Function of a Random Variable

Given a function (measurable) $g : \mathbb{R} \to \mathbb{R}$

$Y = g(X)$ is also a random variable.

Given the distribution of X, we can determine the distribution of $Y = g(X)$.

The cdf of Y :

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$$

$$= P(X \in g^{-1}(-\infty, y])$$

<u>Examples</u>: Let X be a r.v. with cdf $F_X(x)$. Let $Y_1 = aX+b$ where a and b are real numbers. $a \neq 0$.

If $a > 0$, then

$$F_{Y_1}(y) = P(Y_1 \le y) = P(aX + b \le y)$$
$$= P\left(X \le \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right)$$

Let $a < 0$. Then

$$P(aX + b \le y) = P\left(X \ge \frac{y-b}{a}\right)$$
$$= 1 - P\left(X \le \frac{y-b}{a}\right) + P\left(X = \frac{y-b}{a}\right)$$
$$= 1 - F_X\left(\frac{y-b}{a}\right) + P\left(X = \frac{y-b}{a}\right)$$

2. $Y_2 = |X|$

$$F_{Y_2}(y) = P(|X| \le y) = 0 \text{ if } y < 0$$

For $y > 0$

$$F_{Y_2}(y) = P(-y \le X \le y)$$

$$= P(X \le y) - P(X \le -y) + P(X = -y)$$

$$= F_X(y) - F_X(-y) + P(X = -y)$$

3. $Y_3 = X^2$

$$F_{Y_3}(y) = P(X^2 \le y) = 0, \quad y < 0$$

For $y > 0$

$$F_{Y_3}(y) = P(-\sqrt{y} \le X \le \sqrt{y})$$

$$= P(X \le \sqrt{y}) - P(X \le -\sqrt{y}) + P(X = -\sqrt{y})$$

$$= F_X(\sqrt{y}) - F_X(-\sqrt{y}) + P(X = -\sqrt{y})$$

4. $Y_4 = \max(X, 0)$

$F_{Y_4}(y) = P(\max(X,0) \leq y) = 0$ if $y < 0$

$\underline{y > 0}$: $= \begin{cases} P(X \leq 0) & \text{if } y = 0 \\ P(X \leq 0) + P(0 < X \leq y) & \text{if } y > 0 \end{cases}$

So $F_{Y_4}(y) = \begin{cases} 0 & \text{if } y < 0 \\ F_X(y), & y \geq 0 \end{cases}$

Case of Discrete Random Variables

Let X be a discrete r.v. with pmf $p_X(x_i)$, $x_i \in \mathcal{X}$.

$Y = g(X)$

for $Y = y_j \rightarrow g^{-1}(y_j) \rightarrow x_{j_1}, \dots x_{j_r}$

r inverse images

$p_Y(y_j) = P(Y = y_j) = P(X = x_{j_1}) + \dots$

$$+ P(X = x_{j_r})$$

Example: Let $p_X(-2) = \frac{1}{5}$, $p_X(-1) = \frac{1}{6}$

$p_X(0) = \frac{1}{5}$, $p_X(1) = \frac{1}{15}$, $p_X(2) = \frac{11}{30}$

Let $Y = X^2 \rightarrow 0, 1, 4$

$$p_Y(0) = P(X^2 = 0) = P(X = 0) = \frac{1}{5}$$

$$p_Y(1) = P(X^2 = 1) = P(X = -1) + P(X = 1)$$

$$= \frac{7}{30}$$

$$p_Y(4) = P(X^2 = 4) = P(X = -2) + P(X = 2)$$

$$= \frac{17}{30}$$

If X is continuous, then we use a different approach.

