3.2 边缘分布

一、边缘分布函数

二、离散型随机变量的边缘分布律

三、连续型随机变量的边缘分布

一、边缘分布函数

问题: 已知 (X,Y) 的分布, 如何确定 X,Y 的分布?

$$F(x,y) = P\{X \le x, Y \le y\}, \quad F(x) = P\{X \le x\},$$

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

(X,Y) 关于 X 的边缘分布函数.

定义 设 $F(x, y)$ 为随机变量 (X, Y) 的分布函数，

则 $F(x, y) = P\{X \le x, Y \le y\}$.

令 $y \to +\infty$, 称 $P\{X \le x\} = P\{X \le x, Y < +\infty\} = F(x, +\infty)$

为随机变量 (X, Y) 关于 X 的边缘分布函数.

记为 $F_X(x) = F(x, +\infty)$.

同理令 $x \to +\infty$,

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

为随机变量 (X, Y) 关于 Y 的边缘分布函数.

例1.设（X,Y）的分布函数为

$$F(x,y)=\frac{1}{\pi^2}\left(\frac{\pi}{2}+\arctan\frac{x}{2}\right)\left(\frac{\pi}{2}+\arctan\frac{y}{3}\right)$$

1.关于X和Y的边缘分布函数　　2.求 $P(X>2)$

解: $1.F_X(x)=F(x,+\infty)=\frac{1}{\pi^2}\left(\frac{\pi}{2}+\arctan\frac{x}{2}\right)\cdot\pi$

$$=\frac{1}{2}+\frac{1}{\pi}\arctan\frac{x}{2}$$

$$F_Y(y)=F(+\infty,y)=\frac{1}{2}+\frac{1}{\pi}\arctan\frac{y}{3}$$

$$2.P(X>2)=1-P(X\le 2)=1-F_X(2)=\frac{1}{4}$$

二、离散型随机变量的边缘分布律

定义 设二维离散型随机变量(X,Y)的联合分布律为

$$P\{X=x_i, Y=y_j\}=p_{ij}, i,j=1,2,\cdots.$$

记

$$p_{i\bullet}=\sum_{j=1}^{\infty}p_{ij}=P\{X=x_i\},\quad i=1,2,\cdots,$$

$$p_{\bullet j}=\sum_{i=1}^{\infty}p_{ij}=P\{Y=y_j\},\quad j=1,2,\cdots,$$

分别称 $p_{i\bullet}\ (i=1,2,\cdots)$ 和 $p_{\bullet j}\ (j=1,2,\cdots)$ 为(X,Y)关于 X 和关于 Y 的边缘分布律.

$Y \backslash X$	x_1	x_2	$\cdots$	x_i	$\cdots$
y_1	p_{11}	p_{21}	$\cdots$	p_{i1}	$\cdots$
y_2	p_{12}	p_{22}	$\cdots$	p_{i2}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	
y_j	p_{1j}	p_{2j}	$\cdots$	p_{ij}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	

$$P\{X = x_i\} = \sum_{j=1}^{+\infty} p_{ij},\ i = 1,2,\cdots;\quad P\{Y = y_j\} = \sum_{i=1}^{+\infty} p_{ij},\ j = 1,2,\cdots.$$

因此离散型随机变量关于X和Y的边缘分布函数分别为

$$F_X(x) = F(x,+\infty) = \sum_{x_i \le x} \sum_{j=1}^{+\infty} p_{ij},\quad F_Y(y) = F(+\infty, y) = \sum_{y_j \le y} \sum_{i=1}^{+\infty} p_{ij}.$$

例2 已知下列分布律求其边缘分布律.

$Y \backslash X$	0	1
0	$\frac{12}{42}$	$\frac{12}{42}$
1	$\frac{12}{42}$	$\frac{6}{42}$

解

$Y \backslash X$	0	1	$p_{\cdot j} = P\{Y = y_j\}$
0	$\frac{12}{42}$ +	$\frac{12}{42}$	$\frac{4}{7}$
1	+ $\frac{12}{42}$ +	+ $\frac{6}{42}$	$\frac{3}{7}$
$p_{i\cdot} = P\{X = x_i\}$	$\frac{4}{7}$	$\frac{3}{7}$	1

