Chapter 4 Bivariate Distributions (二元分布)

Section 4.3 CONDITIONAL DISTRIBUTIONS

➢ Motivation

Let X and Y have the joint probability mass function $f(x,y):\mathrm{D}\to[0,1]$.
The marginal *pmf* of X and Y are

$$f_X(x):\ D_X\to[0,1] \text{ and } f_Y(y):\ D_Y\to[0,1].$$

$D_X=\{all\ possible\ values\ of\ X\ in\ D\}$, $D_Y=\{all\ possible\ values\ of\ Y\ in\ D\}$.
By definition,

$$f(x,y)=P(X=x,Y=y)\triangleq P(\{X=x,Y=y\}).$$
$$f_X(x)=P(X=x)\triangleq P(\{X=x\})=\textstyle\sum_{y\in D_Y}f(x,y).$$
$$f_Y(y)=P(Y=y)\triangleq P(\{Y=y\})=\textstyle\sum_{x\in D_X}f(x,y).$$

Let $A=\{X=x\},\ B=\{Y=y\}$, $A\cap B=\{X=x\}\cap\{Y=y\}\triangleq\{X=x,Y=y\}$.
Then recall the conditional probability of event A given event B is

$$P(A|B)=\frac{P(A\bigcap B)}{P(B)}=\frac{f(x,y)}{f_Y(y)} \text{ (under the assumption } f_Y(y)>0).$$

Definition 4.3-1 [conditional probability mass functiom]

Conditional pmf of *X* given $Y=y$ is defined by

$$g(x|y) = \frac{f(x,y)}{f_Y(y)}, \qquad \text{provided that } f_Y(y) > 0$$

Similarly, **conditional pmf of *Y*** given that $X=x$ is defined

$$h(y|x) = \frac{f(x,y)}{f_X(x)}, \qquad \text{provided that } f_X(x) > 0$$

Example 1:

Let the joint *pmf* of X and Y be defined by

$$f(x,y) = \frac{x+y}{21}, \qquad x = 1,2,3, \qquad y = 1,2.$$

We have shown

$$f_X(x) = \sum_{y \in D_Y} f(x,y) = \sum_{y=1}^{2} \frac{x+y}{21} = \frac{2x+3}{21}, \qquad x = 1,2,3.$$

$$f_Y(y) = \sum_{x \in D_X} f(x,y) = \sum_{x=1}^{3} \frac{x+y}{21} = \frac{y+2}{7}, \qquad y = 1,2.$$

Then the conditional *pmf* of X given $Y = y$ is

$$g(x|y) = \frac{f(x,y)}{f_Y(y)} = \left(\frac{x+y}{21}\right) \Big/ \left(\frac{y+2}{7}\right) = \frac{x+y}{3(y+2)}, \qquad x = 1,2,3 \qquad y = 1,2.$$

and the conditional *pmf* of Y given $X = x$ is

$$h(y|x) = \frac{f(x,y)}{f_X(x)} = \left(\frac{x+y}{21}\right) \Big/ \left(\frac{2x+3}{21}\right) = \frac{x+y}{2x+3}, \qquad x = 1,2,3 \qquad y = 1,2.$$

- Conditional *pmf* is a well-defined *pmf*

① $h(y|x) \geq 0.$

② $\sum_{y \in D_Y} h(y|x) = \sum_{y \in D_Y} \frac{f(x,y)}{f_X(x)} = \frac{\sum_{y \in D_Y} f(x,y)}{f_X(x)} = \frac{f_X(x)}{f_X(x)} = 1.$

$f(x,y) = 0$ if $(x,y) \notin D.$

- Conditional mean and conditional variance

Let $u(Y)$ be a function of Y. Then the **conditional expectation** of $u(Y)$ is given by

$$E(u(Y)|X=x) = \sum_{y \in D_Y} u(y)h(y|x).$$

When $u(Y) = Y,$

$$E(Y|X=x) = \sum_{y \in D_Y} yh(y|x).$$

Conditional mean

When $u(Y) = \left[Y - E(Y|X=x)\right]^2,$

$$Var(Y|X=x) \triangleq E\left\{\left[Y - E(Y|X=x)\right]^2 \middle| X=x\right\} = \sum_{y \in D_Y} \left[y - E(Y|X=x)\right]^2 h(y|x).$$