<u>Theorem</u> : Let X be a continuous r.v. with density function $f_X(x)$. Let

$y = g(x)$ be a differentiable function for all x and either $g'(x) > 0 \ \forall x$ or $g'(x) < 0 \ \forall x$. Then $Y = g(X)$ is a continuous r.v. with pdf of Y as

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| ,$$

$$\alpha < y < \beta$$

where $\alpha = \min(g(-\infty), g(+\infty))$

and $\beta = \max\left(g(-\infty), g(+\infty)\right)$

<u>Proof</u>: Let $g'(x) > 0 \ \forall \ x$

Then g is strictly increasing and so it is one-to-one function and g^{-1} is strictly increasing i.e $\frac{d}{dy} g^{-1}(y) > 0 \ \forall y$

The cdf of Y

$$F_Y(y) = P(Y \leq y) = P\left(g(X) \leq y\right)$$

$$= P\left(X \leq g^{-1}(y)\right)$$

$$= F_X\left(g^{-1}(y)\right)$$

So the pdf of Y is

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \cdot \frac{d}{dy} g^{-1}(y)$$

$$= f_X\left(g^{-1}(y)\right) \left| \frac{d}{dy} g^{-1}(y) \right|$$

Let $g'(x) < 0 \ \forall \ x$. Then $g(x)$ is strictly decreasing and so $g^{-1}(y)$ is strictly decreasing and $\frac{d}{dy} g^{-1}(y) < 0 \ \forall \ y$

The cdf of Y

$$F_Y(y) = P(g(X) \le y) = P(X \ge g^{-1}(y))$$

$$= 1 - P(X \le g^{-1}(y)) + \underbrace{\boxed{P(X = g^{-1}(y))}}_{0}$$

$$= 1 - F_Y(g^{-1}(y))$$

The pdf of Y is then

$$f_Y(y) = - f_Y(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$$

$$= f_Y(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|.$$

<u>Examples</u>: 1. Let $X \sim N(\mu, \sigma^2)$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad x \in \mathbb{R},$$

$\mu \in \mathbb{R}, \sigma > 0$

Put $Y = e^X$. $\quad x = \log y$

$dx = \frac{1}{y} dy$

Then the pdf of Y is

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}\, y} \cdot e^{-\frac{1}{2\sigma^2}(\log y - \mu)^2}, \quad y > 0$$

This is log-normal distn.

2. Let $X \sim U(0,1)$

$Y = -\log_e X$, $x = e^{-y}$

$$\frac{dx}{dy} = -e^{-y}$$

$$f_X(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

The pdf of Y is

$$f_Y(y) = \begin{cases} e^{-y}, & 0 < y < \infty \end{cases}$$

$$\Bigg(0 \qquad , \qquad y > 0$$

This is exponential distⁿ

<u>Theorem</u> : Let X be a continuous r.v. with pdf $f_X(x)$. Let $Y = g(x)$ be a differentiable function (for almost all x) and assume that $g'(x)$ is continuous and non-zero at all but a finite number of values of x. Then for

every real y,

(a) $\exists$ real inverses $x_1(y), \ldots, x_n(y)$

$\ni \quad g(x_r(y)) = y, \quad g'(x_r(y)) \neq 0,$

$r = 1, \ldots n$

(b) $\nexists$ any $x \ni g(x) = y$,

Then Y is a continuous r.v. with

pdf

$$f_Y(y) = \begin{cases} \sum_{r=1}^{n} f(x_r(y)) \left| g'(x_r(y)) \right|^{-1} & \text{if (a)} \\ 0, & \text{if (b)} \end{cases}$$

<u>Examples</u>: Let $X \sim N(0,1)$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}$$

Let $Y = X^2$

$$x_1(y) = -\sqrt{y}, \qquad x_2(y) = +\sqrt{y}$$

$$-\frac{1}{2\sqrt{y}} \qquad \frac{1}{2\sqrt{y}}$$

$$f_Y(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} e^{-y/2} \cdot \dfrac{1}{2\sqrt{y}} \\ + \dfrac{1}{\sqrt{2\pi}} e^{-y/2} \cdot \dfrac{1}{2\sqrt{y}}, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

$$= \frac{1}{\sqrt{2\pi y}} e^{-y/2}, \qquad y > 0$$

$$= \frac{1}{2^{1/2}\,\Gamma\frac{1}{2}} y^{\frac{1}{2}-1} e^{-y/2}, \quad y>0$$

which is a Gamma$\left(\frac{1}{2}, \frac{1}{2}\right)$

dist^n.

