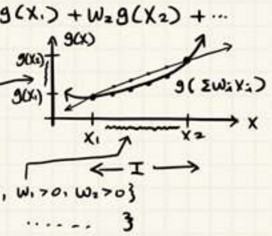
$= \frac{\cancel{\sigma_X\sigma_Y} E[Z_XZ_Y] + \sigma_X\mu_Y \cancel{E[Z_X]}^{0} + \sigma_Y\sigma_X \cancel{E[Z_Y]}^{0} + \cancel{\mu_X\mu_Y - \mu_X\mu_Y}}{\cancel{\sigma_X\sigma_Y}}$

$= E[Z_XZ_Y] \in [-1,1]$

Def: g is a "convex function" on an interval I (a subset of reals) if for all $x_1, x_2, \cdots \in I$ and all $w_1, w_2, \cdots \in (0,1)$ s.t. $\Sigma w_i = 1$ AKA the "weights"

$g(w_1x_1 + w_2x_2 + \cdots) \leq w_1g(x_1) + w_2g(x_2) + \cdots$

$g(\Sigma w_ix_i) \leq \Sigma w_ig(x_i)$



$w_1 + w_2 = 1$

$\{w_1x_1 + w_2x_2 : w_1 + w_2 = 1, w_1 > 0, w_2 > 0\}$

$\{w_1g(x_1) + w_2g(x_2) : \ \cdots\cdots \}$

Let g be a convex function and X be a discrete r.v. If discrete, we know $Supp[X] = \{x_1, x_2, \cdots\}$ and sum $p(x_i) = 1$ (the PMF)

Thus, we can call the PMF values, the weights i.e. $w_i := p(x_i)$

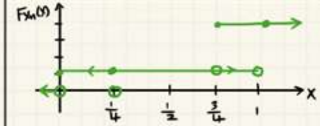
$E[X] = \Sigma x_ip(x_i) = \Sigma g(x_i)p(x_i) = E[g(X)]$

Jensen's Inequality

Convergence of rv's. We will study three different types. First, let's review "Convergence in distribution" We say a sequence of rv's $X_1, X_2, \cdots$ denoted X_n converges in distribution to X denoted: $X_n \xrightarrow{d} X$ means by def'n that the limiting CDF is X's CDF:

$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \quad \forall x$

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



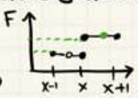
$\Rightarrow X \sim \begin{cases} 0 & wp \ \frac{1}{3} \\ 1 & wp \ \frac{2}{3} \end{cases}$

Conjecture: PMF convergence and CDF convergence are equivalent. This is not true in general. But here's a situation where it is true:

If $Supp[X_n]$ be a subset of $\mathbb{Z}$, the integers and let $Supp[X]$ also be a subset of $\mathbb{Z}$, the integer. Let's prove it.

pf: CDF convergence implies PMF convergence:

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



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

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

pf: PMF convergence implies CDF convergence:

$F_{X_n}(x) := P(X_n \leq x) = \sum_{y=-\infty}^{x} p_{X_n}(y)$

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

$= \sum_{y=-\infty}^{x} p_X(y) = P(X \leq x) = F_X(x)$

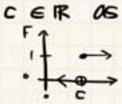
PDF convergence always implies CDF convergence but not vice versa. Here's a counterexample:

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

$\lim f_{X_n}(x) = \infty$! Not a PDF!

Define $X_n \xrightarrow{d} c$, $c \in \mathbb{R}$ as $X_n \xrightarrow{d} X \sim Deg(c)$

$\lim F_{X_n} = \begin{cases} 0 & \text{if } x < c \\ 1 & \text{if } x \geq c \end{cases}$



Convergence in Prob. to a constant.

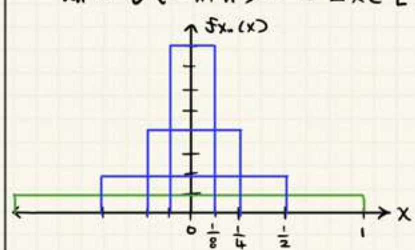
For a sequence of rv's $X_1, X_2, \cdots$ denoted X_n, X_n converges in prob to a constant c,

$X_n \xrightarrow{p} c$ is defined to be:

$\forall \varepsilon > 0 \quad P(|X_n - c| \geq \varepsilon) = 0$

or $\forall \varepsilon > 0 \quad \lim_{n\to\infty} P(|X_n - c| < \varepsilon) = 1$

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



$\varepsilon = 0.0001$

$n \leq 100$

$X_n \sim U(-.01, .01)$

$P(|X_n - 0| \leq 0.0001)$

$= P(X_n \in [-0.0001, 0.0001])$

$= \frac{2}{100} \cdot \frac{2}{100} \neq 1$

↑

$n = 1000$

$X_n \sim U(-.0001, .0001)$

$P(X_n \in [-.0001, .0001])$

$= 1$