# STATISTICS 2

NOTES BY
GAGNEET
KAUR

October

Date:

YOUVA

# WEEK 5

# Joint Distributions

## # Describing Discrete - continuous joint distributions

- $(X, Y)$ : jointly distributed
- $X$ : discrete with range $T_X$ and PMF $p_X(x)$
- For each $x \in T_X$, we have a continuous random variable $Y_x$ with density $f_{Y_x}(y)$.
- $Y_x$ : $Y$ given $X = x$, denoted $(Y | X = x)$
- $f_{Y_x}(y)$ : conditional density of $Y$ given $X = x$, denoted $f_{Y|X=x}(y)$
- Marginal density of $Y$

$$f_Y(y) = \sum_{x \in T_X} p_X(x) \cdot f_{Y|X=x}(y)$$

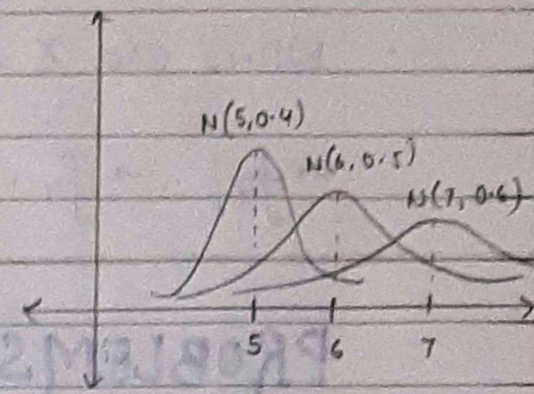
---

discrete uniform conditional densities

# Problem : Let $X \sim$ Uniform $\{0, 1, 2\}$. Let $Y | X = 0 \sim$ Normal$(5, 0.4)$, $Y | X = 1 \sim$ Normal $(6, 0.5)$ and $Y | X = 2 \sim$ Normal $(7, 0.6)$.

- What is the marginal of $Y$?
- Suppose we observe $Y$ to be around $y_0$. What can you say about $X$?

Solution: $f_{Y|X=0}(y) = \dfrac{1}{0.4\sqrt{2\pi}} e^{-\frac{(y-5)^2}{2(0.4)^2}}$




$$f_{Y|X=1}(y) = \frac{1}{0.5\sqrt{2\pi}} e^{-\frac{(y-6)^2}{2(0.5)^2}}$$

$$f_{Y|X=2}(y) = \frac{1}{0.6\sqrt{2\pi}} e^{-\frac{(y-7)^2}{2(0.6)^2}}$$

$$\therefore \quad f_Y(y) = \frac{1}{3}\left(\frac{1}{0.4\sqrt{2\pi}} e^{-\frac{(y-5)^2}{2(0.4)^2}}\right) + \frac{1}{3}\left(\frac{1}{0.5\sqrt{2\pi}} e^{-\frac{(y-6)^2}{2(0.5)^2}}\right) + \frac{1}{3}\left(\frac{1}{0.6\sqrt{2\pi}} e^{-\frac{(y-7)^2}{2(0.6)^2}}\right)$$

---

# CONDITIONAL PROBABILITY OF DISCRETE GIVEN CONTINUOUS

Suppose $X$ and $Y$ are jointly distributed with $X \in T_X$ being discrete with PMF $p_X(x)$ and conditional densities $f_{Y|X=x}(y)$ for $x \in T_X$. The conditional probability of $X$ given $Y = y_0 \in \text{supp}(Y)$ is defined as

$$P(X = x \mid Y = y_0) = \frac{p_X(x)\, f_{Y|X=x}(y_0)}{f_Y(y_0)}$$

where $f_Y$ is the marginal density of $Y$.

$P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$

- Similar to Bayes' rule : $P(X=x \mid Y=y_0)\, f_Y(y_0) = f_{Y|X=x}(y_0)\, p_X(x)$

- $X \mid Y = y_0$ : 'conditioned' discrete random variable

- When are $X$ and $Y$ independent? $f_{Y|X=x}$ is independent of $x$

$\rightarrow f_{Y|X} \rightarrow f_Y = f_{Y|X=x}$ and $P(X=x|Y=y_0) = p_X(x)$

---

## PROBLEMS

(1) Let $X \sim$ Uniform $\{-1, 1\}$. Let $Y|X=-1 \sim$ Uniform $[-2, 2]$, $Y|X=1 \sim Exp(5)$. Find the distribution of $X$ given $Y=-1$, $Y=1$, $Y=3$.