2. $f_X(x) = \frac{1}{2} e^{-|x|}, \quad x \in R$

$Y = |X|, \qquad g_1(y) = -y, \quad g_2(y) = y$

$\frac{d g_1(y)}{dy} = -1, \qquad \frac{d g_2(y)}{dy} = 1$

$$f_Y(y) = \begin{cases} \frac{1}{2} e^{-y} \cdot 1 + \frac{1}{2} e^{-y} \cdot 1, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

$$= e^{-y}, \quad y > 0$$

Probability Integral Transform

Let X be a continuous r.v. with cdf $F_X(x)$. Let $U = F_X(X)$.

Then $U \sim U[0,1]$.

Conversely, let $U \sim U[0,1]$

and F be an absolutely continuous cdf. Then

$X = F^{-1}(U)$ is a continuous r.v. with cdf F.

Jointly Distributed Random Variables

Many times we are taking multiple measurements / outcomes in a single random experiment. This gives rise to jointly distributed r.v.s

For example, on a patient, we

can have. $X_1 \to$ pulse rate,

$X_2 \to$ temperature, $X_3 \to$ age.

$X_4 \to$ weight, $X_5 \to$ systolic BP

$X_6 \to$ dystolic BP

$\underline{X} = (X_1, X_2, \ldots X_6)$ is a

6-dimensional random vector

Performance of a student

$X_1 \rightarrow$ perf in subject A_1

$X_2 \rightarrow$ . . . . . A_2

$\vdots$

$X_6 \rightarrow$ . . . A_6

$$\underline{X} = (X_1 \ldots X_6) \ldots$$

So a k-dimensional r. vector is a measurable function

$\underline{X} : \Omega \to \mathbb{R}^k$

We first consider $k = 2$.

Suppose we have two-dimensional r.v. (X, Y). In general the prob. distn of (X, Y) is described by the joint cdf $F_{(X,Y)}(x, y)$

$$F_{X,Y}(x,y) = P(X \leq x, Y \leq y).$$

$F_{X,Y}(\cdot\,,\,\cdot)$ Satisfies the following

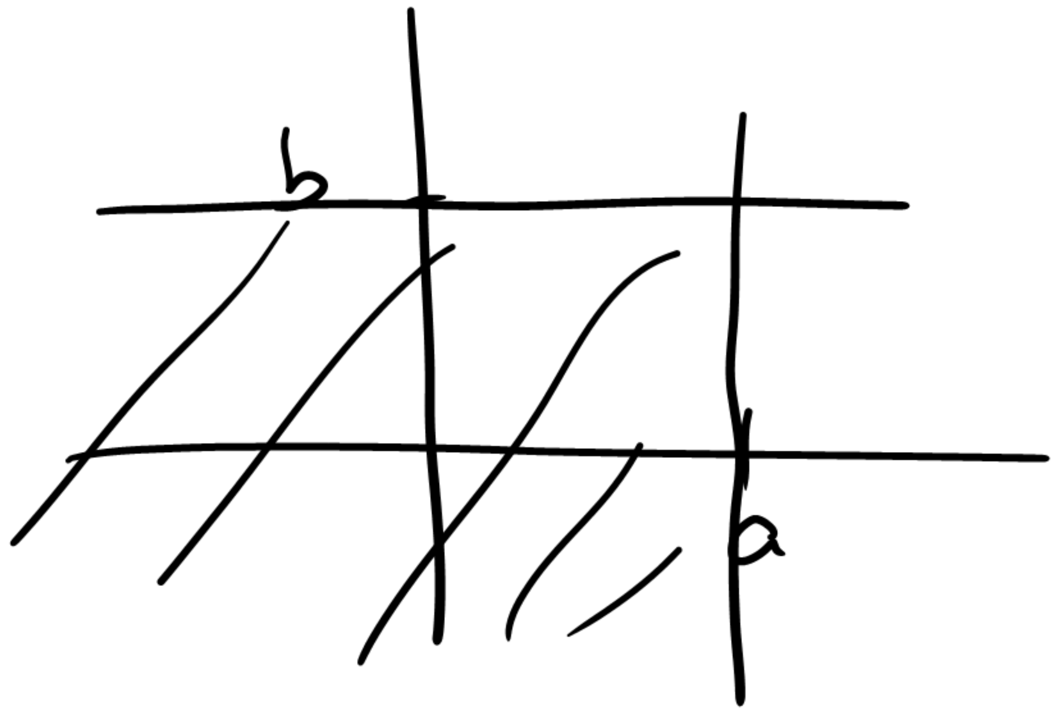
properties :

