第二节 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

四、小结

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？

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

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

(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，

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.

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

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

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

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}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	

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

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

例1 **已知下列分布律, 求其边缘分布律.**

$Y \backslash X$	0	1
0	$\dfrac{16}{49}$	$\dfrac{12}{49}$
1	$\dfrac{12}{49}$	$\dfrac{9}{49}$

解

$Y \backslash X$	0	1	$p_{\bullet j}$
0	$\frac{12}{42}$ +	$\frac{12}{42}$	$\frac{4}{7}$
	+	+	
1	$\frac{12}{42}$ +	$\frac{6}{42}$	$\frac{3}{7}$
$p_{i\bullet}$	$\frac{4}{7}$	$\frac{3}{7}$	1

注意

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

例2 一整数 N 等可能地在 $1,2,3,\cdots,10$ 十个值中取一个值. 设 $D=D(N)$ 是能整除 N 的正整数的的个数, $F=F(N)$ 是能整除 N 的素数的个数 (注意1不是素数). 试写出 D 和 F 的联合分布律, 并求边缘分布律.

解

N	1	2	3	4	5	6	7	8	9	10
D	1	2	2	3	2	4	2	4	3	4
F	0	1	1	1	1	2	1	1	1	2

由此得 D 和 F 的联合分布律与边缘分 布律:

$F \backslash D$	1	2	3	4	$P\{F=j\}$
0	$1/10$	0	0	0	$1/10$
1	0	$4/10$	$2/10$	$1/10$	$7/10$
2	0	0	0	$2/10$	$2/10$
$P\{D=i\}$	$1/10$	$4/10$	$2/10$	$3/10$	1

或将边缘分布律表示为

D	1	2	3	4
p_k	$1/10$	$4/10$	$2/10$	$3/10$

F	0	1	2
p_k	$1/10$	$7/10$	$2/10$

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

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

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

称其为随机变量 (X,Y) 关于 X 的边缘概率密度.

同样，Y也是一个连续型随机变量，其概率密度为

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

$$=\int_{-\infty}^{y}\left[\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x\right]\mathrm{d}y,$$

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

Y 的边缘概率密度.

例3 设随机变量 X 和 Y 具有联合概率密度

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

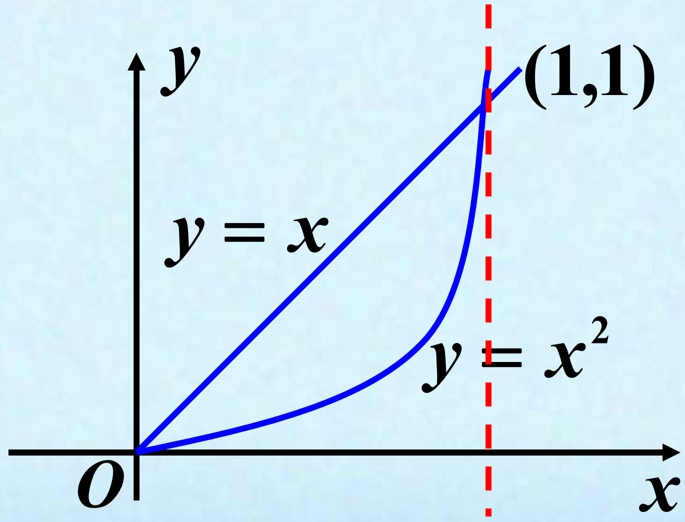
求边缘概率密度 $f_X(x), f_Y(y)$.

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

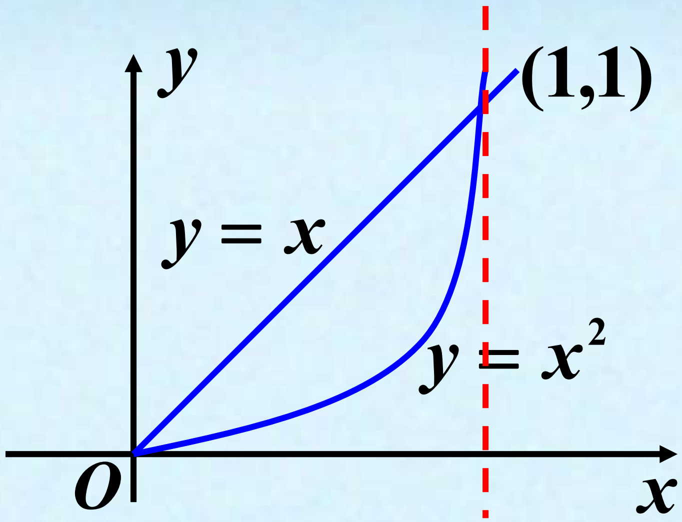
当 $0\le x\le 1$ 时,

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

$$=\int_{x^2}^{x} 6\,\mathrm{d}\,y$$



$= 6(x - x^2).$

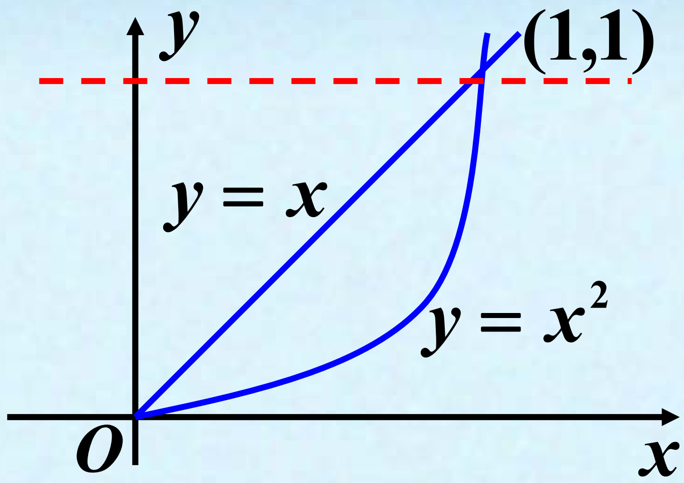


当 $x < 0$ 或 $x > 1$时,

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

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4 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 $\sigma_1>0,\sigma_2>0,$ $-1<\rho<1.$ 我们称 (X,Y)为服从参数为 $\mu_1,\mu_2,\sigma_1,$

σ_2, ρ的二维正态分布.(**这五个参数的意义将在下一章说明**), 记为$(X,Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. **试求二维正态随机变量的边缘概率密度.**

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

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

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

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

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

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

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

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

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.

请同学们思考

边缘分布均为正态分布的随机变量， 其联合分布一定是二维正态分布吗?

答 不一定. 举一反例以示证明.

解答: 令 (X,Y) 的联合密度函数为

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

显然, (X,Y) 不服从正态分布 , 但是

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

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.

四、小 结

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

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

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

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

联合分布 $\overset{\longrightarrow}{\not\longleftarrow}$ **边缘分布**

思考题

设(X,Y)在区域 $G=\{(x,y)\,|\,1<x^2+y^2<4\}$内服从均匀分布，试求 (X,Y)的边缘概率密度.

$$f_X(x)=\begin{cases}\dfrac{2}{3\pi}\left(\sqrt{4-x^2}-\sqrt{1-x^2}\right), & |x|\le 1,\\ \dfrac{2}{3\pi}\sqrt{4-x^2}, & 1<|x|<2,\\ 0, & |x|\ge 2,\end{cases}$$

作业

5，6，8，9