Lecture 08

- Marginal PMF, PDF, CDF
- Some common 2D PDF
- Conditional PMF, PDF, CDF

(条件分布律，分布函数，概率密度函数)

Marginal Cumulative Density Function (边缘概率分布函数)

Def. Given $F(X,Y)$ as the joint CDF of (X,Y), the marginal CDF of (X,Y) are defined as

$$F_X(x) = P(X \le x) = P(X \le x, Y < +\infty) = F(x, +\infty)$$

Marginal CDF of X

$$F_Y(y) = P\{Y \le y\} = P(X < +\infty, Y \le y) = F(+\infty, y)$$

Marginal CDF of Y

$F(x,y)$ $\xrightarrow{\text{Marginalization (边缘化)}}$ $F_X(x), F_Y(y)$

Note:
Marginal CDF vs. CDF
$F_X(x)$ vs. $F(x)$

Depend on how many variables are in the context!

2-Dimensional

$$F(x,y) \quad \underset{F(x,y)=\int_{-\infty}^{y}\left[\int_{-\infty}^{x} f(u,v)du\right]dv}{\overset{f(x,y)=\frac{\partial^2 F(x,y)}{\partial x \partial y}}{\rightleftarrows}} \quad f(x,y)$$

???

Let $y \to +\infty$,

$F(x,+\infty) = F_X(x)$

Let $x \to +\infty$,

$F(+\infty,y) = F_Y(y)$

$$F_X(x), F_Y(y) \quad \underset{F(x)=\int_{-\infty}^{x} f(t)\, dt}{\overset{f(x)=\frac{dF(x)}{dx}}{\rightleftarrows}} \quad f_X(x), f_Y(y)$$

1-Dimensional

Marginal Probability Density Function

Def. Given $f(X,Y)$ as the joint PDF of (X,Y),

$$F_X(x) = F(x,+\infty) = \int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty} f(x,y)dy\right]dx, \qquad -\infty < x < +\infty$$

Taking derivative of both sides, it yields the marginal PDF of X

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy\,, \qquad -\infty < x < +\infty$$

Similarly, we obtain the marginal PDF of Y

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx\,, \qquad -\infty < y < +\infty$$