注意

$$\text{联合分布} \underset{\not\longleftarrow}{\longrightarrow} \text{边缘分布}$$

例3. 求(X,Y)关于X和Y的边缘分布律。

X \ Y	-1	0	4
1	0.17	0.05	0.21
3	0.04	0.28	0.25

解：

X \ Y	-1	0	4	$P\{X=x_i\}=p_{i.}$
1	0.17	0.05	0.21	0.43
3	0.04	0.28	0.25	0.57
$P\{Y=y_j\}=p_{.j}$	0.21	0.33	0.46	1

因此关于X的边缘分布律为

X	1	3
p	0.43	0.57

关于*Y*的边缘分布律为

Y	-1	0	4
p	0.21	0.33	0.46

三、连续型随机变量的边缘分布

定义 对于连续型随机变量(X,Y)，设它的概率密度为$f(x,y)$，由于

$$F_X(x)=F(x,+\infty)=\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty}f(u,v)\,\mathrm{d}v\right]\mathrm{d}u,$$

记
$$f_X(x)=\int_{-\infty}^{\infty}f(x,v)\,\mathrm{d}v,$$

称其为随机变量(X,Y)关于X的边缘概率密度

同理可得 Y 的边缘分布函数

$$F_Y(y)=F(+\infty,y)=\int_{-\infty}^{y}\left[\int_{-\infty}^{+\infty}f(u,v)\,\mathrm{d}u\right]\mathrm{d}v,$$

Y 的边缘概率密度

$$f_Y(y)=\int_{-\infty}^{+\infty}f(u,y)\,\mathrm{d}u.$$

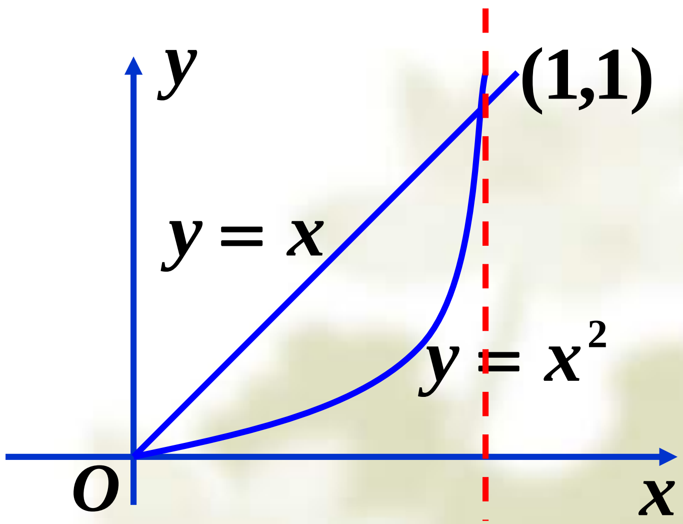
$$\boxed{\begin{aligned}f_X(x)&=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}y,\\ f_Y(y)&=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}x.\end{aligned}}$$

例4 设随机变量 X 和 Y 具有联合概率密度

$$f(x,y)=\begin{cases}6, & x^2\le y\le x,\\ 0, & 其他.\end{cases}$$

求边缘概率密度 $f_X(x)$, $f_Y(y)$.

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



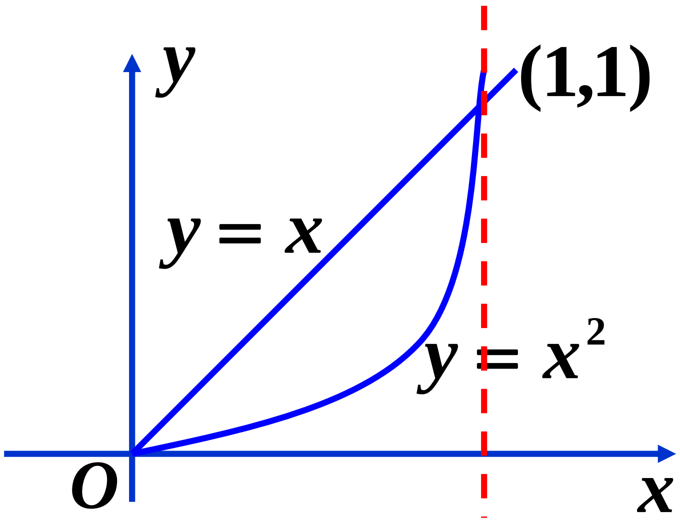
当 $0\le x\le 1$ 时,

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

$$=\int_{x^2}^{x}6\,\mathrm{d}y=6(x-x^2).$$

当 $x<0$ 或 $x>1$时,

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y=0.$$



因而得

$$f_X(x)=\begin{cases} 6(x-x^2), & 0\le x\le 1, \\ 0, & \text{其他}. \end{cases}$$

当$0\le y\le 1$时,

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x=\int_y^{\sqrt{y}} 6\,\mathrm{d}x=6(\sqrt{y}-y).$$

