CSE 107 11-14-24

- Mid2 : Thur 11/21

Sums & Convolutions

Let X, Y be independent, discrete r.v.s, and let

$$Z = X + Y$$

Then

$$p_Z(z) = P(Z = z)$$
$$= P(X + Y = z)$$

$$= \sum_{\{(x,y) \mid x+y=z\}} P(X=x, Y=y)$$

$$= \sum_{\{(x, z-x) \mid x \in \text{range}(X)\}} P(X=x, Y=z-x)$$

$$= \sum_{x} P(X=x) \cdot P(Y=z-x) \quad \left\{ \begin{array}{l} \text{by} \\ \text{ind.} \end{array} \right.$$

$$= \sum_{x} P_X(x) \cdot P_Y(z-x)$$

We call this the <u>convolution</u> <u>Product</u> of P_X and P_Y.

<u>notation</u>

$$(p_X * p_Y)(z) = \sum_x p_X(x) \cdot p_Y(z-x)$$

Thus: $p_Z = p_{X+Y} = p_X * p_Y$

<u>note</u>: $p_X * p_Y = p_Y * p_X$

Now let X, Y be independent (jointly) continuous r.v.s, and

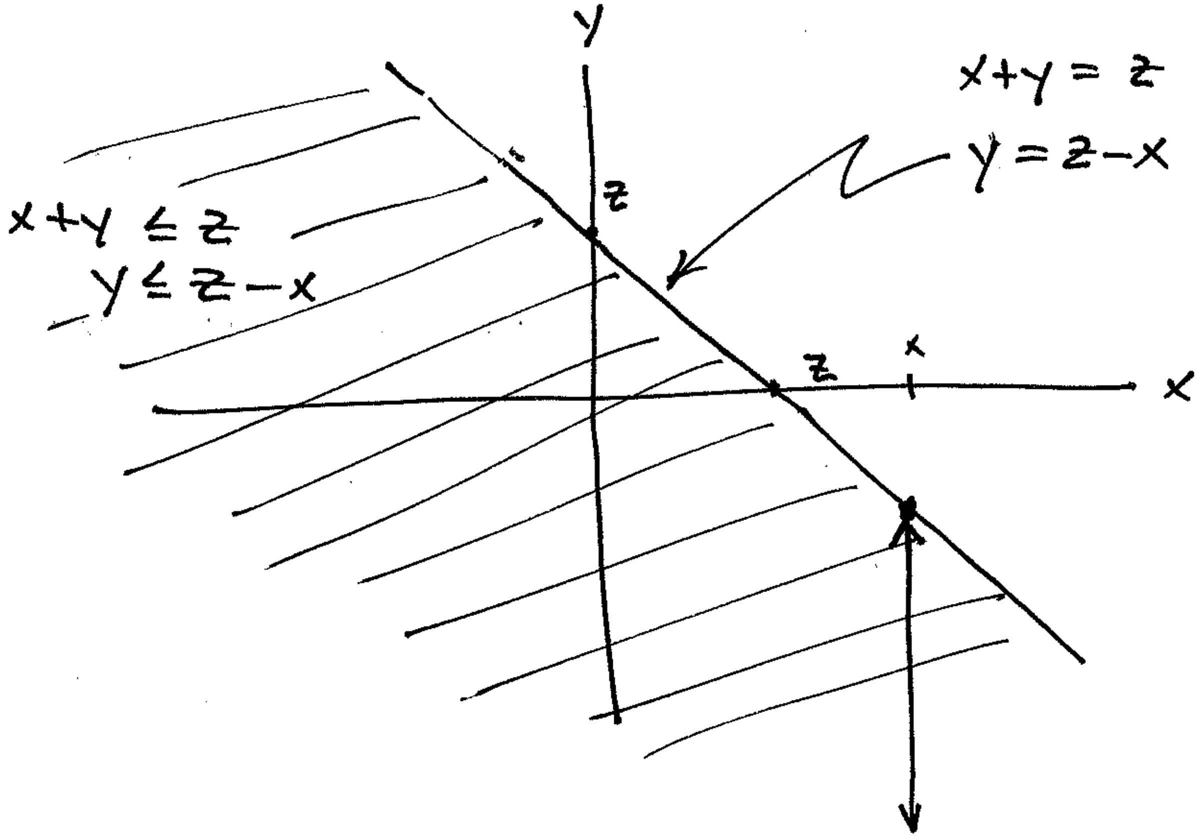
$$Z = X + Y$$

Then

$$F_Z(z) = P(Z \leq z)$$

$$= P(X + Y \leq z)$$

$$= \iint_{\{(x,y) \mid x+y \leq z\}} f_{X,Y}(x,y)\, dy\, dx$$



so

$$F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f_X(x) \cdot f_Y(y)\, dy\, dx \quad \Big\} \text{ by ind.}$$

$$= \int_{-\infty}^{\infty} f_X(x) \cdot \left(\int_{-\infty}^{z-x} f_Y(y)\, dy \right) dx$$

$$= \int_{-\infty}^{\infty} f_X(x) \cdot F_Y(z-x)\, dx$$

differentiate w.r.t. z !

