Mathematical Statistics 1

Ch.4 Bivariate Distributions

Jungsoon Choi

jungsoonchoi@hanyang.ac.kr

Table of Contents

Ch.4.1 Bivariate Distributions - Discrete Case

1.1 Discrete Case

Definition 4.1-1

Let X and Y be two random variables defined on a discrete probability space. Let $\mathcal{X}$ be the corresponding two-dimensional space of X and Y, the two random variables of the discrete type. The joint probability mass function (joint pmf) of X and Y is

$$f_{XY}(x, y) = P(X = x, Y = y)$$

Properties

1. $0 \leq f_{XY}(x, y) \leq 1.$
2. $\sum_x \sum_y f_{XY}(x, y) = 1.$
3. $P[(X, Y) \in A] = \sum_{(x,y) \in A} f_{XY}(x, y)$, where $A \subset \mathcal{X}$.
4. $E[g(X, Y)] = \sum_{(x,y) \in \mathbb{R}^2} g(x, y) f_{XY}(x, y)$

(X, Y) discreate

$f_{XY}(x,y)$ joint pmf

$P(X=x, Y=y)$

maginal dist $\left\langle \begin{array}{l} f_X(x) \\ f_Y(y) \end{array} \right.$

$\sum_{\forall y} f_{XY}(x,y)$

$\sum_{\forall x} f_{XY}(x,y)$

joint $E[g(X,Y)]$ $\sum_{\forall x}\sum_{\forall y} g(x,y) f_{XY}(x,y)$

joint mgf $M_{XY}(t_1,t_2) = E[e^{\overbrace{t_1X+t_2Y}^{g(X,Y)}}]$ $\sum_{\forall x}\sum_{y} (e^{t_1x+t_2y}) f_{XY}(x,y)$

X & Y : indep.

$\hookrightarrow f_{XY}(x,y) = f_X(x) f_Y(y)$ for $\forall x, y$

continuous.

joint pdf

$\int f_{XY}(x,y)\,dy$

$\int f_{XY}(x,y)\,dx$

$\iint g(x,y) f_{XY}(x,y)\,dx\,dy$

$\iint (e^{t_1x+t_2y}) f_{XY}(x,y)\,dx\,dy$

Example 4.1-1

Roll a pair of unbiased dice. For each of the 36 sample points with probability $1/36$, let X denote the smaller and Y the larger outcome on the dice. Please provide the joint pmf of X and Y.

Ex. 4.1-1

Outcomes.

Outcomes	X	Y
(1,1)	1	1
(1, 2)	1	2
⋮	⋮	⋮
(1.6)	1	6
(2.1)	1	2
(2,2)	2	2
⋮	⋮	⋮
(6.6)	6	6

$f_{XY}(1,1) = \frac{1}{36}$

$f_{XY}(1,2) = \frac{2}{36}$

$$f_{XY}(x,y) = \begin{cases} \frac{2}{36} & 1 \le X < Y \le 6 \\ \frac{1}{36} & 1 \le x = y \le 6 \end{cases}$$

Definition 4.1-2 - marginal pmf

Let X and Y have joint probability mass function $f_{XY}(x,y)$ with space $\mathbb{R}^2$. The probability mass function of X alone, which is called the marginal probability mass function of X, is defined by

$$f_X(x) = Pr(X = x) = \sum_{y \in \mathbb{R}} f_{XY}(x,y), \quad x \in \mathbb{R}.$$

Similarly, the marginal probability mass function of Y is defined by

$$f_Y(y) = Pr(Y = y) = \sum_{x \in \mathbb{R}} f_{XY}(x,y), \quad y \in \mathbb{R}.$$

independence

The random variable X and Y are independent if and only if

$$P(X = x, Y = y) = P(X = x)P(Y = y),$$

or, equivalently, for all $x,\ y$,

$$f_{XY}(x, y) = f_X(x)f_Y(y), \quad (x, y) \in \mathbb{R}^2$$

Otherwise X and Y are dependent.