当 $y<0$ 或 $y>1$时,

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x=0.$$ 得 $$f_Y(y)=\begin{cases} 6(\sqrt{y}-y), & 0\le y\le 1, \\ 0, & \text{其他}. \end{cases}$$

例5 已知(X, Y)的概率密度为

$$\mathbf{f}(\mathbf{x},\boldsymbol{y})=\begin{cases}\mathbf{1}, & |\boldsymbol{y}|<\boldsymbol{x},\mathbf{0}<\boldsymbol{x}<\mathbf{1}\\ \mathbf{0} & \text{其他}\end{cases},$$

求边缘概率密度$f_X(x), f_Y(y)$

解

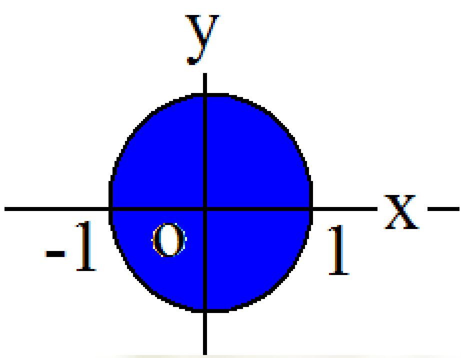
$$f_X(\mathbf{x})=\begin{cases}\mathbf{2}\boldsymbol{x}, & \mathbf{0}<\boldsymbol{x}<\mathbf{1}\\ \mathbf{0} & \text{其他}\end{cases}$$

$$f_Y(\mathbf{y})=\begin{cases}\mathbf{1}-|\boldsymbol{y}|, & |\boldsymbol{y}|<\mathbf{1}\\ \mathbf{0} & \text{其他}\end{cases}$$

例 6： 设二维随机变量(X,Y)服从单位圆$x^2+y^2\le 1$上的均匀分布,求(X,Y)关于 X 和 Y 的边缘分布.

解 （X，Y）的联合密度函数

$$f(x,y)=\begin{cases}\dfrac{1}{\pi} & x^2+y^2\le 1\\ 0 & 其它\end{cases}$$



则（X，Y）关于X的边缘密度函数

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\begin{cases}\displaystyle\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}dy & -1<x<1\\ 0 & 其它\end{cases}=\begin{cases}\dfrac{2}{\pi}\sqrt{1-x^2} & -1<x<1\\ 0 & 其它\end{cases}$$

（X，Y）关于Y的边缘密度函数

$$f_Y(y)=\begin{cases}\dfrac{2}{\pi}\sqrt{1-y^2} & -1<y<1\\ 0 & 其它\end{cases}$$

例7 设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

$$\cdot\exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

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

其中$\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$都是常数，且$\sigma_1>0,\sigma_2>0,$

$-1<\rho<1.$

试求二维正态随机变量的边缘概率密度.

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}},\quad f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}\mathrm{e}^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}},\quad -\infty<x,y<+\infty.$$

解 $f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y,$ 由于

$$\frac{(y-\mu_2)^2}{\sigma_2^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(x-\mu_1)^2}{\sigma_1^2}$$

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

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

令 $t=\dfrac{1}{\sqrt{1-\rho^2}}\left(\dfrac{y-\mu_2}{\sigma_2}-\rho\dfrac{x-\mu_1}{\sigma_1}\right),$

则有 $$f_X(x)=\frac{1}{2\pi\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t,$$

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

同理可得

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

二维正态分布的两个边缘分布都是一维正态分布,

并且都不依赖于参数 ρ.

令 (X,Y) 的联合密度函数为

$$f(x,y)=\frac{1}{2\pi}\mathrm{e}^{-\frac{x^2+y^2}{2}}(1+\sin x\sin y),$$

虽然正态分布（X，Y）不服从二维正态分布

$$f_X(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}},\quad f_Y(y)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}.$$

因此边缘分布均为正态分布的随机变量, 其联合分布不一定是二维正态分布.