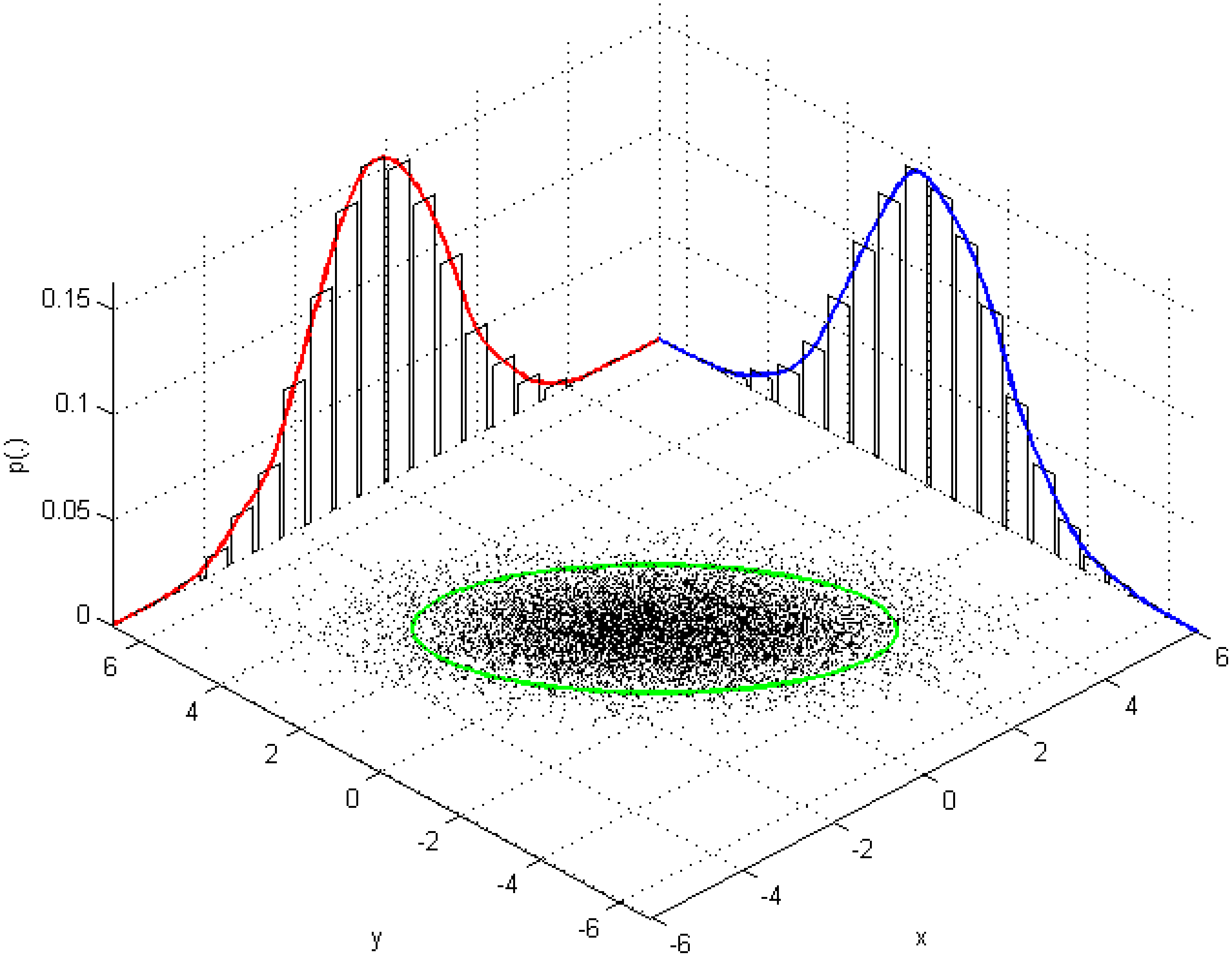


Illustration of marginal PDF or PMF

2-Dimensional

$F(x,y)$

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

$$F(x,y) = \int_{-\infty}^{y} \left[\int_{-\infty}^{x} f(u,v)du \right] dv$$

$f(x,y)$

???

Let $y \to +\infty$,

$F(x,+\infty) = F_X(x)$

Let $x \to +\infty$,

$F(+\infty,y) = F_Y(y)$

???

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx$$

$F_X(x), F_Y(y)$

$$f(x) = \frac{dF(x)}{dx}$$

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

$f_X(x), f_Y(y)$

1-Dimensional

Probability concept explained: Marginalization

Probability Concepts Explained: Marginalization, https://towardsdatascience.com/probability-concepts-explained-marginalisation-2296846344fc

- Marginalisation is a method that requires summing over the possible values of one variable to determine the marginal contribution of another.
- Suppose we're interested in people's happiness under different weather condition. We can write this mathematically as P☺(weather).
- The problem is that people live in different region, and we can't just get rid of it in the measurement. What we're actually measuring is P☺(weather, region).
- Marginalization tells us that we can calculate the quantity we want if we sum over all regions. P☺(weather) = P☺(weather, region=R_1) + P☺(weather, region=R_2) +…+ P☺(weather, region=R_n).
- Note: Compare with the total probability law.

Common 2D PDF

Uniform distribution (均匀分布),

$$f(x,y) = \begin{cases} \dfrac{1}{A}, & (x,y) \in G \\ 0 & \text{otherwise} \end{cases}$$

G is an area in arbitrary shape.

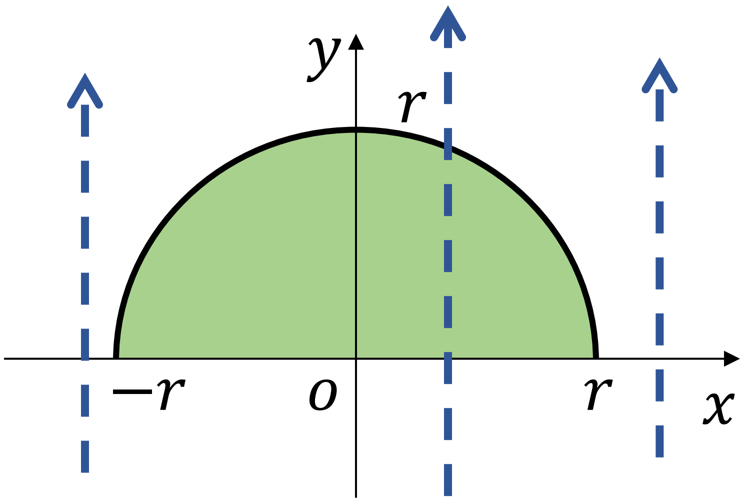
$$F(x,y) = \iint\limits_{R} f(x,y)dxdy = \frac{1}{A} \cdot S$$

where R denotes the ROI, and S is the area of R.

Ex. Let G be the domain of $(X, Y) \in \{x^2 + y^2 \leq r^2 \cap y \geq 0\}$, (X, Y) follows uniform distribution, find the marginal PDF of X and Y.

Ex. Let G be the domain of $(X,Y) \in \{x^2 + y^2 \le r^2 \cap y \ge 0\}$, (X,Y) follows uniform distribution, find the marginal PDF of X and Y.

Sol. $$f(x,y) = \begin{cases} \dfrac{2}{\pi r^2}, & (x,y) \in G, \\ 0, & \text{otherwise.} \end{cases}$$

