第二节 多维随机变量及其分布(2)

- 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、内容小结

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,

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.

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

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

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

因此得离散型随机变量关于X和Y的边缘分布函数分别为

$$F_X(x)=F(x,+\infty)=\sum_{x_i\le x}\sum_{j=1}^{\infty}p_{ij},$$

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

例1 已知下列分布律求其边缘分布律.

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

解

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

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

参见上页数据

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 $p(x, y)$, 由于

$$F_X(x) = F(x, +\infty) = \int_{-\infty}^{x} [\int_{-\infty}^{+\infty} p(x, y)\,dy]\,dx,$$

记

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

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

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

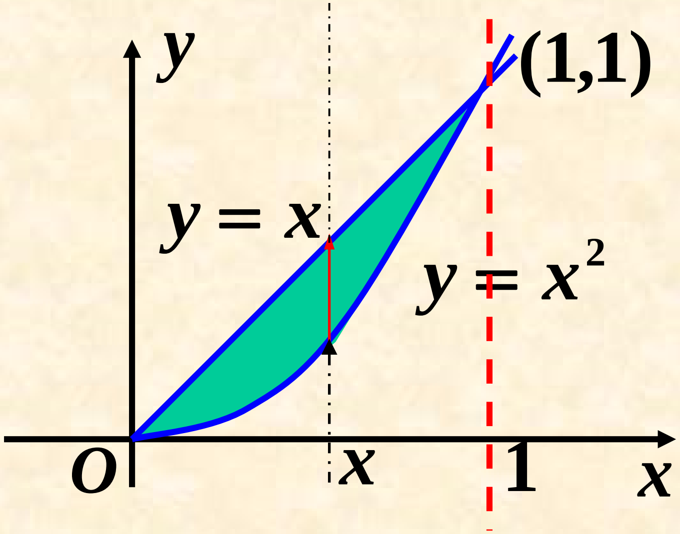
$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)\,\mathrm{d}x.$$

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

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

求边缘概率密度$p_X(x), p_Y(y)$.



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

当$0\le x\le 1$时,

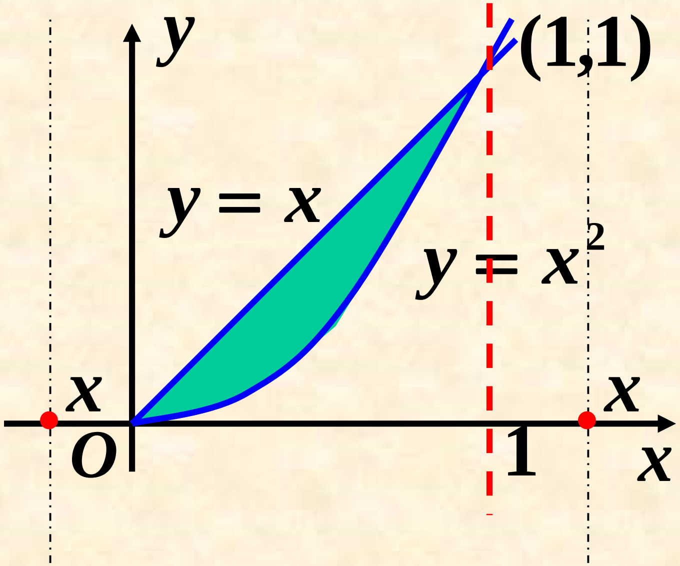
$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\,\mathrm{d}y$$