Example 4.1-2

Let the joint pmf of X and Y be defined by

$$f_{XY}(x, y) = \frac{x + y}{21}, \quad x = 1, 2, 3, \quad y = 1, 2.$$

- Find the marginal pmfs of X and Y.
- Are the random variables X and Y independent?
- Compute $E(XY)$.

Ex 4.1-2 $\quad f_{XY}(x,y) = \frac{x+y}{21}, \quad x = 1,2,3 \quad y = 1,2$

1) $f_X(x) = \sum_{y=1}^{2} f_{XY}(x,y) = \sum_{y=1}^{2} \frac{(x+y)}{21} = \frac{(x+1)+(x+2)}{21} = \frac{2x+3}{21} \quad x = 1,2,3$

marginal pmf.

$f_Y(y) = \sum_{x=1}^{3} \frac{(x+y)}{21} = \frac{(1+y)+(2+y)+(3+y)}{21} = \frac{3y+6}{21} = \frac{y+2}{7} \quad y = 1,2$

2) $f_X(1) f_Y(1) = \frac{5}{21} \times \frac{3}{7} = \frac{5}{49} \neq f_{XY}(1,1) = \frac{2}{21}$

$\therefore$ X & Y are dep.

3) $E[XY] = \sum_{x=1}^{2} \sum_{y=1}^{2} (xy)\left(\frac{x+y}{21}\right) = \sum_{x=1}^{3} x\left[\frac{x+1}{21} + 2\frac{x+2}{21}\right]$

$= \sum_{x=1}^{3} x\left(\frac{3x+5}{21}\right) = \frac{8+2\times 11+3\times 14}{21} = \frac{72}{21} = \frac{24}{7}$

1.2 Trinomial Distribution

Definition of trinomial pmf

We have three mutually exclusive ways for an experiment to terminate: perfect, seconds, defective. We repeat the experiment n independent times, the probabilities p_1, p_2, $p_3 = 1 - p_1 - p_2$ of perfect, seconds, and defective, respectively., remain the same from trial to trial. In the n trials, let X_1=number of perfect items, X_2=number of seconds, and $X_3 = n - X_1 - X_2$=number of defectives. The trinomial pmf is

$$\begin{aligned} f_{X_1X_2}(x_1, x_2) &= P(X_1 = x_1, X_2 = x_2) \\ &= \frac{n!}{x_1!x_2!(n - x_1 - x_2)!} p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{n - x_1 - x_2} \end{aligned}$$

where x_1 and x_2 are nonnegative integers such that $x_1 + x_2 \leq n$.

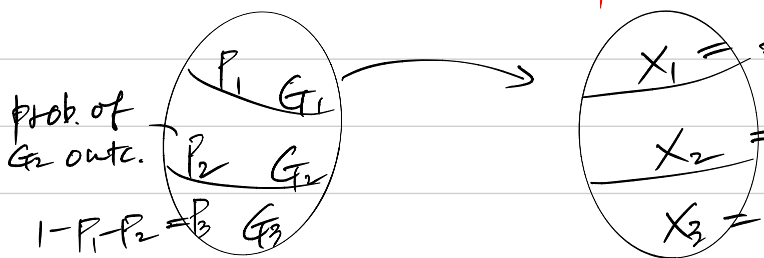
X = # of "s" outcomes with "p"

$n-X$ = # of "f" outcomes with "1-p"

pmf:

$$f_X(x) = P(X=x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0, 1, \cdots, n$$

Trinomial dist^n (pmf)



X_1 = # of G_1 outcomes with "p_1"

X_2 = # of G_2 outcomes with "p_2"

X_3 = # of G_3 outcomes with "$1-p_1-p_2$"

