Recall that if X_i's are indep, then $F_{X_1,\ldots,X_d} = F_{X_1} \times \cdots \times F_{X_d}$
$i=1,\ldots,d$

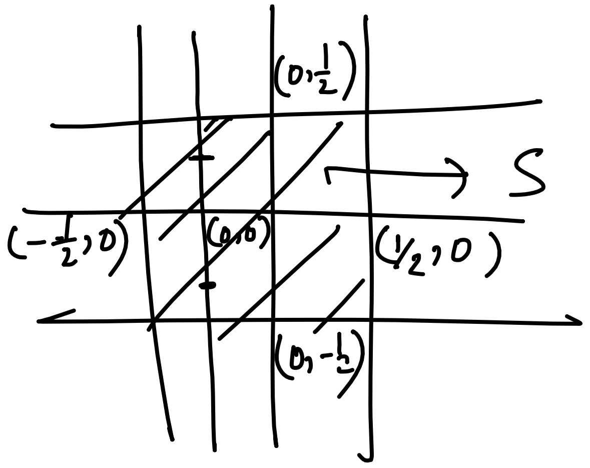
and $f_{X_1,\ldots,X_d} = f_{X_1} \times \cdots \times f_{X_d}$.

Remark :- If $(X,Y) \sim N_2(0,0,1,1,\rho)$, then $\rho = 0 \Rightarrow X \perp\!\!\!\perp Y$

Example :-

$$S = \left\{ (x,y) : -\frac{1}{2} \le x \le \frac{1}{2},\ -\frac{1}{2} \le y \le \frac{1}{2} \right\}.$$

$(X,Y) \sim$ Uniform over S.



The joint density of (X,Y) is

$$f_{X,Y}(x,y) = 1 \quad \text{if } (x,y) \in S$$
$$= 0, \quad \text{o.w.}$$

Marginal p.d.f :-

$$f_X(x) = \int_{-1/2}^{1/2} 1\, dy = 1 \quad \text{if } x \in \left[\frac{1}{2}, \frac{1}{2}\right].$$

Similarly, $f_Y(y) = \int_{-1/2}^{1/2} 1\, dx = 1$ if $y \in \left[\frac{1}{2}, \frac{1}{2}\right]$.

Observe that $\forall x, y$.
$f_{X,Y}(x,y) = f_X(x) \times f_Y(y)$
Hence, here $X \perp\!\!\!\perp Y$.

Conditional distⁿ :- (Discrete case)

X : Discrete Random variable.

Y : Discrete Random variable.

Conditional p.m.f. of X given Y is defined as

$$P[X = x_i \mid Y = y_j] = \frac{P[X = x_i, Y = y_j]}{P[Y = y_j]}$$

Example :- Recall that earlier example

Experiment :- A coin is tossing thrice.

X : No. of heads in the first toss.

Y : Total no. of heads.

$x \backslash y$	0	1	2	3
0	1/8	2/8	1/8	0
1	0	1/8	2/8	1/8

Want to compute $P[X = 0 \mid Y = 1]$

$$= \frac{P[X = 0, Y = 1]}{P[Y = 1]} = \frac{2/8}{3/8} = \frac{2}{3}.$$

Continuous case:—

(X, Y): Bivariate Random vector.

$\hookrightarrow f_{X,Y}$: Joint density of (X, Y).

f_X: Marginal density of X.

f_Y: Marginal density of Y.

Now, the conditional density of $Y|X$ is defined as

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} \quad \text{———} \quad (*)$$

Remark:— $X \perp\!\!\!\perp Y \iff f_{X,Y}(x,y) = f_X(x)\, f_Y(y) \quad \forall x, y \quad (**)$

Also, from $(*)$, we have $f_{X,Y}(x,y) = f_X(x)\, f_{Y|X}(y|x) \quad \forall x, y. \quad (***)$

Hence $(**)$ & $(***)$ $\Rightarrow f_{Y|X}(y|x) = f_Y(y)$ if $X \perp\!\!\!\perp Y$.