Solution: $f_Y(y) = \frac{1}{2}\left(f_{Y|X=-1}(y)\right) + \frac{1}{2}\left(f_{Y|X=1}(y)\right)$

$\Rightarrow f_Y(y) = \underbrace{\frac{1}{2} \times \frac{1}{4}}_{y \in [-2,2]} + \frac{1}{2} \times 5e^{-5y} \rightarrow y > 0$

$$\Rightarrow f_Y(y) = \begin{cases} 0 & , \quad y < -2 \\ \frac{1}{2} \times \frac{1}{4} & , \quad -2 < y < 0 \\ \frac{1}{2} \times \frac{1}{4} + \frac{1}{2} \times 5e^{-5y} & , \quad 0 < y < 2 \\ \frac{1}{2} \times 5e^{-5y} & , \quad y > 2 \end{cases}$$

Now,

$$X|Y=-1 : \quad P(X=-1|Y=-1) = \frac{p_X(-1) \cdot f_{Y|X=-1}(-1)}{f_Y(-1)}$$

(i) $\Rightarrow P(X=-1|Y=-1) = \dfrac{\frac{1}{2} \times \frac{1}{4}}{\frac{1}{2} \times \frac{1}{4}} = 1$

(ii) $\Rightarrow P(X=1|Y=-1) = \dfrac{p_X(1) \cdot f_{Y|X=1}(-1)}{f_Y(-1)} = \dfrac{\frac{1}{2} \times 0}{\frac{1}{2} \times \frac{1}{4}} = 0$

$X \mid Y=1$ :

$$P(X=-1 \mid Y=1) = \frac{1/2 \times 1/4}{1/2 \times 1/4 + 1/2 \times 5e^{-5}}$$

$$P(X=1 \mid Y=1) = 1 - P(X=-1 \mid Y=1) = \frac{1/2\, 5e^{-5}}{1/2 \times 1/4 + 1/2\, 5e^{-5}}$$

$X \mid Y=3$ :

$$P(X=-1 \mid Y=3) = 0$$

$$P(X=1 \mid Y=3) = 1$$

---

(2) Suppose 60% of adults in the age group of 45-50 in a country are male and 40% are female. Suppose the height (in cm) of adult males in that age group in the country is Normal ($\overset{\mu=}{160}, \overset{\sigma=}{10}$), and that of females is Normal ($\overset{\mu=}{150}, \overset{\sigma=}{5}$). A random person is found to have a height of 155 cm. Is that person more likely to be male or female?

Solution: $X \sim \{\overset{0.6}{M}, \overset{0.4}{F}\}$ $\qquad Y \mid X=M \sim$ Normal $(160, \sigma=10)$

$\qquad\qquad\qquad\qquad\quad Y \mid X=F \sim$ Normal $(150, \sigma=5)$

$X \mid Y=155$ : $$P(X=M \mid Y=155) = \frac{0.6 \times \frac{1}{10\sqrt{2\pi}} e^{-\frac{5^2}{2 \cdot 10^2}}}{0.6 \times \frac{1}{10\sqrt{2\pi}} e^{-5^2/2\cdot 10^2} + \frac{0.4}{\sqrt{2\pi}\cdot 5} e^{-5^2/2\cdot 5^2}}$$

$$P(X=F \mid Y=155) = 1 - P(X \mid M \mid Y=155)$$

(8) Let $Y = X + Z$, where $X \sim$ Uniform $\{-3, -1, 1, 3\}$ and $Z \sim$ Normal $(0, \sigma^2)$ are independent. What is the distribution of $Y$? Find the distribution of $(X | Y = 0.5)$