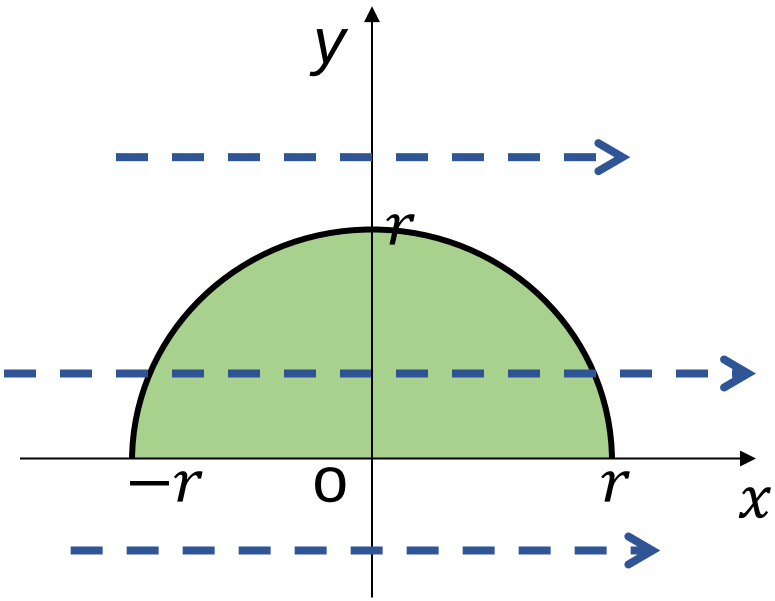


$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy$$

$$= \begin{cases} \displaystyle\int_0^{\sqrt{r^2-x^2}} \frac{2}{\pi r^2} dy, & -r < x < r, \\ 0, & \text{otherwise.} \end{cases} = \begin{cases} \dfrac{2}{\pi r^2}\sqrt{r^2 - x^2}, & |x| < r \\ 0, & \text{otherwise.} \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx$$

$$= \begin{cases} \int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} \frac{2}{\pi r^2} dx\,, & 0 < y < r \\ 0, & \text{otherwise} \end{cases}$$



$$= \begin{cases} \frac{4}{\pi r^2}\sqrt{r^2 - y^2}, & 0 < y < r \\ 0, & \text{otherwise} \end{cases}$$

Common 2D PDF

Two-dimensional Normal Distribution,

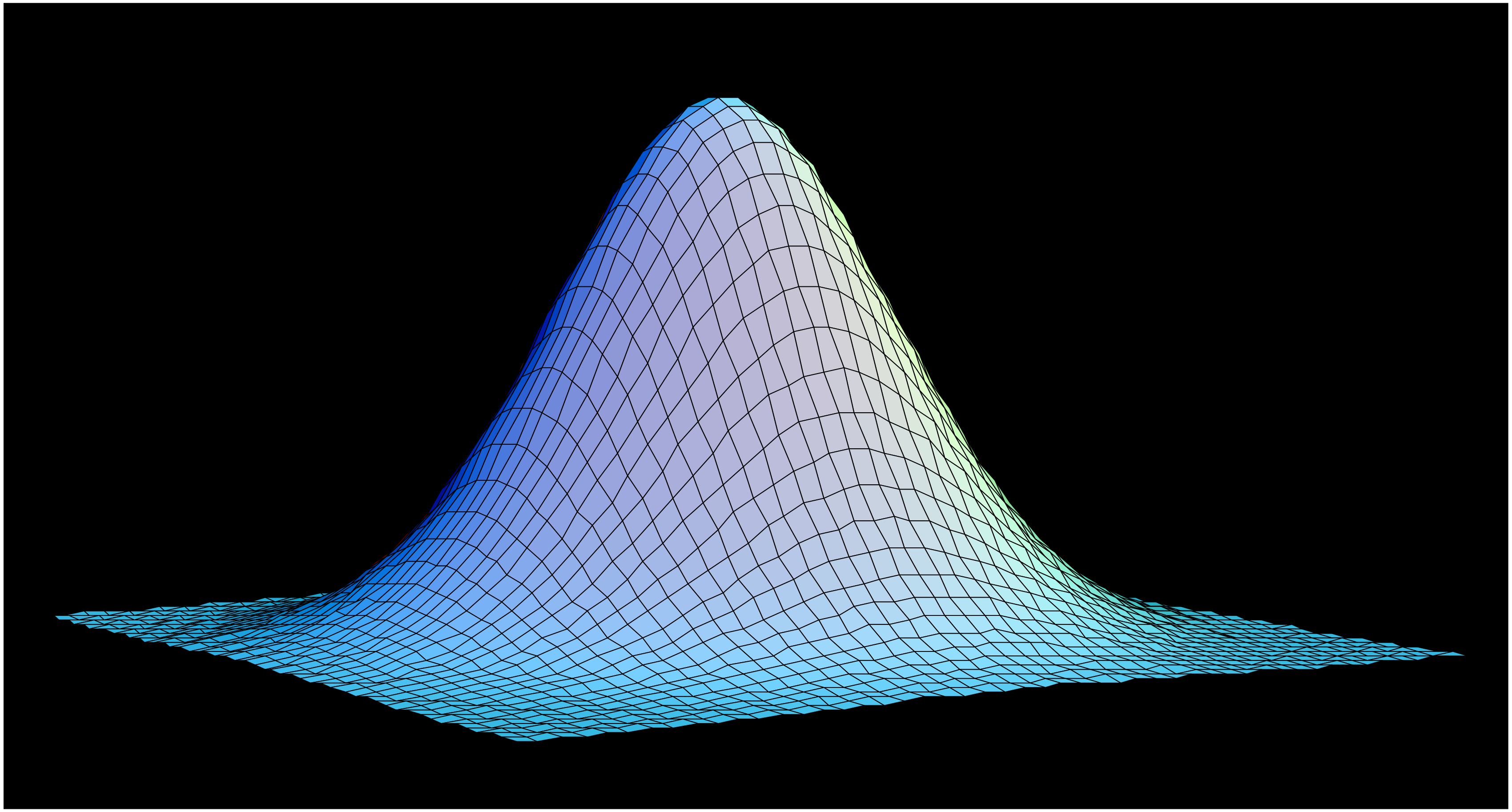
$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1{}^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2{}^2}\right]}$$