Conditional variance

Example 1 (c.n.t.)

$$E(Y|X=3) = \sum_{y \in D_Y} yh(y|3) = \sum_{y=1}^{2} y \cdot \frac{y+3}{9} = \frac{14}{9}.$$

$$Var(Y|X=3) = \sum_{y \in D_Y} \left[y - E(Y|X=3)\right]^2 h(y|3) = \sum_{y=1}^{2} \left(y - \frac{14}{9}\right)^2 \cdot \frac{y+3}{9} = \frac{20}{81}.$$

Section 4.4 Bivariate Distribution of continuous type

- Idea: (bivariate) ***discrete*** RV → (bivariate) ***continuous*** RV

Definition 4.4-1 [joint probability density function (joint *pdf*)]

Let X and Y be two **continuous** RVs. The function $f(x, y)$: D → $[0, +\infty)$ is called the **joint probability density function** (**joint *pdf***) of X *if*:

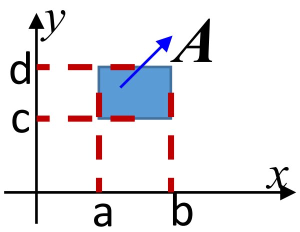
① $f(x,y) \geq 0; \quad (x,y) \in D.$

② $\iint_D f(x,y)dxdy = 1;$

③ $P\left[(X,Y) \in A\right] \triangleq P(\{(x,y) \in A\}) = \iint_A f(x,y)dxdy, \quad A \subseteq D.$

Motivation: The outcome is a tuple of 2 scalars whose range are intervals or union of intervals

Remark :



- Very often, we extend the definition domain of $f(x,y)$ from D to $R \times R$ by letting $f(x,y) = 0$ for $(x,y) \notin D$ and thus $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = 1.$
- In this course, we only consider a special space A in the ③ of the definiton: A is ***rectangular*** with its line segments parallel to the coordinate axis. In this case, $A = \{(x,y) | a \leq x \leq b,\ c \leq y \leq d\}$. Then the double integral becomes

$$P((x,y) \in A) = \int_a^b \int_c^d f(x,y)dydx.$$

Remark3 : Joint pdf can be seen as an extension of joint pmf by extending the 'summation' to 'integral'.

- mass → density
- pmf → pdf
- joint pmf → joint pdf
- marginal pmf → marginal pdf
- conditional pmf→ conditional pdf

summation → integral

Mean

Variance

Covariance

Correlation

Definition 4.4-2 [Marginal *pdf*]

The **marginal probability density function of** X or Y is defined by

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy : D_X \to [0,+\infty)$$

$$x \in D_X = \{\text{all possible values of } x \text{ in } D\}.$$

$$f_X(x) : R \to [0,+\infty) \text{ by letting } f_X(x) = 0 \text{ for } x \notin D_X$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx : D_Y \to [0,+\infty)$$

$$y \in D_Y = \{\text{all possible values of } y \text{ in } D\}.$$

$$f_Y(y) : R \to [0,+\infty) \text{ by letting } f_Y(y) = 0 \text{ for } y \notin D_Y$$

Definition 4.4-3 [Mathematical expectation]

Let $u(X,Y)$ be a function of X and Y whose marginal pdf is given by $f(x,y)$. Thus the mathematical expectation of $u(X,Y)$ is defined by

$$E[u(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} u(x,y)f(x,y)dxdy$$

- When $u(X,Y)=X$,

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x,y)\,dx\,dy=\int_{-\infty}^{+\infty} xf_X(x)\,dx.$$

- When $u(X,Y)=(X-E(X))^2$,

$$Var(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(X-E(X))^2 f(x,y)\,dx\,dy=\int_{-\infty}^{+\infty}(X-E(X))^2 f_X(x)\,dx.$$

Example 1 (Page 155)

Let X and Y have the joint pdf $f(x,y)=\frac{4}{3}(1-xy)$ with $0\le x\le 1,\ 0\le y\le 1$.

Compute $f_X(x), f_Y(y), E(X)$ and $Var(X)$.