Example:- Suppose, $f_{X,Y}(x,y) = \lambda^2 e^{-\lambda y}$, $0 \le x \le y < \infty$

$$f_X(x) = \int_x^\infty f_{X,Y}(x,y)\,dy = \int_x^\infty \lambda^2 e^{-\lambda y}\,dy$$

Marginal density of X

$$= \lambda \cdot e^{-\lambda x}, \quad x \ge 0.$$

$$f_Y(y) = \int_0^y f_{X,Y}(x,y)\,dx = \int_0^y \lambda^2 e^{-\lambda y}\,dx.$$

Marginal density of Y

$$= \lambda^2 y \cdot e^{-\lambda y}.$$

Conditional density of $X|Y$ is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\lambda^2 e^{-\lambda y}}{\lambda^2 y \cdot e^{-\lambda y}} = \frac{1}{y} \text{ if } 0 \le x \le y.$$

Example:- Suppose $(X, Y) \sim N_2(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$.

Already, we have seen $X \sim N(\mu_x, \sigma_x^2)$

$Y \sim N(\mu_y, \sigma_y^2)$.

Want to derive $f_{\underbrace{Y|X=x}_{\text{Writing is NOT perfect}}}(y|x)$ [Conditional density of Y given X].

$$f_{Y|X=x} = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{\text{Joint density of } N_2(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)}{\text{Marginal density of } N(\mu_x, \sigma_x^2)}$$

Check

$Z \sim N(\mu_Z, \sigma_Z^2)$

$f_Z(z) = \frac{1}{\sqrt{2\pi}\sigma_Z} e^{-\frac{1}{2}\left(\frac{z-\mu_Z}{\sigma_Z}\right)^2}$

$$= \frac{1}{\sqrt{2\pi}\sigma_y\sqrt{1-\rho^2}} e^{-\left[\frac{1}{2\sigma_y^2(1-\rho^2)}\left\{x - \left(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)\right)\right\}^2\right]}.$$

$$\mapsto N\left(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x),\ \sigma_y^2(1-\rho^2)\right)$$

Hence,

$$E[Y \mid X = x] = \mu_y + \rho \frac{\sigma_y}{\sigma_x}(x - \mu_x)$$

$$Var[Y \mid X = x] = \sigma_y^2 (1 - \rho^2).$$

This is the definition of linear regression.

$Y = \beta_0 + \beta_1 x + \varepsilon$

$E[Y|x] = \beta_0 + \beta_1 x + \underbrace{E[\varepsilon|x]}_{0}$
$= \beta_0 + \beta_1 x$

$$\mu_y + \rho \frac{\sigma_y}{\sigma_x} x - \mu_x \rho \frac{\sigma_y}{\sigma_x}$$

$$= \beta_0 + \beta_1 x, \text{ where}$$

$$\beta_0 = \mu_y - \mu_x \rho \frac{\sigma_y}{\sigma_x}$$

$$\beta_1 = \rho \frac{\sigma_y}{\sigma_x}.$$

Remark:-

Instead of defining model in simple linear regression, one can define in the following way :-

$$(X, Y) \sim N_2(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho).$$

Then $E[Y|X = x]$ can be consider as the simple linear regression model.

Expectation & Product Moments :—

$\underset{\sim}{X}$: $(X_1, \dots, X_d)$: Random vector of d-dimension.

$g(X_1, \dots, X_d) : \mathbb{R}^d \to \mathbb{R}$.

Then $E[g(X_1, \dots, X_d)] = \underbrace{\sum_{x_1, \dots, x_d} g(x_1, \dots, x_d)\, P[X_1 = x_1, \dots, X_d = x_d]}_{\mapsto \text{ Discrete case}}$

p.m.f.

for continuous case,

$$E[g(X_1, \dots, X_d)] = \int \cdots \int g(x_1, \dots, x_d)\, f_{X_1, \dots, X_d}(x_1, \dots, x_d)\, dx_1\, dx_2 \cdots dx_d$$

p.d.f.

provided $E|g(X_1, \dots, X_d)|$ exists.

$$\parallel$$

$$\int |g(x_1, \dots, x_d)|\, f_{X_1, \dots, X_d}(x_1, \dots, x_d)\, dx_1 \cdots dx_d < \infty.$$

A few special case :-