$$-\infty < x < \infty, \quad -\infty < y < \infty$$

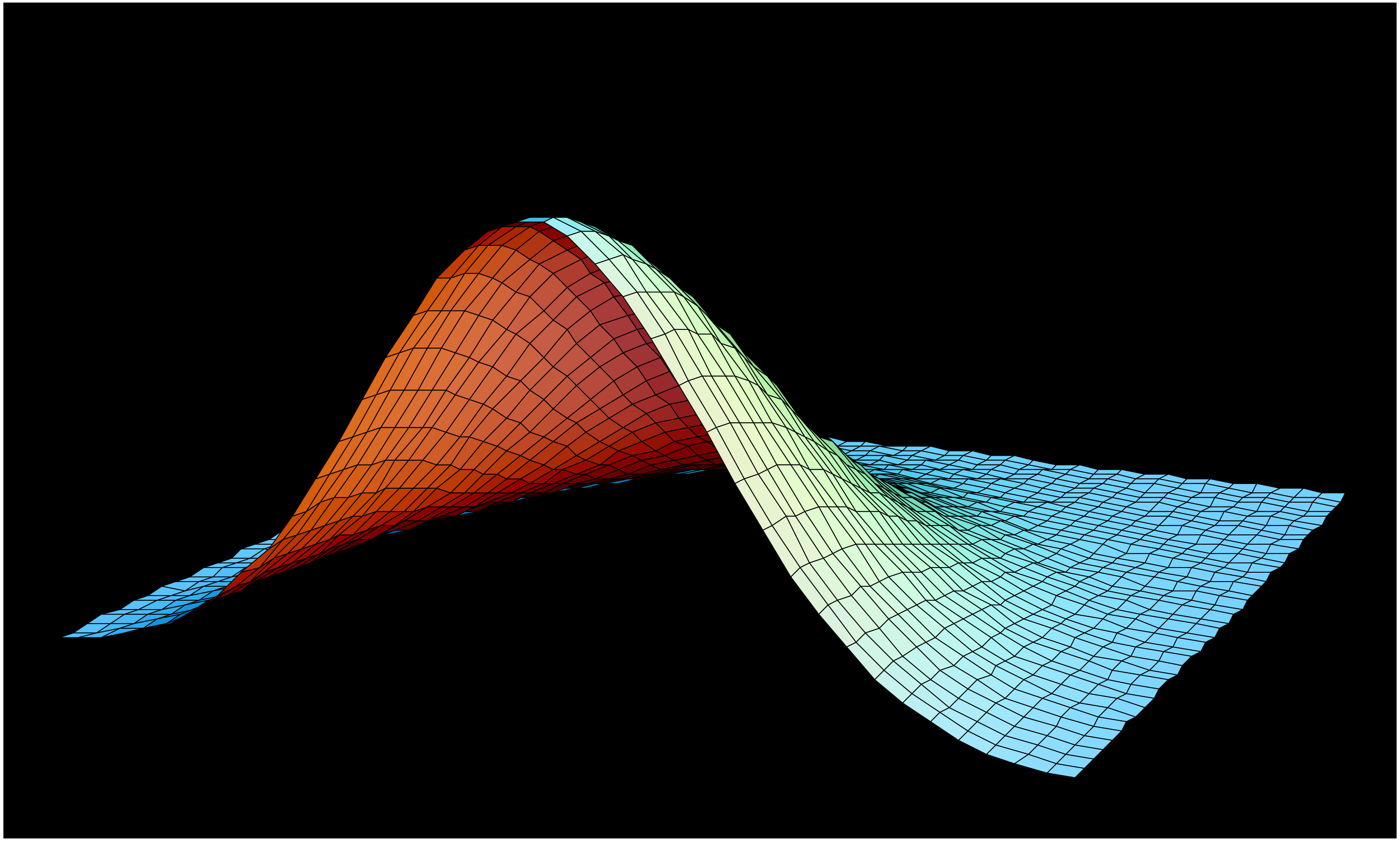
where $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ are constant, and $\sigma_1 > 0, \sigma_2 > 0, |\rho| \leq 1$. 2D normal distribution is denoted as

$$(X, Y) \sim \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$$

where ρ is the correlation coefficient between X and Y (to be covered in a few weeks).



2D PDF of normal distribution



Section view of 2D normal distribution

Ex. Let $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, find the marginal PDF of X and Y.

Sol.

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}},-\infty<x<+\infty$$

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}},-\infty<y<+\infty$$

The marginal PDF of X and Y follows normal distribution.

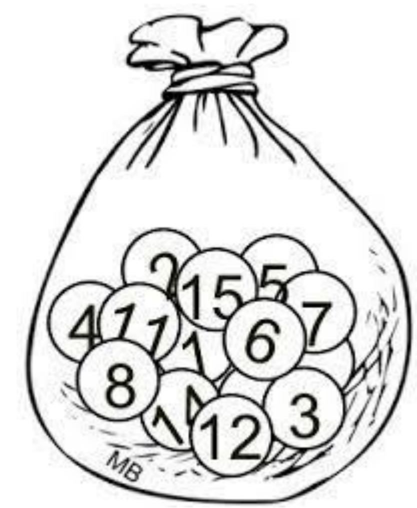
$$X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$$

Verify the result by yourselves.