$X_3 = n - X_1 - X_2$

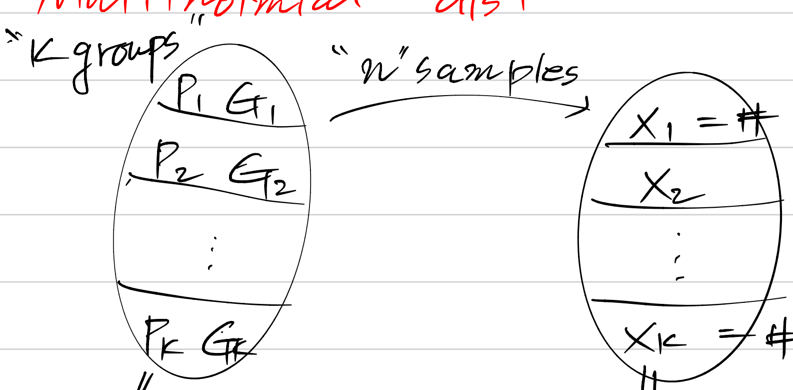
$x_1 + x_2 \le n$

$$f_{X_1 X_2}(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$$

$$= \frac{n!}{x_1!\, x_2!\, (n-x_1-x_2)!} p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{n-x_1-x_2}$$

$\hookrightarrow \binom{n}{x_1}\binom{n-x_1}{x_2}$

Multinoimial dist^n



X_1 = # of G_1 outcomes.

X_K = # of G_K outcomes.

$P_K = 1 - P_1 - \cdots - P_{K-1}$

$X_K = n - X_1 - X_2 - \cdots - X_{K-1}$

$$f_{X_1 \cdots X_{K-1}}(x_1, \cdots, x_{K-1}) = \frac{n!}{x_1!\, x_2! \cdots x_{K-1}!\, \underbrace{(n-x_1-\cdots-x_{K-1})}_{x_K!}} p_1^{x_1} p_2^{x_2} \cdots \underbrace{(1-p_1-p_2\cdots p_{K-1})}_{P_K}^{\overbrace{n-x_1-\cdots-x_{K-1}}^{x_K}}$$

Example 4.1-8

In manufacturing a certain item, it is found that in normal production about 95% of the items are good ones, 4% are "seconds,", and 1% are defective. A company has a program of quality control by statistical methods, and each hour an online inspector observes 20 items selected at random, counting the number X of second and the number Y of defectives. Suppose that the production is normal. Find the probability that, in this sample of size $n = 20$, at least two seconds or at least two defective items are discovered.

"20" samples

0.95 good
0.04 seconds
0.01 defective

X = # of seconds

Y = # of defectives

$$f_{XY}(x,y) = \frac{20!}{x!\,y!\,(20-x-y)!} \times 0.04^{x} \times 0.01^{y} \times 0.95^{(20-x-y)}$$

$$P(X \geq 2 \text{ or } Y \geq 2) = 1 - P(0 \leq X \leq 1,\ 0 \leq Y \leq 1)$$

$$= 1 - P(X=Y=0) - P(X=Y=1) - P(X=0, Y=1) - P(X=1, Y=0).$$

Ch.4.4 Bivariate Distributions - continuous Case

4.1 Continuous Case

Definition

Let X and Y be two continuous-type random variables. The joint probability density function (joint pdf) of X and Y is an integrable function $f_{XY}(x, y)$ satisfying the following properties:

- $f_{XY}(x, y) \geq 0$
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx\, dy = 1.$
- $P((X, Y) \in A) = \int \int_A f_{XY}(x, y) dxdy$

Definition for continuous r.v (marginal pdf)

Let X and Y have joint probability density function $f_{XY}(x, y)$ with space $\mathbb{R}^2$. The probability density function of X alone, which is called the marginal probability density function of X, is defined by

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy, \quad x \in \mathbb{R}.$$