1. $\lim_{x \to -\infty} F_{X,Y}(x,y) = 0$

2. $\lim_{y \to -\infty} F_{X,Y}(x,y) = 0$

3. $\lim_{y \to +\infty} F_{X,Y}(x, y) = F_X(x)$

4. $\lim_{x \to +\infty} F_{X,Y}(x, y) = F_Y(y)$

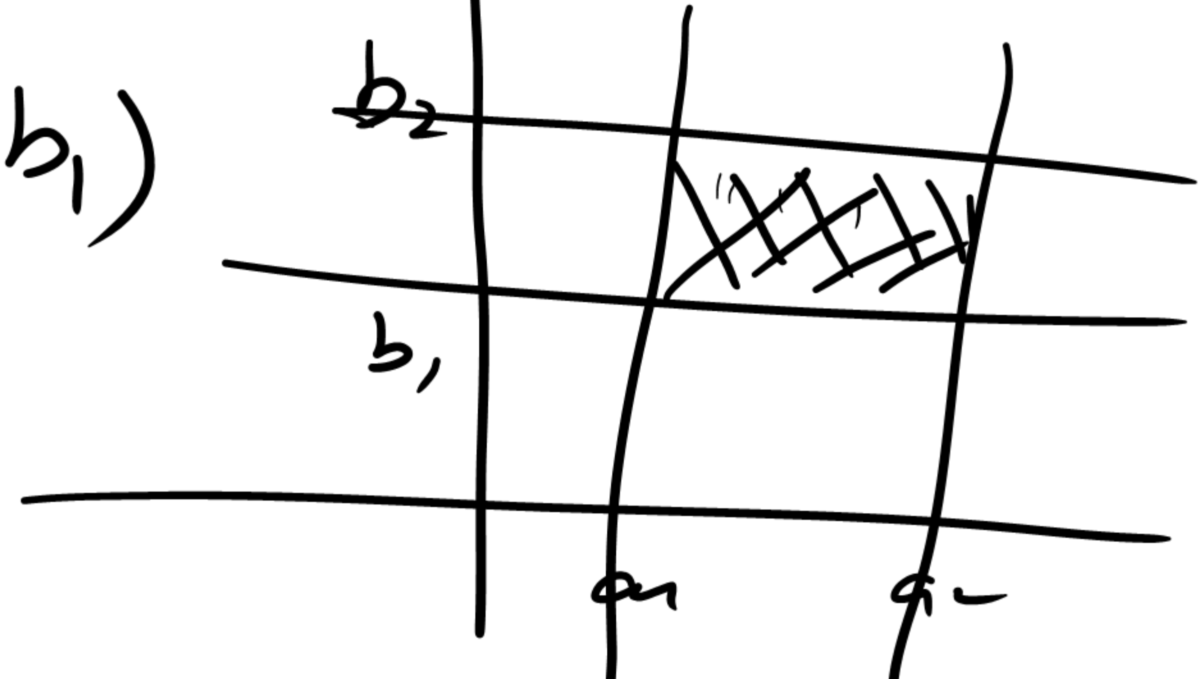


$$P(X \le a, Y \le b) = F_{X,Y}(a,b)$$

$$P(a_1 < X \le a_2, \ b_1 < Y \le b_2)$$

$$= F_{X,Y}(a_2, b_2) - F_{X,Y}(a_2, b_1) - F_{X,Y}(a_1, b_2)$$

$$+ F_{X,Y}(a_1, b_1)$$



Special Cases: When both (X, Y) are discrete $(X, Y) \in \mathcal{X} \times \mathcal{Y}$.

$\rightarrow$ finite or at most countable.

The joint pmf of (X, Y) is

$$p_{X,Y}(x_i, y_j) = P(X = x_i, Y = y_j)$$

$$p_{X,Y}(x_i, y_j) \geq 0 \; \forall (x_i, y_j)$$

$$\sum_{(x_i, y_j)} \sum p_{X,Y}(x_i, y_j) = 1.$$

<u>Example</u>: 10 cars $\rightarrow$ 3 Type II defe

5 are good   2 Type I defect

Suppose 2 cars are randomly selected.