$$f_Z(z) = \frac{d}{dz} \int_{-\infty}^{\infty} f_X(x) F_Y(z-x)\, dx$$

$$= \int_{-\infty}^{\infty} f_X(x) \cdot \frac{d}{dz} F_Y(z-x)\, dx$$

$$= \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(z-x)\, dx$$

We call this operation the <u>convolution</u> <u>product</u> of f_X and f_Y, i.e.

$$(f_X * f_Y)(z) = \int_{-\infty} f_X(x) f_Y(z-x)\, dx$$

so

$$f_{X+Y} = f_X * f_Y = f_Y * f_X$$

Ex.

let X, Y be ind. cont. unif. on $[a, b]$, so

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

likewise for $f_Y(y)$. Let

$$Z = X + Y$$

Then

$$f_Z(z) = (f_X * f_Y)(z)$$

$$= \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(z-x)\, dx$$

The integ-and is non-zero and eq. to $\frac{1}{(b-a)^2}$ iff both

$$a \leq x \leq b \quad \text{and} \quad a \leq z-x \leq b$$

i.e.

$$a \leq x \leq b \quad \text{and} \quad z-b \leq x \leq z-a$$

(and 0 otherwise)

i.e. iff

$$\max(a, z-b) \le x \le \min(b, z-a).$$

Thus

$$f_Z(z) = \int_{\max(a,z-b)}^{\min(b,z-a)} \frac{1}{(b-a)^2}\, dx$$

$$= \frac{\min(b, z-a) - \max(a, z-b)}{(b-a)^2}$$

The interval of integration is empty if either

$$z-a < a \iff z < 2a$$

or

$$z-b > b \iff z > 2b$$

so to be non-zero, we need

$$2a \le z \le 2b$$

Hence

$$f_Z(z) = \begin{cases} \dfrac{\min(b, z-a) - \max(a, z-b)}{(b-a)^2} & \text{if } 2a \le z \le 2b \\ 0 & \text{otherwise} \end{cases}$$

<u>Ex.</u> let X, Y be indep. normal r.v.s with

mean: μ_X, μ_Y

Variance: σ_X^2, σ_Y^2

Thus

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma_X} \cdot e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}}$$

similarly for Y. Let

$$Z = X + Y$$

Then

$$f_Z(z) = (f_X * f_Y)(z)$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\cdot\sigma_X} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} \cdot \frac{1}{\sqrt{2\pi}\,\sigma_Y} e^{-\frac{(z-x-\mu_y)^2}{2\sigma_y^2}} \, dx$$

$= \text{-----} \; exercise \; \text{-----}$

$$= \frac{1}{\sqrt{2\pi}\cdot\sqrt{\sigma_x^2+\sigma_y^2}}\cdot e^{-\frac{(z-(\mu_x+\mu_y))^2}{2(\sigma_x^2+\sigma_y^2)}}$$

Thus Z is normal with

mean: $\mu_x + \mu_y$

Variance: $\sigma_X^2 + \sigma_Y^2$

It follows that if X, Y are ind. normal, then so is

$$aX + bY.$$

4.2 Covariance & Correlation

<u>Defn</u>

The covariance of r.v.s X, Y is

$$\mathrm{cov}(X,Y) = E\big[(X - E[X])(Y - E[Y])\big]$$

observe

$$\mathrm{cov}(X,X) = \mathrm{var}(X)$$

We say X, Y are <u>uncorrelated</u> iff

$$\mathrm{cov}(X,Y) = 0$$

- If $Cov(X,Y) > 0$: X, Y <u>positively correlated</u>

- If $Cov(X,Y) < 0$: X, Y <u>negatively correlated</u>

observe:

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

$$= E[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y]$$

$$= E[XY] - \mu_X E[Y] - \mu_Y E[X] + \mu_X \mu_Y$$

$$= E[XY] - \mu_X \mu_Y - \cancel{\mu_X \mu_Y} + \cancel{\mu_X \mu_Y}$$

$$= E[XY] - E[X]E[Y].$$

<u>note</u>:

$$X, Y \text{ independent} \Rightarrow E[XY] = E[X]\cdot E[Y]$$

$$\Rightarrow \operatorname{cov}(X, Y) = 0$$

$$\Rightarrow X, Y \text{ uncorrelated.}$$

<u>Warning</u>: the converse is false!

However, the contrapositive is true

$$\boxed{\begin{array}{l}X, Y \text{ either} \\ \text{pos. or neg.} \\ \text{correlated}\end{array}} \Rightarrow \boxed{\begin{array}{l}X, Y \text{ are} \\ \text{not indep.}\end{array}}$$

__Defn__

The __correlation coefficient__ of

X, Y is

$$\rho(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\cdot \mathrm{Var}(Y)}}$$

this is a normalized version of $\mathrm{cov}(\cdot,\cdot)$. It can be shown

$$\underbrace{-1}_{\substack{\uparrow \\ \text{Perfectly} \\ \text{neg. corr.}}} \le \rho(X,Y) \le \underbrace{1}_{\substack{\uparrow \\ \text{Perfectly} \\ \text{Pos. corr.}}}$$