Solution: $Y | X = -3 \rightarrow (-3 + Z) \sim N(-3, \sigma^2)$

$Y | X = -1 \rightarrow (-1 + Z) \sim N(-1, \sigma^2)$

$Y | X = 1 \rightarrow (1 + Z) \sim N(1, \sigma^2)$

$Y | X = 3 \rightarrow (3 + Z) \sim N(3, \sigma^2)$

October 5, 2021

| M T W T F S S | |
|---|---|
| Page No.: | YOUVA |
| Date: | |

# JOINTLY CONTINUOUS DISTRIBUTION

## * Joint Density in two Dimensions

A function $f(x,y)$ is said to be a joint density $f^n$ if-

→ $f(x,y) \geq 0$ , i.e., $f$ is non-negative

→ $\int\int_{-\infty}^{\infty} f(x,y)\, dx \cdot dy = 1$

→ Technical : $f(x,y)$ is piecewise continuous in each variable

- For every joint density $f(x,y)$, there exist two jointly distributed continuous random variables $X$ and $Y$ such that, for any two dimensional region $A$,

$$P\left((X,Y) \in A\right) = \iint_A f(x,y)\, dx \cdot dy$$

- $f(x,y)$ also denoted $f_{XY}(x,y)$ is also called the joint density of $X$ and $Y$
- $supp(X,Y) = \{(x,y) : f_{XY}(x,y) > 0\}$

Example : Uniform in the unit square

Let $X$ and $Y$ have joint density

$$f_{XY}(x,y) = \begin{cases} 1 & , \; 0 < x < 1 , \; 0 < y < 1 \\ 0 & , \; \text{otherwise} \end{cases}$$

→ To compute the probability, find the area of the region.

Joint

- $P(0 < X < 0.1,\ 0 < Y < 0.1) = (0.1)^2 = 0.01$
- $P(0.5 < X < 0.6,\ 0 < Y < 0.1) = (0.1)^2 = 0.01$
- $P(0.9 < X < 1,\ 0.9 < Y < 1) = (0.1)^2 = 0.01$

Marginal

- $P(0 < X < 0.1) = 0.1$
- $P(0.5 < Y < 0.6) = 0.1$
- $P(X > Y) = 0.5$
- $P(X > 2Y) = 0.25$
- $P(X^2 + Y^2 < 0.25) = \frac{1}{4} \times \pi \left(\frac{1}{2}\right)^2 = \frac{\pi}{16}$

---

# 2D UNIFORM DISTRIBUTION

For some (reasonable) region $D$ in $\mathbb{R}^2$ with total area $|D|$. We say that $(X,Y) \sim \text{Uniform}(D)$ if they have joint density

$$f_{XY}(x,y) = \begin{cases} 1/|D| & , (x,y) \in D \\ 0 & , \text{otherwise} \end{cases}$$

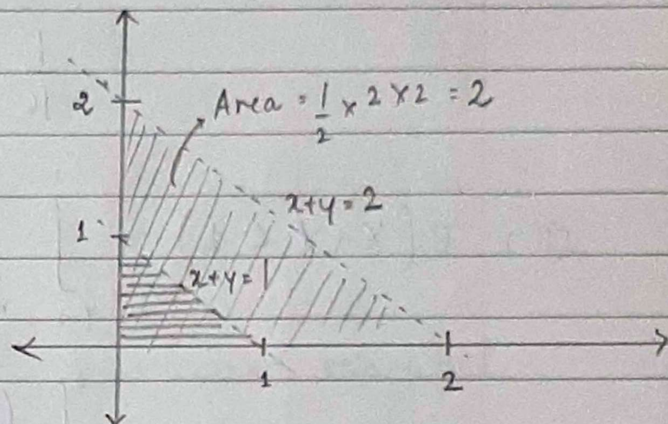
- Rectangle : $D = [a,b] \times [c,d] = \{(x,y) : a < x < b,\ c < y < d\}$
- Circle : $D = \{(x,y) : (x - x_0)^2 + (y - y_0)^2 \le r^2\}$
- For any sub-region $A$ of $D$, $P((X,Y) \in A) = \frac{|A|}{|D|}$

- Uniform distribution is a good approximation for flat histograms.

