西北工業大學
NORTHWESTERN POLYTECHNICAL UNIVERSITY

概率论与数理统计

徐爽

西北工业大学理学院概率统计教研室

第二节 多维随机变量及其分布(2)

- 五、边缘分布函数
- 六、离散型随机变量 的边缘分布律
- 七、连续型随机变量的边缘分布

五、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 $x\to+\infty$, $P\{X<+\infty,Y\le y\}=P\{Y\le y\}$

即$F_Y(y)=F(+\infty,y)$

为随机变量(X,Y)关于Y的边缘分布函数.

边缘分布也称为边沿分布或边际分布。

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

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

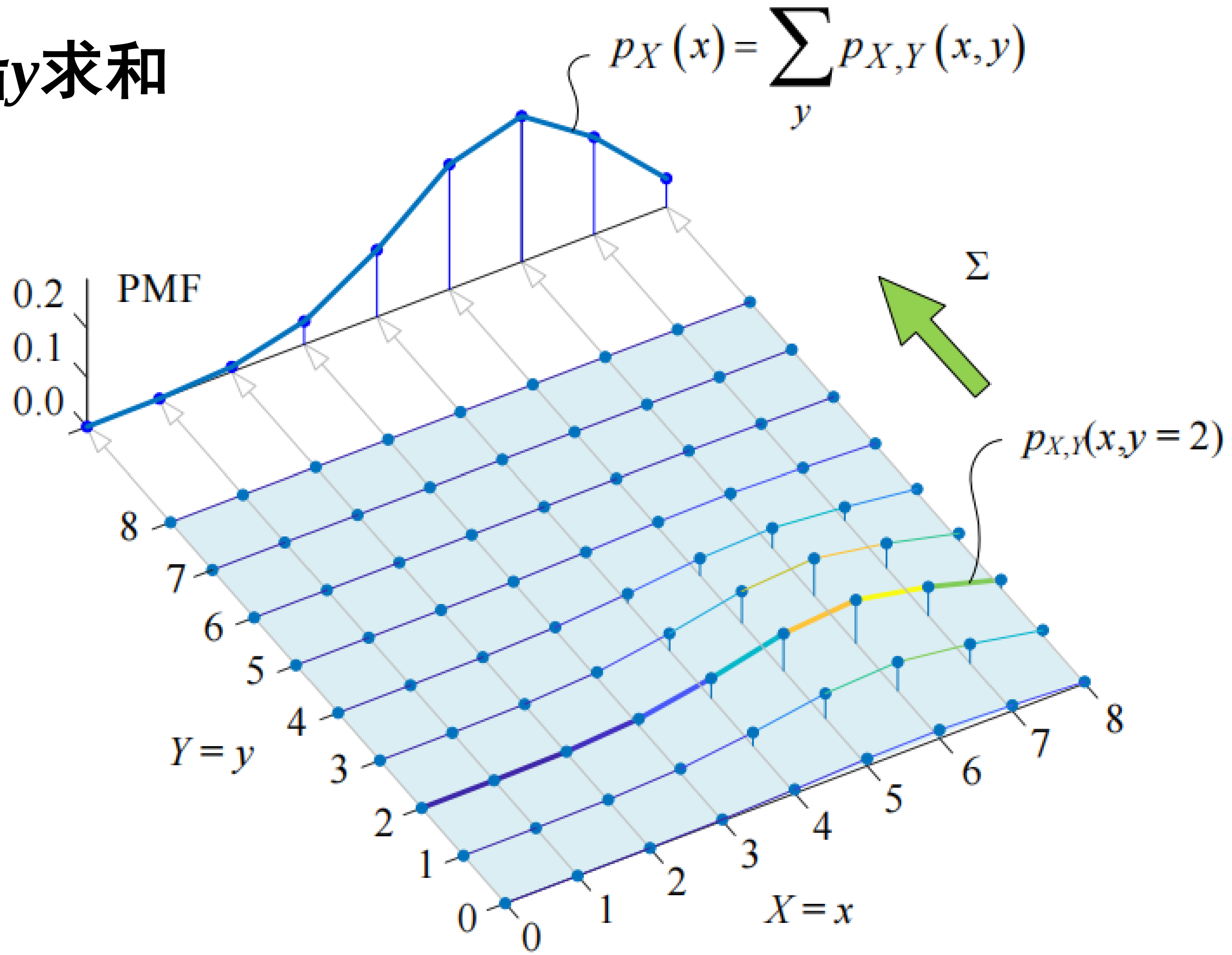
记

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

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