Solution:

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_0^1\frac{4}{3}(1-xy)dy=\frac{4}{3}-\frac{4}{3}(x)(\frac{1}{2}y^2)\Big|_0^1=\frac{4}{3}(1-\frac{1}{2}x)$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_0^1\frac{4}{3}(1-xy)dx=\frac{4}{3}(1-\frac{1}{2}y)\quad\leftarrow \text{Due to the symmetry.}$$

$$E(X)=\int_{-\infty}^{+\infty} xf_X(x)dx=\int_0^1 x\frac{4}{3}(1-\frac{1}{2}x)dx=\frac{4}{3}\left[\frac{1}{2}x^2\Big|_0^1-\frac{1}{6}x^3\Big|_0^1\right]=\frac{4}{9}$$

$$Var(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 f_X(x)dx=\int_0^1(x-\frac{4}{9})^2\frac{4}{3}(1-\frac{1}{2}x)dx=\frac{13}{162}.$$

You should verify the details by yourself!

Quiz

Let X and Y have the joint pdf $f(x,y)=\frac{3}{2}x^2(1-|y|)$ with $-1<x<1,\ \ -1<y<1$.

$A=\{(x,y)|0<x<1, 0<y<x\}$. Compute $E(X)$ and $P(A)$.

Solution :

$$f_X(x)=\int_{-1}^{1}\frac{3}{2}x^2(1-|y|)dy=\int_{0}^{1}\frac{3}{2}x^2(1-y)dy+\int_{-1}^{0}\frac{3}{2}x^2(1+y)dy$$

$$=\frac{3}{2}x^2\left[y-\frac{1}{2}y^2\right]_0^1+\frac{3}{2}x^2\left[y+\frac{1}{2}y^2\right]_{-1}^0=\frac{3}{2}x^2\times\frac{1}{2}+\frac{3}{2}x^2\times\frac{1}{2}=\frac{3}{2}x^2.$$

$$E(X)=\int_{-1}^{1}xf_X(x)dx=\int_{-1}^{1}\frac{3}{2}x^3dx=\left[\frac{3}{8}x^4\right]_{-1}^{1}=0.$$

We have two ways to compute $P(A)$:

$$P(A)=\iint_A f(x,y)dxdy=\int_0^1\int_0^x\frac{3}{2}x^2(1-|y|)dydx=\int_0^1\int_0^x\frac{3}{2}x^2(1-y)dydx=\int_0^1\left[\frac{3}{2}x^2(y-\frac{1}{2}y^2)\right]_0^x dx$$

$$=\int_0^1(\frac{3}{2}x^3-\frac{3}{4}x^4)dx=\left[\frac{3}{8}x^4-\frac{3}{20}x^5\right]_0^1=\frac{9}{40}.$$

$$P(A)=\iint_A f(x,y)dydx=\int_0^1\int_y^1\frac{3}{2}x^2(1-|y|)dxdy=\int_0^1\left[\frac{x^3}{2}(1-|y|)\right]_y^1 dy=\int_0^1(\frac{1}{2}-\frac{y^3}{2})(1-|y|)dy$$

$$=\int_0^1\left[\frac{y^4}{2}-\frac{y^3}{2}-\frac{y}{2}+\frac{1}{2}\right]dy=\left[\frac{y^5}{10}-\frac{y^4}{8}-\frac{y^2}{4}+\frac{1}{2}y\right]_0^1=\frac{1}{10}-\frac{1}{8}-\frac{1}{4}+\frac{1}{2}=\frac{9}{40}.$$

Definition 4.4-4 [independent Continuous Variables]

Two continuous variables X and Y are **independent** if and only if,

$$f(x,y)=f_X(x)f_Y(y), \qquad x\in D_X, y\in D_Y$$

Otherwise, X and Y are said to be **dependent**.

Example 1 (Revisited)

Since $f(x,y)=\frac{4}{3}(1-xy)\neq\left[\frac{4}{3}(1-\frac{1}{2}x)\right]\left[\frac{4}{3}(1-\frac{1}{2}y)\right]=f_X(x)f_Y(y)$, X and Y are *dependent*.