Similarly, the marginal probability density function of Y is defined by

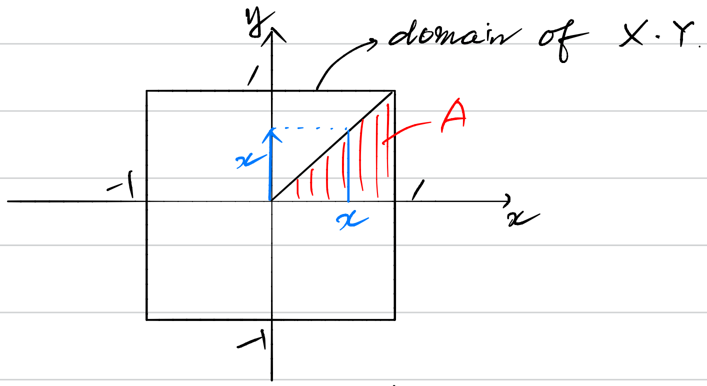
$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx, \quad y \in \mathbb{R}.$$

Example 4.4-2

Let the joint pdf of X and Y be defined by

$$f_{XY}(x,y) = \frac{3}{2}x^2(1-|y|), \quad -1 < x < 1, \quad -1 < y < 1.$$

- Let $A = \{(x,y) : 0 < x < 1, 0 < y < x\}$. Find $P[(X,Y) \in A]$.
- Find the marginal pdfs of X and Y.



1) $P[(X,Y)\in A] = \iint_A f_{XY}(x,y)\,dx\,dy = \int_0^1\int_0^x \frac{3}{2}x^2(1-|y|)\,dy\,dx$

$= \int_0^1\int_0^x \frac{3}{2}x^2(1-y)\,dy\,dx = \int_0^1 \frac{3}{2}x^2\left[y-\frac{1}{2}y^2\right]_0^x dx$

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

$= \frac{3}{2}\left[\frac{1}{4}x^4-\frac{1}{10}x^5\right]_0^1 = \frac{3}{2}\left(\frac{1}{4}-\frac{1}{10}\right) = \frac{3}{2}\,\frac{3}{20} = \frac{9}{40}$

2) $f_X(x) = \int_{-1}^{1} f_{XY}(x,y)\,dy = \int_{-1}^{1}\frac{3}{2}x^2(1-|y|)\,dy$

$= \int_{-1}^{0}\frac{3}{2}x^2(1+y)\,dy + \int_0^1\frac{3}{2}x^2(1-y)\,dy = \frac{3}{2}x^2 \quad -1<x<1$

$f_Y(y) = \int_{-1}^{1} f_{XY}(x,y)\,dx = \int_{-1}^{1}\frac{3}{2}x^2(1-|y|)\,dx = 1-|y|\ ,\ -1<y<1$

Properties

1. $E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{XY}(x,y)\, dx\, dy$
2. The random variable X and Y are independent if and only if

$$f_{XY}(x,y) = f_X(x) f_Y(y), \quad (x,y) \in \mathbb{R}^2$$

Example 4.4-3

Let the joint pdf of X and Y be defined by

$$f_{XY}(x, y) = 2, \quad 0 \leq x \leq y \leq 1.$$

- Find $P(0 \leq X \leq 0.5, 0 \leq Y \leq 0.5)$.
- Find the marginal pdfs of X and Y.
- Find the expected values, $E(X)$, $E(Y)$, and $E(Y^2)$.
- Are the random variables X and Y independent?

$$\int_0^{\frac{1}{2}}\int_0^{y} 2\,dx\,dy = \int_0^{\frac{1}{2}} 2y\,dy = \frac{1}{4}$$

$$f_X(x) = \int_x^1 2\,dy \qquad 0 \le x \le 1$$

$$2(1-x) = 2-2x.$$

$$f_Y(y) = \int_0^y 2\,dx = 2y \qquad 0 \le y \le 1$$

$$E(X) = \int_0^1\int_x^1 2x\,dy\,dx = \int_0^1 2x(1-x)\,dx$$

$$= \int_0^1 2x - 2x^2\,dx = \left[x^2 - \frac{2}{3}x^3\right]_0^1 = 1-\frac{2}{3} = \frac{1}{3}$$

$$\int_0^1 (2-2x)x\,dx$$

$$\int_0^1\int_0^y 2x\,dx\,dy = \int_0^1 y^2\,dy = \left[\frac{1}{3}y^3\right]_0^1 = \frac{1}{3}$$