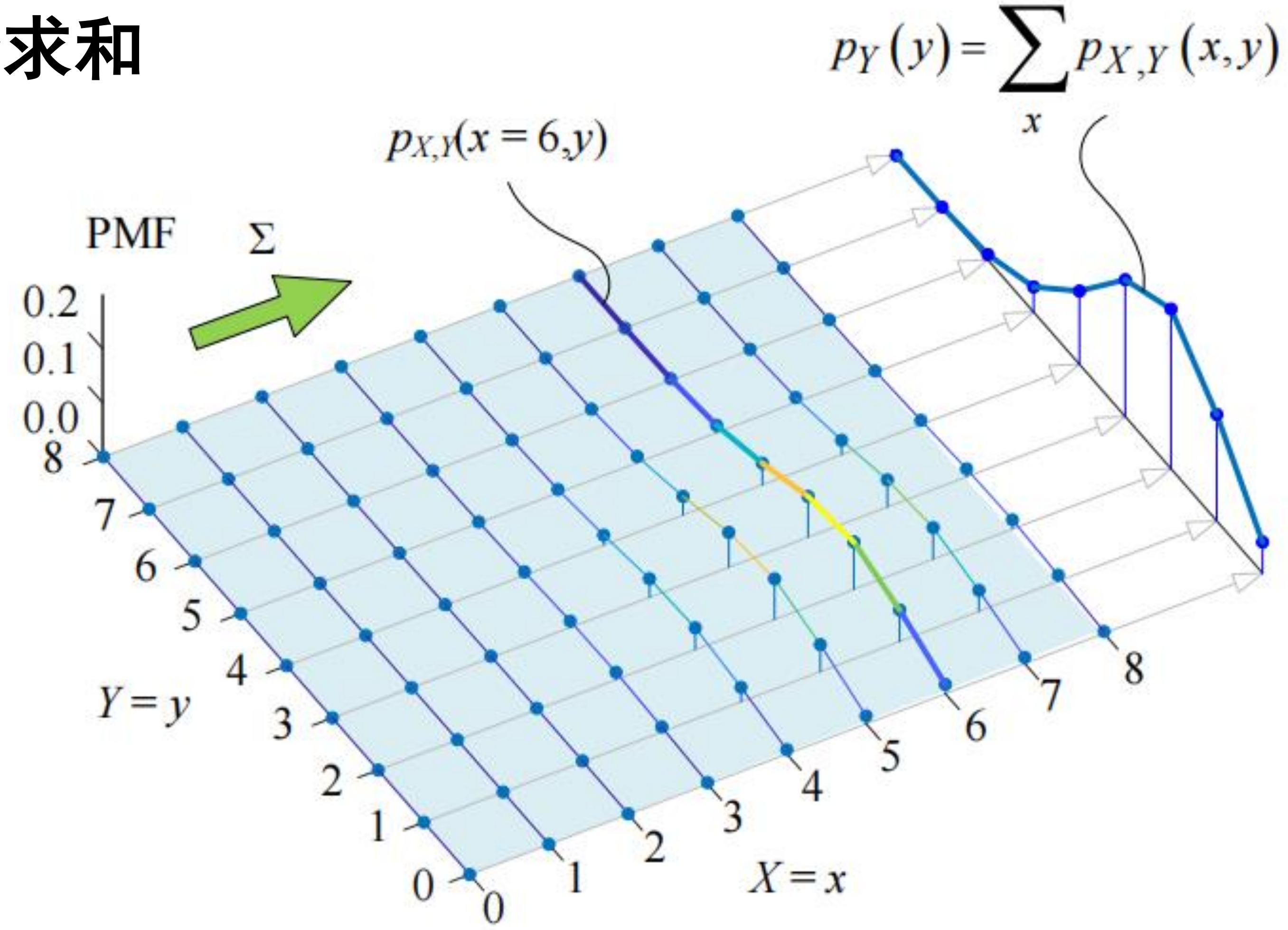
分别称 $p_{i\bullet}\ (i=1,2,\cdots)$ 和 $p_{\bullet j}\ (j=1,2,\cdots)$ 为(X,Y)关于X和关于Y的边缘分布律。

沿y求和



$p_X(x) = \sum_y p_{X,Y}(x,y)$
PMF
0.2
0.1
0.0
Σ
$p_{X,Y}(x,y=2)$
$Y = y$
$X = x$
0 1 2 3 4 5 6 7 8
0 1 2 3 4 5 6 7 8

沿x求和


$p_Y(y) = \sum_x p_{X,Y}(x,y)$
$p_{X,Y}(x=6,y)$
PMF
Σ
0.2
0.1
0.0
8
7
6
5
4
3
2
1
0
$Y=y$
0
1
2
3
4
5
6
7
8
$X=x$

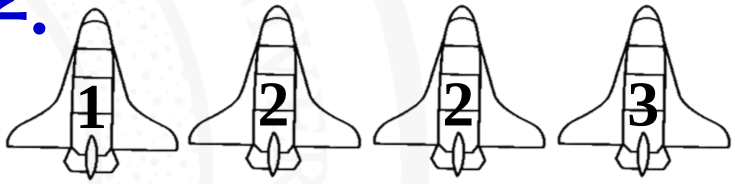
离散型随机变量关于X和Y的边缘分布函数

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 一机场有四架飞机,它们依次标有数字1,2,2,3,设每次任务派出一架飞机后**不再返回**，并且每架飞机被提取的可能性相同,以X,Y分别记第一次、第二次取得飞机上标有的数字.