Definition 4.4-5 [Covariance and correlation coefficient]

The **covariance** of X and Y is given by

$$Cov(X,Y)=E\left[(X-E(X))(Y-E(Y))\right]=E(XY)-E(X)E(Y),$$

$$\text{where } E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)dxdy$$

The correlation cofficients is defined as

$$\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}.$$

Definition 4.4-6 [conditional probability density function]

Let X and Y have the joint pdf $f(x, y)$ and marginal pdfs are $f_X(x)$ and $f_Y(y)$. Then the **conditional pdf**, **mean**, and **variance** of Y, given that $X=x$, are

$$h(y|x)=\frac{f(x,y)}{f_X(x)} \text{ for } f_X(x)>0,$$

$$E(Y|X=x)=\int_{-\infty}^{+\infty} yh(y|x)dy,$$

$$Var(Y|X=x)=E(\left[Y-E(Y|X=x)\right]^2|X=x)=\int_{-\infty}^{+\infty}(y-E(Y|X=x))^2h(y|x)dy$$

$$=E(Y^2|X=x)-\left[E(Y|X=x)\right]^2$$

Example 2 (Page 160)

Let X and Y be continuous RVs that have

$$f(x,y)=2, 0\le x\le y\le 1,$$

Question:

(a) Sketch the support of X and Y.

(b) Compute the marginal pmfs $f_X(x)$ and $f_Y(y)$.

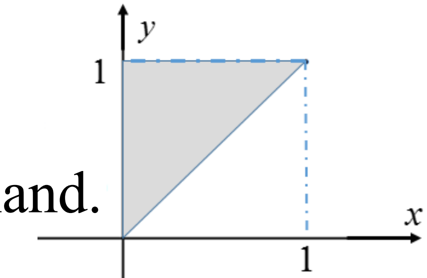
(c) Compute the conditional pdf, conditional mean, conditional variance of Y, given $X=x$.

(d) Compute $P(\frac{3}{4}\le Y\le\frac{7}{8}\Big|X=\frac{1}{4})$.

Example 2 (c.n.t.)

Solution :

(a) The graph for the support of X and Y is listed righthand.



(b) $$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_x^1 f(x,y)dy=2(1-x) \qquad 0\le x\le 1.$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_0^y f(x,y)dy=2y \qquad 0\le y\le 1.$$

(c) $$h(y|x)=\frac{f(x,y)}{f_X(x)}=\frac{2}{2(1-x)}=\frac{1}{1-x}, \quad 0\le x\le y\le 1.$$

$$E(Y|X=x)=\int_{-\infty}^{+\infty} yh(y|x)dy=\int_x^1 y\frac{1}{1-x}dy=\frac{1}{1-x}\left[\frac{1}{2}y^2\right]_x^1=\frac{1}{2}(x+1).$$

$$Var(Y|X=x)=\int_{-\infty}^{+\infty}\left[y-E(Y|X=x)\right]^2 h(y|x)dy$$

$$=\int_x^1\left[y-\frac{1}{2}(x+1)\right]^2\frac{1}{1-x}dy=\frac{1}{1-x}\left[y-\frac{1}{2}(1+x)\right]^3\Big|_x^1=\frac{(1-x)^2}{12}.$$

(d) $$P\left(\frac{3}{4}\le Y\le\frac{7}{8}\middle|X=\frac{1}{4}\right)=\int_{3/4}^{7/8} h\left(y\middle|\frac{1}{4}\right)dy=\int_{3/4}^{7/8}\frac{1}{1-1/4}dy=\frac{1}{8}\times\frac{4}{3}=\frac{1}{6}.$$

Section 4.5 Bivariate Normal Distribution

Definition 4.5-1 [Bivariate Normal]

Let X and Y be two continuous RVs and have the joint pdf

$$f(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-p^2}}\exp\left[-\frac{1}{2}q(x,y)\right]$$

where $q(x,y)=\frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2-2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\geq 0$

$\mu_X=E(X), \mu_Y=E(Y), \sigma_X=\sqrt{Var(X)}, \sigma_Y=\sqrt{Var(Y)}$, ρ is the **correlation coefficient**.