$X \rightarrow$ no of cars with Type I defect

$Y \rightarrow$ no of cars with Type II defect

$X \to 0, 1, 2, \qquad Y \to 0, 1, 2$

$$X + Y \leq 2$$

$$p_{X,Y}(0,0) = \frac{\binom{5}{2}}{\binom{10}{2}} = \frac{2}{9}$$

$$p_{X,Y}(0,1) = \frac{\binom{5}{1}\binom{3}{1}}{\binom{10}{2}} = \frac{1}{3}$$

$$p_{X,Y}(0,2) = \frac{\binom{3}{2}}{{}^{10}C_2} = \frac{1}{15}$$

$$p_{X,Y}(1,0) = \frac{\binom{2}{1}\binom{5}{1}}{\binom{10}{2}} = \frac{2}{9}$$

$$p_{X,Y}(1,1) = \frac{\binom{2}{1}\binom{3}{1}}{\binom{10}{2}} = \frac{2}{15}$$

$$p_{X,Y}(2,0) = \frac{\binom{2}{2}}{\binom{10}{2}} = \frac{1}{45}$$

$x \backslash y$	0	1	2	$p_X(x_i)$
0	$10/45$	$15/45$	$3/45$	$28/45$
1	$10/45$	$6/45$	0	$16/45$
2	$1/45$	0	0	$1/45$
$p_Y(y_j)$	$21/45$	$21/45$	$3/45$	

$$P(X \leq 1, Y \leq 1) = p_{X,Y}(0,0) + p_{X,Y}(0,1) + p_{X,Y}(1,0) + p_{X,Y}(1,1)$$

$$= \frac{10}{45} + \frac{15}{45} + \frac{10}{45} + \frac{6}{45} = \frac{41}{45}$$

$$P(X=0) = P(X=0, Y=0) + P(X=0, Y=1)$$
$$+ P(X=0, Y=2)$$
$$= \frac{10}{45} + \frac{15}{45} + \frac{3}{45} = \frac{28}{45}$$

The marginal pmf of X is

$$p_X(x_i) = \sum_{y_j \in \mathcal{Y}} p_{X,Y}(x_i, y_j)$$

The marginal pmf of Y is

$$p_Y(y_j) = \sum_{x_i \in \mathcal{X}} p_{X,Y}(x_i, y_j)$$

The conditional pmf of X given $Y = y_j$

$$p_{X|Y=y_j}(x_i | y_j) = \frac{p_{X,Y}(x_i, y_j)}{p_Y(y_j)}, \quad x_i \in \mathcal{X}$$

The conditional pmf of Y given $X = x_i$

$$p_{Y|X=x_i}(y_j|x_i) = \frac{p_{X,Y}(x_i, y_j)}{p_X(x_i)}, \quad y_j \in \mathcal{Y}.$$

<u>Example</u>: The conditional pmf of Y given $X=0$

$$p_{Y|X=0}(0|0) = \frac{p_{X,Y}(0,0)}{p_X(0)} = \frac{10/45}{28/45} = \frac{10}{28}$$

$$p_{Y|X=0}(1|0) = \frac{p_{X,Y}(0,1)}{p_X(0)} = \frac{15}{28}$$

$$p_{Y/X=0}(2|0) = \frac{p_{X,Y}(0,2)}{p_X(0)} = \frac{3}{28}$$

$$p_{Y/X=2}(0|2) = \frac{p_{X,Y}(2,0)}{p_X(2)} = 1$$

$$p_{Y/X=1}(0|1) = \frac{p_{X,Y}(1,0)}{p_X(1)} = \frac{10/45}{16/45} = \frac{10}{16}$$

$$p_{Y|X=1}(1|1) = \frac{p_{X,Y}(1,1)}{p_X(1)} = \frac{6}{16}$$

Ex* Find conditional pmf of X given $Y=0$, $Y=1$ & $Y=2$.