Lecture 08

- Marginal PMF, PDF, CDF
- Some common 2D PDF
- Conditional PMF, PDF, CDF

(条件分布律，分布函数，概率密度函数)



Conditional PMF vs. Marginal PMF for discrete R.V.s
条件分布律 vs. 边缘分布率

	y_1	y_2	…	y_j	…	$p_{i\cdot}$
x_1	p_{11}	p_{12}	…	p_{1j}	…	$p_{1\cdot}$
x_2	p_{21}	p_{22}	…	p_{2j}	…	$p_{2\cdot}$
⋮	⋮	⋮		⋮		⋮
x_i	p_{i1}	p_{i2}	…	p_{ij}	…	$p_{i\cdot}$
⋮	⋮	⋮		⋮		⋮
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	…	$p_{\cdot j}$	…	1

The practical implication of the Marginal PMF?

$$P\left(X = x_i \mid Y = y_j\right)$$

$$P\left(Y = y_j \mid X = x_i\right)$$

Conditional probability?

Discrete conditional distributions

Recall the definition of the conditional probability of event E given event F:

$$P(E|F) = \frac{P(EF)}{P(F)}$$

For discrete random variables X, Y, the conditional probability mass function (PMF) **of $\boldsymbol{X}$** given $Y = y_j$ is

$$P\left(X = x_i \mid Y = y_j\right) = \frac{P\left(X=x_i, Y=y_j\right)}{P\left(Y=y_j\right)},$$

Note: $P\left(Y = y_j\right) \neq 0$, and $i = 1,2, \ldots$

Different notation, same meaning:

$$p_{X|Y}\left(x_i|y_j\right) = \frac{p_{X,Y}(x_i, y_j)}{p_Y(y_j)} = \frac{p_{ij}}{p_{\cdot j}}$$

Discrete conditional distributions

For discrete random variables X, Y, the conditional probability mass function (PMF) **of $\boldsymbol{X}$** given $Y = y_j$ is

$$P(X = x_i \mid Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)}, \qquad i = 1,2, \ldots$$

The conditional cumulative distribution function (CDF) **of $\boldsymbol{X}$** given $Y = y$ is

$$F_{X|Y}(x_i|y_j) = P(X \le x_i | Y = y_j) = \sum_{x \le x_i} P_{X|Y}(x|y_j), \qquad i = 1,2, \ldots$$

Note: Compare with the unconditional case, a condition $P(Y = y_j)$ is considered in both PMF and CDF.

Conditional PMF vs. Marginal PMF for discrete R.V.s
条件分布律 vs. 边缘分布率

	y_1	y_2	…	y_j	…	$p_{i\cdot}$
x_1	p_{11}	p_{12}	…	p_{1j}	…	$p_{1\cdot}$
x_2	p_{21}	p_{22}	…	p_{2j}	…	$p_{2\cdot}$
⋮	⋮	⋮		⋮		⋮
x_i	p_{i1}	p_{i2}	…	p_{ij}	…	$p_{i\cdot}$
⋮	⋮	⋮		⋮		⋮
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	…	$p_{\cdot j}$	…	1

Marginal PMF: the PMF of a multi-dimensional variable along a dimension of interest;

Conditional PMF: the PMF of a multi-dimensional variable under a given condition.

Notations in Conditional PMF vs. Marginal PMF

$$P_X(x) = p_{i\cdot} = \sum_{j=1}^{+\infty} p_{ij} \; (i = 1,2, \dots)$$

Marginal PMF of X

$$P_Y(y) = p_{\cdot j} = \sum_{i=1}^{+\infty} p_{ij} \; (j = 1,2, \dots)$$

Marginal PMF of Y

$$P\{X = x_i | Y = y_j\} = \frac{p_{ij}}{p_{\cdot j}},$$
$$i = 1,2, \dots$$

Conditional PMF of X

$$P\{Y = y_j | X = x_i\} = \frac{p_{ij}}{p_{i\cdot}},$$
$$j = 1,2, \dots$$

Conditional PMF of Y

Note:

1) Conditional PMF = Joint PMF ÷ Marginal PMF;
2) Conditional PMF = Conditional Probability + Mass Function.

条件概率 + 分布律

Ex. A Safety Officer for an auto insurance company was interested in learning how the extent of an individual's injury in an automobile accident relates to the type of safety restraint the individual was wearing at the time of the accident. As a result, the Safety Officer used statewide ambulance and police records to compile the following two-way table of joint probabilities:

Extent of Injury (X)	Type of Restraint (Y)			$P_X(x)$
	None (0)	Belt only (1)	Belt and Harness (2)	
None (0)	0.065	0.075	0.06	0.20
Minor (1)	0.175	0.16	0.115	0.45
Major (2)	0.135	0.10	0.065	0.30
Death (3)	0.025	0.015	0.01	0.05
$P_Y(y)$	0.40	0.35	0.25	1.00

The Safety Officer was interested in answering the following questions:

1) What is the probability that a randomly selected person in an automobile accident was wearing a seat belt and had only a minor injury? 0.16
2) If a randomly selected person wears no restraint, what is the probability of death? 0.025/0.40
3) If a randomly selected person sustains no injury, what is the probability the person was wearing a belt and harness? 0.06/0.20

Quick check

$$P\left(X = x_i | Y = y_j\right) = \frac{P\left(X = x_i, Y = y_j\right)}{P\left(Y = y_j\right)}$$

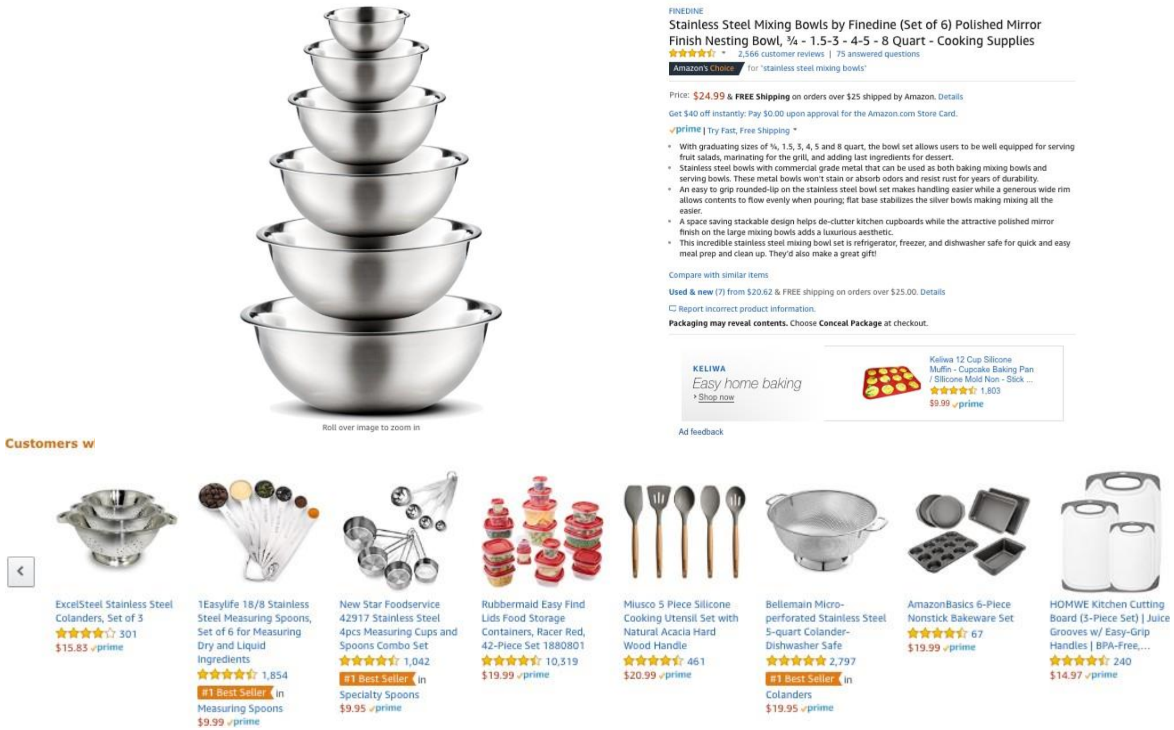
Number or function?

1. $P(X = 2|Y = 5)$
   number
2. $P(X = x_i|Y = 5)$
   1-D function
3. $P\left(X = 2|Y = y_j\right)$
   1-D function
4. $P\left(X = x_i|Y = y_j\right)$
   2-D function

True or false?

1. $\sum_i P(X = x_i|Y = 5) = 1$
   true
2. $\sum_j P\left(X = 2|Y = y_j\right) = 1$
   false
3. $\sum_i \sum_j P\left(X = x_i|Y = y_j\right) = 1$
   false
4. $\sum_i \left(\sum_j P\left(X = x_i|Y = y_j\right) P\left(Y = y_j\right)\right) = 1$
   true

Conditional probability in Amazon



$$P(\text{buy item } X \mid \text{bought item } Y)$$

The conditional relationship between purchasing different items.

Conditional PDF and CDF for continuous R.V.s

Conditional PDF of X given $Y = y$.

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

Joint PDF

Marginal PDF

$$F_{X|Y}(x|y) = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)} du$$

Conditional CDFs

$$F_{Y|X}(y|x) = \int_{-\infty}^{y} \frac{f(x,v)}{f_X(x)} dv$$

Conditional PDF of Y given $X = x$.

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

Joint PDF

Marginal PDF

Conditional PDF and CDF for continuous R.V.s

Conditional PMF $p_{X|Y}(x|y) = \dfrac{p_{X,Y}(x,y)}{p_Y(y)}$ $\xrightarrow{\text{Divide the histograms into small bins}}$ $f_{X|Y}(x|y) = \dfrac{f(x,y)}{f_Y(y)}$ Conditional PDF

Conditional PDF $f_{X|Y}(x|y) = \dfrac{f(x,y)}{f_Y(y)}$

Multiple lhs by Δ_x and rhs by $\Delta_x \Delta_y / \Delta_x$.