# PROBLEMS

Uniform (1) Let $(X, Y) \sim$ Uniform $(D)$, where $D = \{(x, y) : x + y < 2, x > 0, y > 0\}$. Sketch the support and compute $P(X + Y < 1)$, $P(X + 2Y > 1)$

Solution : $f_{XY}(x, y) = \begin{cases} 1/2, & (x, y) \in D \\ 0, & \text{otherwise} \end{cases}$




$P(X + Y < 1) = \frac{1/2}{2} = \frac{1}{4}$

$P(X + 2Y > 1) = 2 - (1/2 \times 1 \times 1/2) = \frac{7}{4} \times \frac{1}{2} = \frac{7}{8}$

Non-uniform (2) Let $(X, Y)$ have joint density $f_{XY}(x, y) = \begin{cases} x + y, & 0 < x, y < 1 \\ 0, & \text{otherwise} \end{cases}$

Show that the above is a valid density.
Find $P(X < 1/2$ and $Y < 1/2)$, $P(X + Y < 1)$.

Solution : $\int_{y=0}^{1} \int_{x=0}^{1} (x + y) \, dx \, dy = \int_{y=0}^{1} \left[ \frac{x^2}{2} + yx \right]_0^1 dy = \int_{y=0}^{1} \left( \frac{1}{2} + y \right) dy$

$= \left[ \frac{y^2}{2} + \frac{y}{2} \right]_0^1 = 1$, Thus it is a valid density.

(i) $P(X < 1/2 \text{ and } Y < 1/2) = \int_0^{1/2}\int_0^{1/2} (x+y)\, dx\, dy$

$\Rightarrow P(X < 1/2, Y < 1/2) = \int_0^{1/2} \left[\frac{x^2}{2} + yx\right]_0^{1/2} dy$

$= \int_0^{1/2} \left[\frac{1}{8} + \frac{y}{2}\right] dy = \left[\frac{y}{8} + \frac{y^2}{4}\right]_0^{1/2}$

$= \frac{1}{16} + \frac{1}{16} = \frac{1}{8}$

(ii) $P(X+Y < 1) = \int_{y=0}^{1}\int_{x=0}^{1-y} (x+y)\, dx\, dy = \int_{y=0}^{1} \left[\frac{x^2}{2} + yx\right]_0^{1-y} dy$

$= \int_{y=0}^{1} \frac{(1-y)^2}{2} + y(1-y)\; dy$

$= \int_{y=0}^{1} \frac{1}{2}(1+y^2-\cancel{2y} + \cancel{2y} - 2y^2)\; dy$

$= \int_{y=0}^{1} \frac{1}{2}(1-y^2)\, dy = \frac{1}{2}\int_{y=0}^{1} (1-y^2)\, dy$

$= \frac{1}{2}\left[y - \frac{y^3}{3}\right]_0^1 = \frac{1}{2}\left[1 - \frac{1}{3}\right] = \frac{2}{6} = \frac{1}{3}$

October 6, 2021

# Marginal Densities

## MARGINAL DENSITY

Suppose $(X, Y)$ have joint density $f_{XY}(x,y)$. Then,

- $X$ has the marginal density $f_X(x) = \int_{y=-\infty}^{\infty} f_{XY}(x,y)\, dy$

- $Y$ has the marginal density $f_Y(y) = \int_{x=-\infty}^{\infty} f_{XY}(x,y)\, dx$

- The PDF of $X$ and $Y$ individually are called marginal densities.

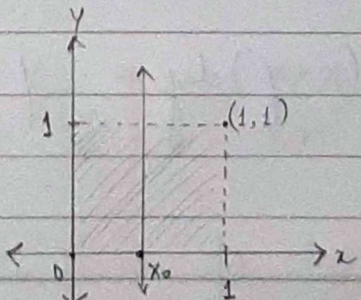
- The joint density exactly determines both the marginal densities.

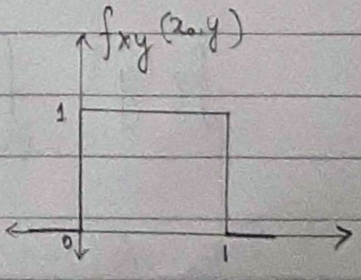
Examples : Marginals do not determine joint