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

当 $x<0$ 或 $x>1$时,

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

$$=\int_{-\infty}^{+\infty} 0\,\mathrm{d}y=0.$$



因而得

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

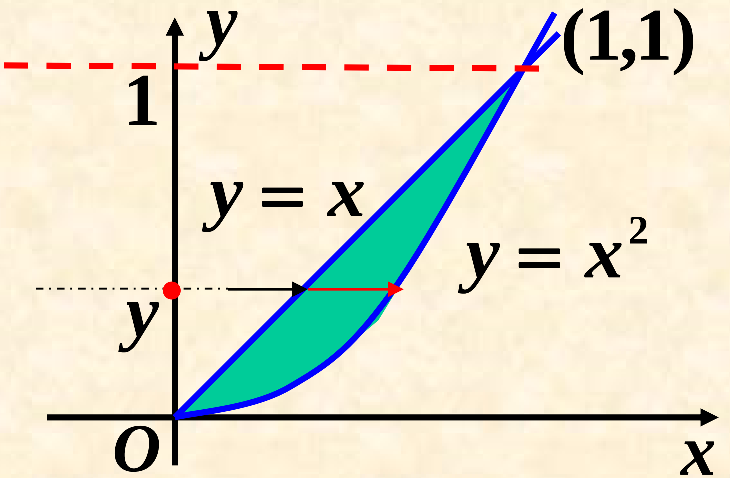
目录 上页 下页 返回 结束

当$0 \le y \le 1$时,

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

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

$$= 6(\sqrt{y} - y).$$



当 $y < 0$ 或 $y > 1$时, $P_Y(y) = \int_{-\infty}^{+\infty} P(x,y)\,\mathrm{d}x = 0.$

得 $$P_Y(y) = \begin{cases} 6(\sqrt{y} - y), & 0 \le y \le 1, \\ 0, & \text{其它}. \end{cases}$$

目录 上页 下页 返回 结束

例4 设二维随机变量 (X,Y) 的概率密度为

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

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

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

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.

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

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

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 $p_X(x)=\frac{1}{2\pi\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}\,dt,$

即

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

同理可得

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

注. 上述推导表明：

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.

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

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

但反之呢？

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.

令(X,Y)的联合密度函数为

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

显然,(X,Y)不服从正态分布,但是

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

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、内容小结

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

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

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

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

联合分布 $\underset{\cancel{\longleftarrow\hspace{2cm}}}{\xrightarrow{\hspace{2cm}}}$ 边缘分布

备份题

例2-1 设袋中有三个球，分别标有数字1, 2, 2. 从袋中任取一球后，不放回袋中，再从袋中任取一球.以X，Y分别表示第一，第二次取得的球上所标的数字，求(X,Y)的边缘分布律.

解

Y \ X	1	2	$p_{\cdot j}$
1	0	$\frac{1}{3}$	$\frac{1}{3}$
2	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
$p_{i\cdot}$	$\frac{1}{3}$	$\frac{2}{3}$	1

$\therefore (X,Y)$关于X的边缘分布律:

X	1	2
$p_{i\cdot}$	$\frac{1}{3}$	$\frac{2}{3}$

(X,Y)关于Y的边缘分布律:

Y	1	2
$p_{\cdot j}$	$\frac{1}{3}$	$\frac{2}{3}$

例3-1 设$(X,Y) \sim p(x,y)=\begin{cases} e^{-y}, & 0<x<y, \\ 0, & \text{其它}. \end{cases}$

求 (1) $p_X(x)$; (2) $P\{X+Y \le 1\}$.

解 当$x>0$时,

$$p_X(x)=\int_{-\infty}^{+\infty} p(x,y)\,\mathrm{d}y=\int_{x}^{+\infty} e^{-y}\,\mathrm{d}y=e^{-x}.$$

当 $x \le 0$ 时, $p_X(x)=\int_{-\infty}^{+\infty} p(x,y)\,\mathrm{d}y=0.$

故 $p_X(x)=\begin{cases} e^{-x}, & x>0, \\ 0, & \text{其它}. \end{cases}$

(2) $P\{X+Y\le 1\}$

$$=\iint\limits_{x+y\le 1} p(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

$$=\int_0^{\frac{1}{2}} dx\int_x^{1-x} e^{-y}\,\mathrm{d}y$$

$$=-\int_0^{\frac{1}{2}} [e^{-(1-x)}-e^{-x}]\,\mathrm{d}x = 1+e^{-1}-2e^{-\frac{1}{2}}.$$