$$E(Y) = \int_0^1\int_x^1 2y\,dy\,dx = \int_0^1 1-x^2\,dx = \left[x - \frac{1}{3}x^3\right]_0^1$$

$$= 1 - \frac{1}{3} = \frac{2}{3}$$

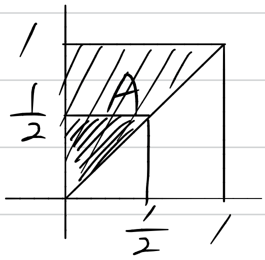
$$= \int_0^1 y \cdot 2y\,dy$$

$$E(Y^2) = \int_0^1\int_x^1 2y^2\,dy\,dx = \int_0^1 \left[\frac{2}{3}y^3\right]_x^1 dx$$

$$= \int_0^1 \frac{2}{3}(1-x^3)\,dx = \frac{2}{3}\left[x - \frac{1}{4}x^4\right]_0^1$$

$$= \frac{2}{3}\cdot\left(1-\frac{1}{4}\right) = \frac{2}{3}\cdot\frac{3}{4} = \frac{1}{2}$$

$$f_X(1)\cdot f_Y(1) = 0 \ne f_{XY} = 2.$$ dep.



$$f_{XY}(x,y) = 2 \qquad 0 \le x \le y \le 1$$

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

$= \int_0^{\frac{1}{2}} (1 - 2x)\,dx = \left[x - x^2\right]_0^{\frac{1}{2}} = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$

2) marginal pdf of X

$= \int_x^1 2\,dy \qquad 0 \le x \le 1$

$= 2(1-x) = 2 - 2x, \quad 0 \le x \le 1 \qquad = f_X(x)$

marginal pdf of Y

$= \int_0^y 2\,dx \qquad 0 \le y \le 1$

$= 2y, \quad 0 \le y \le 1 \qquad = f_Y(y).$

3) $E(X) = \int_0^1 (2-2x)x\,dx = \int_0^1 2x - 2x^2\,dx$

$= \left[x^2 - \frac{2}{3}x^3\right]_0^1 = 1 - \frac{2}{3} = \frac{1}{3}$

$E(Y) = \int_0^1 2y^2\,dy = \left[\frac{2}{3}y^3\right]_0^1 = \frac{2}{3}$

$E(Y^2) = \int_0^1 2y^3\,dy = \left[\frac{1}{2}y^4\right]_0^1 = \frac{1}{2}$

4). If X and Y independent,

$f_{XY}(x,y) = f_X(x) \cdot f_Y(y).$

$f_{XY}(1,1) = 2 \ne 0 = 0 \cdot 2 = f_X(1) \cdot f_Y(1).$

$\therefore$ dependent.

Example 4.4-4

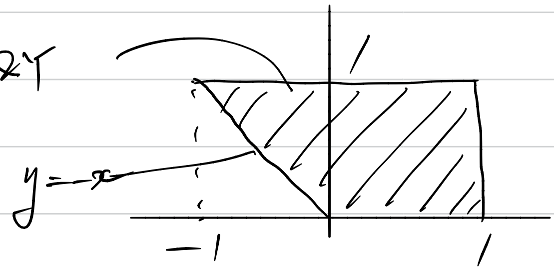
Let the joint pdf of X and Y be defined by

$$f_{XY}(x, y) = cx^2y, \quad -y \leq x \leq 1, \quad 0 \leq y \leq 1.$$

- Determine the value of c.
- Find the marginal pdfs of X and Y.
- Find $P(X \leq 0)$ and $P(0 \leq Y \leq X \leq 1)$.