When both (X,Y) are continuous, the joint pdf of (X,Y) is

$f_{X,Y}(x,y)$:

(i) $f_{X,Y}(x,y) \geq 0 \quad \forall\ (x,y) \in \mathbb{R}^2$

(ii) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx\, dy = 1$$

$\rightarrow dy\, dx$

(iii) If $B \subset \mathbb{R}^2$

$$P((X,Y) \in B) = \iint_B f_{X,Y}(x,y)\, dx\, dy$$

The marginal pdf of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$$

The marginal pdf of Y is

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$

The conditional pdf of X given $Y=y$

$$f_{X|Y=y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

The conditional pdf of Y given $X=x$

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

Examples: 1. Let (X,Y) be jointly distd. continuous r.v's with pdf

$$f_{X,Y}(x,y) = \begin{cases} 10xy^2, & 0<x<y<1 \\ 0, & \text{elsewhere} \end{cases}$$

The marginal pdf of X:

$$f_X(x) = \int_x^1 10xy^2\,dy = \frac{10}{3}x(1-x^3), \quad 0<x<1$$

$$f_X(x) = \begin{cases} \frac{10}{3}x(1-x^3), & 0<x<1 \\ 0, & \text{otherwise} \end{cases}$$

The marginal pdf of Y :

$$f_Y(y) = \int_0^y 10\, x y^2 \, dx = \begin{cases} 5y^4, & 0<y<1 \\ 0, & \text{otherwise} \end{cases}$$

The conditional pdf of X given $Y=y$

$$f_{X|Y=y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \frac{2x}{y^2}, & 0<x<y \\ 0, & \text{ew} \end{cases}$$

where $0<y<1$

The conditional pdf of Y given $X=x$

$$f_{Y|X=x}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} \dfrac{3y^2}{1-x^3}, & x<y<1 \\ 0, & ew \end{cases}$$

for $0 < x < 1$

$$P(X < \tfrac{1}{4}) = \int_0^{1/4} f_X(x)\,dx$$

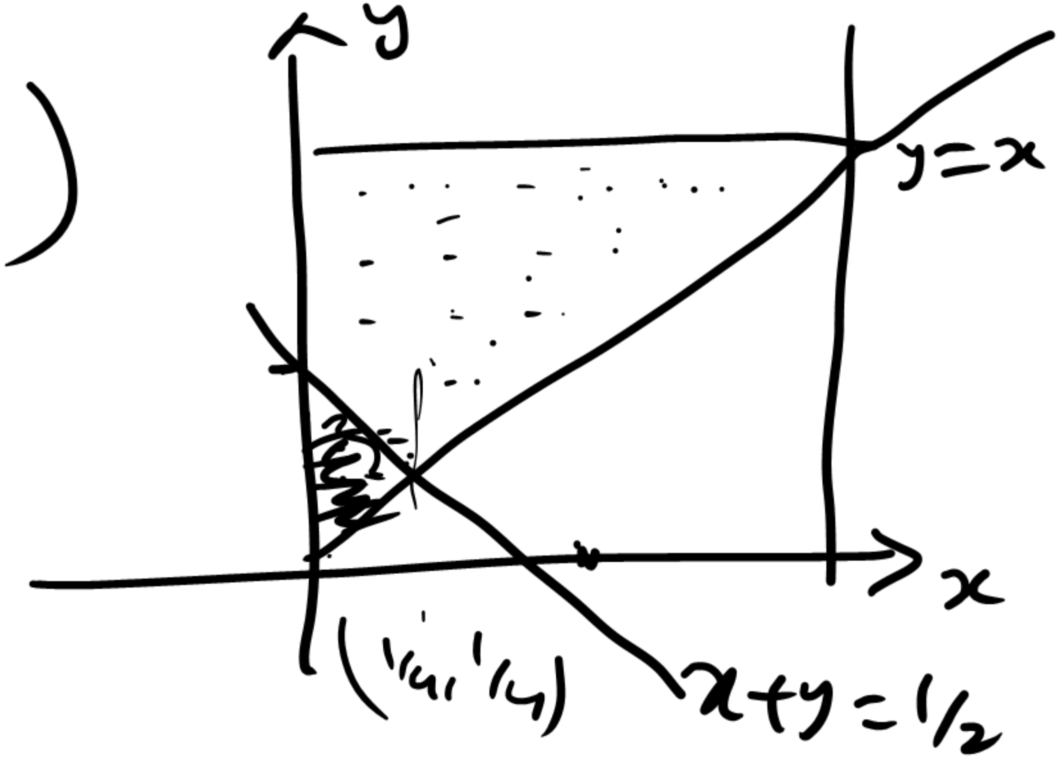
$$= \int_0^{1/4} \frac{10}{3} x(1-x^3)\,dx$$