(1) Uniform on unit square

$$f_X(x_0) = \int_{y=0}^{1} f_{XY}(x,y)\, dy = 1 \quad , \ 0 < x_0 < 1$$

$X \sim$ Uniform $[0, 1]$
$Y \sim$ Uniform $[0, 1]$

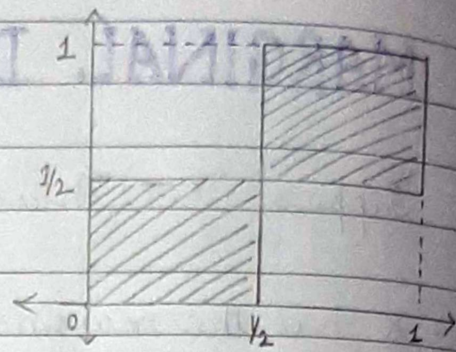


(2) $(X,Y) \sim \text{Uniform}(D)$, where

$$D = [0, 1/2] \times [0, 1/2] \cup [1/2, 1] \times [1/2, 1]$$

$$f_{XY}(x,y) = \begin{cases} 2, & (x,y) \in D \\ 0, & \text{otherwise} \end{cases}$$




$f_X(x_0) = 1, \quad x_0 \in [0, 1/2]$

$f_X(x_1) = 1, \quad x_1 \in [1/2, 1]$

$\therefore \; f_X(x) = 1, \quad 0 < x < 1$

$X \sim \text{Uniform}[0,1]$
$Y \sim \text{Uniform}[0,1]$

---

<u>PROBLEM</u> : Consider the joint density

$$f_{XY}(x,y) = \begin{cases} x+y, & 0 < x,y < 1 \\ 0, & \text{otherwise} \end{cases}$$

Find the marginals.

$$f_X(x) = \int_{y=0}^{1} (x+y)\,dy = \left[xy + \frac{y^2}{2}\right]_0^1 = x + \frac{1}{2}, \quad 0 < x < 1$$

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

---

| M | T | W | T | F | S | S |
|---|---|---|---|---|---|---|

Page No.:
Date:
YOUVA

# Independence

<u>Theorem</u> : $(X,Y)$ with joint density $f_{XY}(x,y)$ are independent if

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

where $f_X(x)$ and $f_Y(y)$ are the marginal densities.

- Given the joint density, the marginals can be computed.
- If the joint density is the product of the marginal densities, then $X$ and $Y$ are independent.
- So, if independent, the marginals determine the joint density.

---

Problem : Suppose $X \sim Exp(\lambda_1)$, $Y \sim Exp(\lambda_2)$ are independent random variables. Find their joint density and compute $P(X > Y)$.

Solution : $f_X(x) = \lambda_1 e^{-\lambda_1 x}$ , $x > 0$

$f_Y(y) = \lambda_2 e^{-\lambda_2 y}$ , $y > 0$

$f_{XY}(x,y) = \lambda_1 e^{-\lambda_1 x} \cdot \lambda_2 e^{-\lambda_2 y}$ ; $x > 0, y > 0$

$\therefore P(X > Y) = \int_{x=0}^{\infty} \int_{y=0}^{x} \lambda_1 e^{-\lambda_1 x} \cdot \lambda_2 e^{-\lambda_2 y} \, dy \, dx = 1 - \dfrac{\lambda_1}{\lambda_1 + \lambda_2}$

# Conditional Density

**Definition :**

Let $(X, Y)$ be random variables with joint density $f_{XY}(x,y)$. Let $f_X(x)$ and $f_Y(y)$ be the marginal densities.

→ For 'a' such that $f_X(a) > 0$, the conditional density of $Y$ given $X = a$, denoted $f_{Y|X=a}(y)$, is defined as

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

→ For 'b' such that $f_Y(b) > 0$, the conditional density of $X$ given $Y = a$, denoted $f_{X|Y=b}(x)$, is defined as

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

**Properties :**

- Both the conditional densities are valid densities in one dimension. So, the 'conditional' random variables $(Y|X=a)$ and $(X|Y=b)$ are well defined.

- Joint = Marginal × conditional, for $x = a$ and $y = b$ such that $f_X(a) > 0$ and $f_Y(b) > 0$

$$f_{XY}(a,b) = f_X(a) \cdot f_{Y|X=a}(b) = f_Y(b) \cdot f_{X|Y=b}(a)$$

- The above is usually written as

$$f_{XY}(x,y) = f_X(x) \cdot f_{Y|X=x}(y) = f_Y(y) \cdot f_{X|Y=y}(x)$$