$\Delta_x, \Delta_y = \varepsilon$

$$f_{X|Y}(x|y)\Delta_x = \frac{f(x,y)\Delta_x\Delta_y}{f_Y(y)\Delta_y} \approx \frac{P\{x \le X \le x + \Delta_x, y \le Y \le y + \Delta_y\}}{P\{y \le Y \le y + \Delta_y\}}$$
$$= P\{x \le X \le x + \varepsilon | y \le Y \le y + \varepsilon\}$$

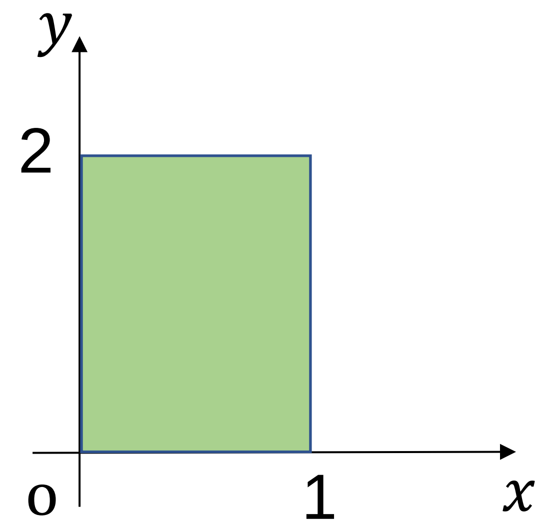
Consider a large range of X.

$$P\{X \le x | Y = y\} = F_{X|Y}(x|y) = \lim_{\varepsilon \to 0^+} P\{X \le x | y \le Y \le y + \varepsilon\} = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)} du$$

Conditional CDF

Ex. Let the joint PDF of (X, Y) as

$$f(x,y) = \begin{cases} x^2 + \frac{1}{3}xy, & 0 < x < 1, \\ & 0 < y < 2, \\ 0, & \text{otherwise} \end{cases}$$

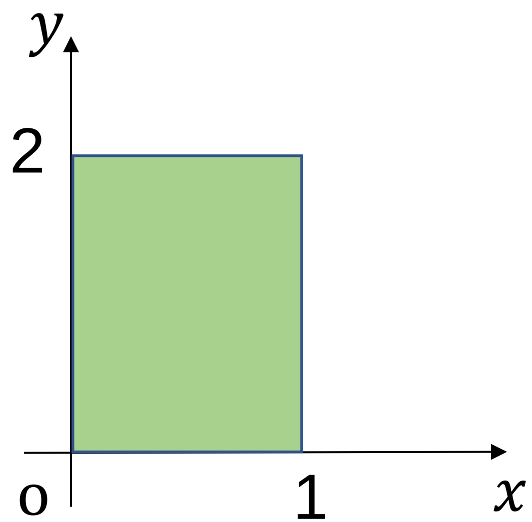


(1) $f_X(x)$, $f_Y(y)$;

(2) $f_{X|Y}(x|y)$, $f_{Y|X}(y|x)$.

Ex. Let the joint PDF of (X, Y) as

$$f(x,y)=\begin{cases} x^2+\frac{1}{3}xy, & 0<x<1,\ 0<y<2, \\ 0, & \text{otherwise} \end{cases}$$



(1) $f_X(x)$, $f_Y(y)$;

(2) $f_{X|Y}(x|y)$, $f_{Y|X}(y|x)$.

Sol. (1) $f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy=\int_0^2\left(x^2+\frac{1}{3}xy\right)dy$

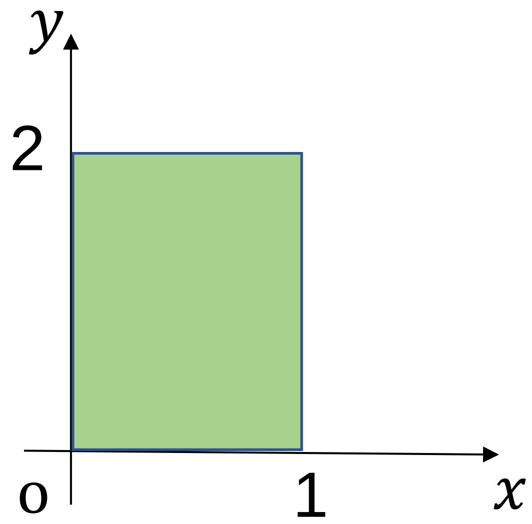
$$=2x^2+\frac{2}{3}x$$

$$f_X(x)=\begin{cases} 2x^2+\frac{2}{3}x, & 0<x<1 \\ 0, & \text{otherwise} \end{cases}$$

Similarly, $f_Y(y)=\begin{cases} \frac{1}{3}+\frac{1}{6}y, & 0<y<2 \\ 0, & \text{otherwise} \end{cases}$

Ex. Let the joint PDF of (X, Y) as

$$f(x,y) = \begin{cases} x^2 + \frac{1}{3}xy, & 0 < x < 1, \\ & 0 < y < 2, \\ 0, & \text{otherwise} \end{cases}$$



(1) $f_X(x)$, $f_Y(y)$;

(2) $f_{X|Y}(x|y)$, $f_{Y|X}(y|x)$.