$f_{XY}(x,y) = cx^2y \qquad -y \le x \le 1, \quad 0 \le y \le 1.$

domain of X&Y



$$\int_0^1\int_{-y}^1 cx^2y\,dx\,dy = 1.$$

$$= \int_0^1 \frac{c}{3}y[x^3]_{-y}^1\,dy = \int_0^1 \frac{c}{3}y(1+y^3)\,dy$$

$$= \int_0^1 \frac{c}{3}(y+y^4)\,dy = \frac{c}{3}\left[\frac{1}{2}y^2 + \frac{1}{5}y^5\right]_0^1$$

$$= \frac{c}{3}\left(\frac{1}{2}+\frac{1}{5}\right) = \frac{c}{3}\cdot\frac{7}{10} = \frac{7}{30}c = 1. \qquad \therefore c = \frac{30}{7}.$$

$$\therefore f_{XY}(x,y) = \frac{30}{7}x^2y$$

$$f_X(x) = \int_0^1 \frac{30}{7}x^2y\,dy = \frac{30}{7}x^2\left[\frac{1}{2}y^2\right]_0^1$$

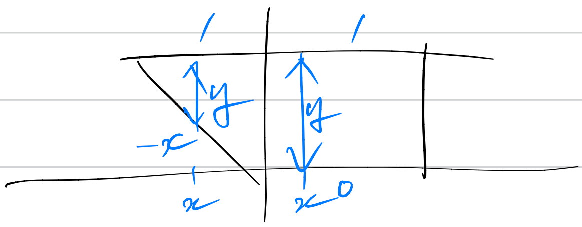
$$= \frac{30}{7}x^2\cdot\frac{1}{2} = \frac{15}{7}x^2 \qquad -1 \le x \le 1. \quad 0 < x \le 1$$

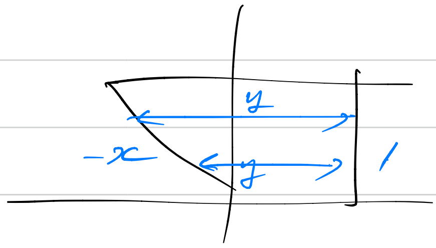
$$f_X(x) = \begin{cases} \int_{-x}^1 f_{XY}(x,y)\,dy & -1 \le x \le 0 \\ \int_0^1 f_{XY}(x,y)\,dy & 0 < x \le 1 \end{cases}$$

$$f_Y(y) = \int_{-y}^1 \frac{30}{7}x^2y\,dx = \frac{30}{7}y\cdot\frac{1}{3}[x^3]_{-y}^1$$

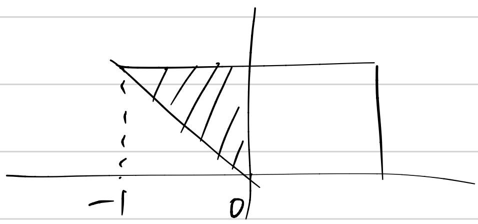
$$= \frac{10}{7}y\cdot(1+y^3) = \frac{10}{7}(y+y^4) \qquad 0 \le y \le 1.$$

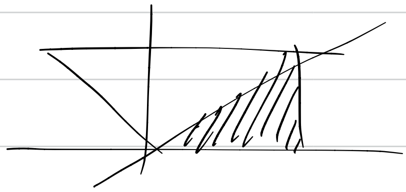
$$f_Y(y) = \frac{10}{7}(y+y^4) \qquad 0 \le y \le 1.$$





$P(X \le 0)$ $\qquad$ $P(0 \le Y \le X \le 1).$ $\qquad$ $f_{XY}(x,y) = \frac{30}{7}x^2y$





$$\int_0^1 \int_{-y}^0 \frac{30}{7}x^2y \, dx\, dy \qquad \int_0^1 \int_0^x \frac{30}{7}x^2y \, dx\, dy.$$

$$\Rightarrow \int_A f_X(x)\, dx.$$

$(x).$

$$= \int_{-1}^0 \frac{15}{7}x^2 \, dx$$

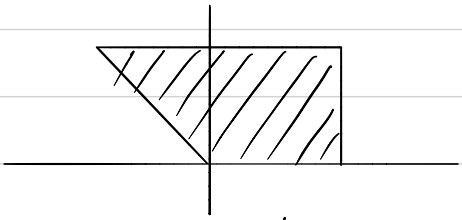
다시 계산 해보기.