$$= \frac{10}{3}\left[\frac{1}{32} - \frac{1}{5\cdot 4^5}\right] = \frac{53}{512}$$

$$P(Y > 3/4) = \int_{3/4}^{1} f_Y(y)\,dy$$

$$= \int_{3/4}^{1} 5y^4\,dy = 1-\left(\frac{3}{4}\right)^5$$

$$P\left(0 < X+Y < \frac{1}{2}\right)$$



$$= \int_0^{1/4}\int_x^{\left(\frac{1}{2}-x\right)} 10\,x y^2 \,dy\,dx$$

$$= \frac{10}{3}\int_0^{1/4} x\left[\left(\frac{1}{2}-x\right)^3 - x^3\right] dx$$

$$= \dots\dots$$

$$P\left(X < \frac{1}{2} \,\middle|\, Y = 3/4\right) = \int_0^{1/2} \frac{32}{9} x \, dx \longrightarrow 4/9$$

$$f_{X|Y=3/4}(x) = \begin{cases} \frac{32}{9} x, & 0 < x < \frac{3}{4} \\ 0, & \text{ew} \end{cases}$$

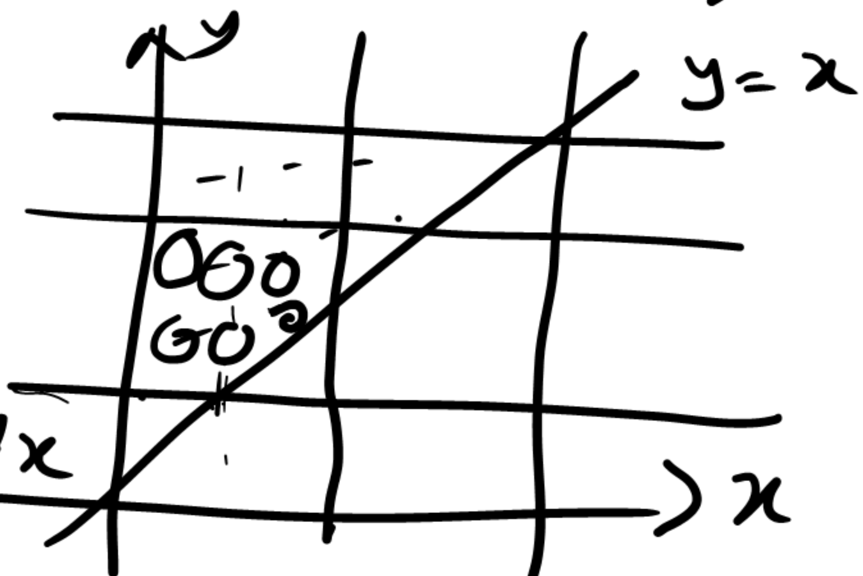
$$P\left(Y < \frac{1}{2} \,\middle|\, X = \frac{1}{4}\right)$$

$$f_{Y|X=1/4}(y) = \begin{cases} \frac{64}{21} y^2, & \frac{1}{4} < y < 1 \\ 0, & \text{ew} \end{cases}$$

$$\int_{1/4}^{1/2} \frac{64}{21}\, y^2\,dy = \frac{1}{9}$$

$$P\left(0 < X < \frac{1}{2},\ \frac{1}{4} < Y < \frac{3}{4}\right)$$

$$= \int_{0}^{1/4}\int_{1/4}^{3/4} 10xy^2\,dy\,dx + \int_{1/4}^{1/2}\int_{x}^{3/4} 10xy^2\,dy\,dx$$



$= \dots \circledast$