Then X and Y are said o be **bivariate normal distributed.**

- Property:

Given that $Y=y$ is normal distribution, The probability distribution of X with **mean** $\mu_X+\frac{\sigma_X}{\sigma_Y}\rho(y-\mu_Y)$ and **variance** $(1-\rho^2)\sigma_X^2$ is given by

$$X|Y=y\sim N(\mu_X+\frac{\sigma_X}{\sigma_Y}\rho(y-\mu_Y),(1-\rho^2)\sigma_X^2).$$

Similarly, $$Y|X=x\sim N(\mu_Y+\frac{\sigma_Y}{\sigma_X}\rho(x-\mu_X),(1-\rho^2)\sigma_Y^2).$$

Example 1 (Page 165)

Observe a group of college students, Let X and Y denote the grade points in high school and the first year in college have a bivariate normal distribution with parameters

$\mu_X = 2.9 \qquad \mu_Y = 2.4 \qquad \sigma_X = 0.4 \qquad \sigma_Y = 0.5 \qquad \rho = 0.8$

Compute $P(2.1 < Y < 3.3)$ and $P(2.1 < Y < 3.3 \mid X = 3.2)$

Solution:

$$P(2.1 < Y < 3.3) = P\left(\frac{2.1-2.4}{0.5} < \frac{Y-2.4}{0.5} < \frac{3.3-2.4}{0.5}\right) = \Phi(1.8) - \Phi(-0.6) = 0.69$$

Note that $Y \mid X = x \sim N\left(\mu_Y + \frac{\sigma_Y}{\sigma_X}\rho(x-\mu_X), (1-\rho^2)\sigma_Y^2\right)$,

when $X = 3.2$, $Y \mid X = 3.2 \sim N(2.7, 0.09)$,

$$P(2.1 < Y < 3.3 \mid X = 3.2) = P\left(\frac{2.1-2.7}{\sqrt{0.09}} < \frac{Y-2.7}{\sqrt{0.09}} < \frac{3.3-2.7}{\sqrt{0.09}}\right)$$

$$= \Phi(2) - \Phi(-2) = 0.95.$$

➢ Geometrical interpretation

Let $z = f(x, y)$ and draw it in x - y - z 3-dimension coordinate system.

① $z = f(x_0, y) = f_X(x_0)h(y|x_0)$

$\rightarrow$ intersection of the surface $z = f(x, y)$ with planes parallel to yz-plate.

② $z = f(x, y_0) = f_Y(y_0)h(x|y_0) = f_Y(y_0)h(x|y_0)$

$\rightarrow$ intersection of the surface $z = f(x, y)$ with planes parallel to xz-plate

They are all bell-shaped, in other words, they have he shape of normal distribution pdf.

③ let $z_0 > 0$ and $z_0 < \dfrac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$

Then the intersection of the surface $z = f(x, y)$ with the plate $z = z_0$ which is parallel to xy-plate is

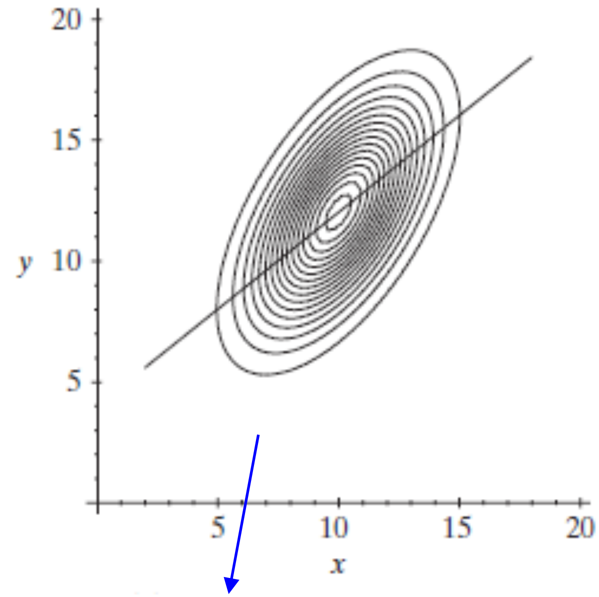
$$\exp\left[-\frac{1}{2}q(x,y)\right] = z_0 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2} \qquad \rightarrow ellipse$$

Taking logarithm yields

$$\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 = -2(1-\rho^2)\ln(z_0 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2})$$

➢ Geometrical interpretation (c.n.t.)

For different z_0, we can draw the ellipse on the xy plate - Pevel curves or contours.



Contours for bivariate normal

Theorem 4.5-1

If X and Y have a bivariate normal distribution with correlation coefficient ρ, then X and Y are independent if and only if $\rho = 0$.