$$\frac{5}{7}\left[x^3\right]_{-1}^0 = \frac{5}{7}$$

$$\int_0^1 \left[\frac{10}{7}x^3\right]_{-y}^0 y \, dy$$

$$= \int_0^1 \frac{10}{7}y^4 \, dy$$

$$= \frac{2}{7}\left[y^5\right]_0^1 = \frac{2}{7}$$



$-y \le x \le 1 \quad 0 \le y \le 1$

$f_{XY}(x,y) = cx^2y$

1) $\int_0^1 \int_{-y}^1 cx^2y \, dx dy = 1$

$\int_0^1 \frac{1}{3} cy [x^3]_{-y}^1 dy = 1$

$\frac{1}{3} c \int_0^1 y \cdot (1+y^3) dy = 1$

$\frac{1}{3} c \int_0^1 y + y^4 dy = 1$

$\frac{1}{3} c \left[\frac{1}{2}y^2 + \frac{1}{5}y^5\right]_0^1 = 1$

$\frac{1}{3} c \left(\frac{1}{2} + \frac{1}{5}\right) = 1$

$\frac{1}{3} \cdot \frac{7}{10} \cdot c = \frac{7}{30} c = 1 \quad \therefore c = \frac{30}{7}$.

$f_{XY}(x,y) = \frac{30}{7} x^2 y$

2) marginal pdf of X

$\int_0^1 \frac{30}{7} x^2 y dy \quad 0 \le x \le 1$

$\int_{-x}^1 \frac{30}{7} x^2 y dy \quad -1 \le x \le 0$

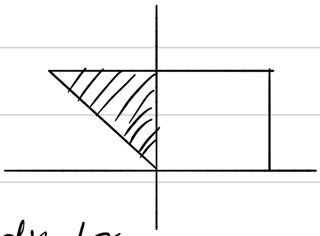
$\frac{15}{7} x^2 \quad 0 \le x \le 1, \quad \frac{15}{7} x^2 (1 - x^2) \quad -1 \le x \le 0$

marginal pdf of Y

$\int_{-y}^1 \frac{30}{7} x^2 y dx \quad 0 \le y \le 1$

$= \frac{1}{3} \cdot \frac{30}{7} y [x^3]_{-y}^1 = \frac{10}{7} y \cdot (1+y^3) = \frac{10}{7}(y + y^4) \quad 0 \le y \le 1$

$P(X \leq 0)$



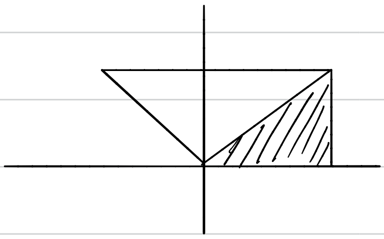
$$\int_{-1}^{0}\int_{-x}^{1} \frac{30}{7}x^2y\,dy\,dx$$

$$= \int_{-1}^{0} \frac{1}{2}\cdot\frac{30}{7}x^2\left[y^2\right]_{-x}^{1} dx \quad = \int_{-1}^{0} \frac{15}{7}x^2(1-x^2)\,dx$$

$$= \int_{-1}^{0} \frac{15}{7}x^2 - \frac{15}{7}x^4\,dx \quad = \frac{1}{3}\cdot\frac{15}{7}\left[x^3\right]_{-1}^{0} - \frac{1}{5}\cdot\frac{15}{7}\left[x^5\right]_{-1}^{0}$$

$$= \frac{5}{7} - \frac{3}{7} = \frac{2}{7}$$

$P(0 \leq Y \leq X \leq 1)$



$$\int_{0}^{1}\int_{0}^{x} \frac{30}{7}x^2y\,dy\,dx$$

$$= \int_{0}^{1} \frac{15}{7}x^4\,dx = \frac{3}{7}\left[x^5\right]_{0}^{1} = \frac{3}{7}$$