(2) $0 < y < 2$,

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(Y)} = \begin{cases} \frac{6x^2 + 2xy}{2 + y}, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

$0 < x < 1$,

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(X)} = \begin{cases} \frac{3x + y}{6x + 2}, & 0 < y < 2 \\ 0, & \text{otherwise} \end{cases}$$

Ex. Assume the joint PDF of (X, Y) is

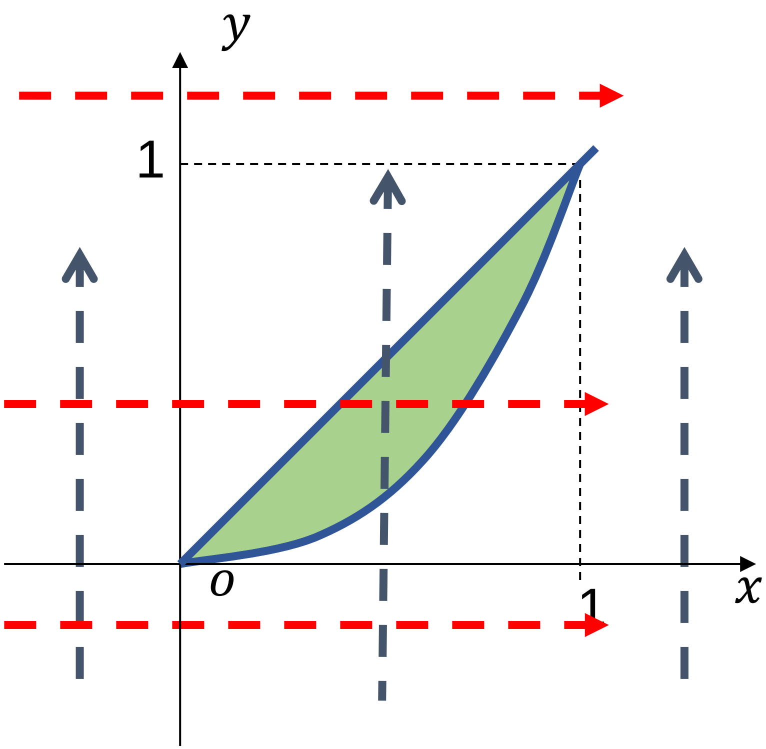
$$f(x, y) = \begin{cases} 24xy, & x^2 \le y \le x \\ 0, & \text{otherwise} \end{cases}$$

Find the conditional PDF of $Y|X$ and $X|Y$.

Ex. Assume the joint PDF of (X, Y) is

$$f(x,y) = \begin{cases} 24xy, & x^2 \le y \le x \\ 0, & \text{otherwise} \end{cases}$$

Find the conditional PDF of $Y|X$ and $X|Y$.



Sol.

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy$$

$$= \begin{cases} \int_{x^2}^{x} 24xy\,dy, & 0 \le x \le 1, \\ 0, & \text{otherwise} \end{cases}$$

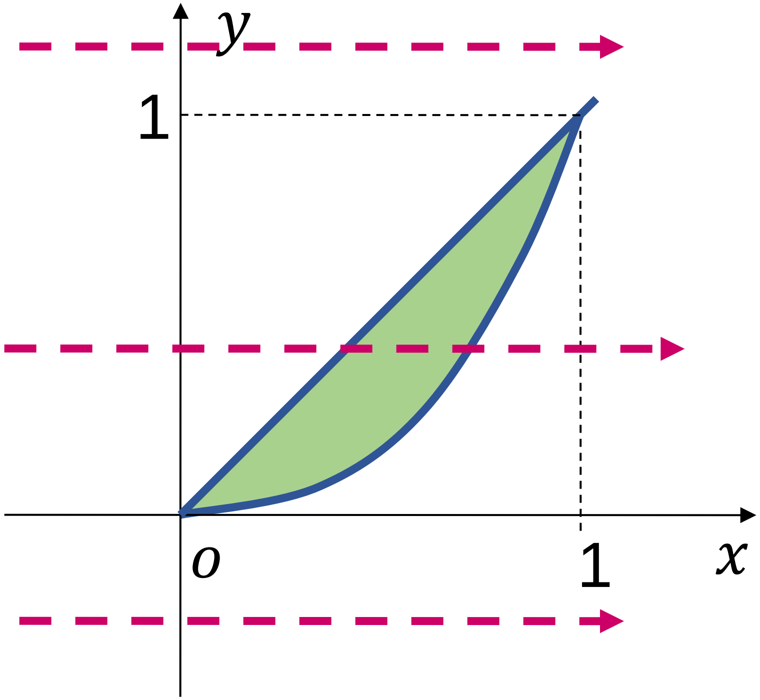
$$= \begin{cases} 12x^3(1-x^2), & 0 \le x \le 1, \\ 0, & \text{otherwise} \end{cases}$$

Given any x $(0 < x < 1)$,

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \begin{cases} \dfrac{2y}{x^2(1-x^2)}, & x^2 \le y \le x \\ 0, & \text{otherwise} \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$$

$$= \begin{cases} 12y(y - y^2), & 0 \le y \le 1, \\ 0, & \text{otherwise} \end{cases}$$



Given any y $(0 < y < 1)$,

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \begin{cases} \frac{2x}{y - y^2}, & y \le x \le \sqrt{y} \\ 0, & \text{otherwise} \end{cases}$$