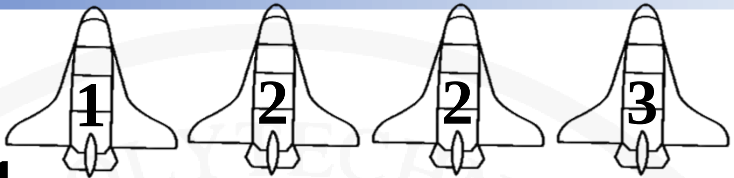


求 (1) (X,Y)的联合分布律;

(2) 关于X和关于Y的边缘分布律.

解 X的所有可能取值为1,2,3;

Y的所有可能取值为1,2,3.所以有



$$P\{X=1,Y=1\}=\frac{1}{4}\times 0=0;\ P\{X=1,Y=2\}=\frac{1}{4}\times\frac{2}{3}=\frac{1}{6};$$

$$P\{X=1,Y=3\}=\frac{1}{4}\times\frac{1}{3}=\frac{1}{12};\ P\{X=2,Y=1\}=\frac{1}{2}\times\frac{1}{3}=\frac{1}{6};$$

$$P\{X=2,Y=2\}=\frac{1}{2}\times\frac{1}{3}=\frac{1}{6};P\{X=2,Y=3\}=\frac{1}{2}\times\frac{1}{3}=\frac{1}{6};$$

$$P\{X=3,Y=1\}=\frac{1}{4}\times\frac{1}{3}=\frac{1}{12};P\{X=3,Y=2\}=\frac{1}{2}\times\frac{1}{3}=\frac{1}{6};$$

$$P\{X=3,Y=3\}=\frac{1}{4}\times 0=0.$$

由上面的计算可得(X,Y)的联合分布律为

$Y \backslash X$	1	2	3
1	0	$\frac{1}{6}$	$\frac{1}{12}$
2	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
3	$\frac{1}{12}$	$\frac{1}{6}$	0

由公式 $p_{i\bullet}=\sum_{j=1}^{\infty}p_{ij}=P\{X=x_i\}$

可得关于 X 的边缘分布律为

X	1	2	3
p	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

关于 X 的边缘分布函数为

$$F_X(x)=\begin{cases}0 & x<1\\ \frac{1}{4} & 1\le x<2\\ \frac{3}{4} & 2\le x<3\\ 1 & x\ge 3\end{cases}$$

同理可得Y的边缘分布律

Y	1	2	3
p	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

关于 X的

边缘分布函数为

$$F_Y(x)=\begin{cases}0 & y<1\\ \frac{1}{4} & 1\le y<2\\ \frac{3}{4} & 2\le y<3\\ 1 & y\ge 3\end{cases}$$

七、连续型随机变量的边缘分布

定义　若(X,Y)为二维连续型随机变量,设密度函数为$p(x,y)$,则X的边缘分布函数

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

则X的边缘密度函数为

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

同理可得 Y 的边缘密度函数

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

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

Y 的边缘密度函数.

例2 设(X,Y)在曲线$y=x^2, y=x$所围成的区域G里服从均匀分布.求联合密度函数和边缘密度函数.

解 区域G的面积$A=\int_0^1 (x-x^2)\mathrm{d}x=\frac{1}{6}$,

由题设知(X,Y)的联合密度函数为

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

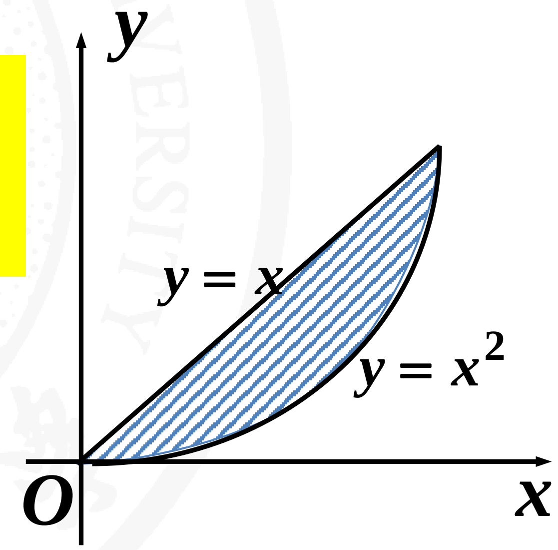


图3-3

从而

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

$$=6\int_{x^2}^{x}\mathrm{d}y=6(x-x^2),\quad 0\le x\le 1$$

即 $p_X(x)=\begin{cases}6(x-x^2),0\le x\le 1,\\ 0, \quad\quad 其他.\end{cases}$

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x=6\int_{y}^{\sqrt{y}}\mathrm{d}x \quad 0\le y\le 1$$

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

即 $p_Y(y)=\begin{cases}6(\sqrt{y}-y),0\le y\le 1,\\ 0, \quad\quad 其他.\end{cases}$

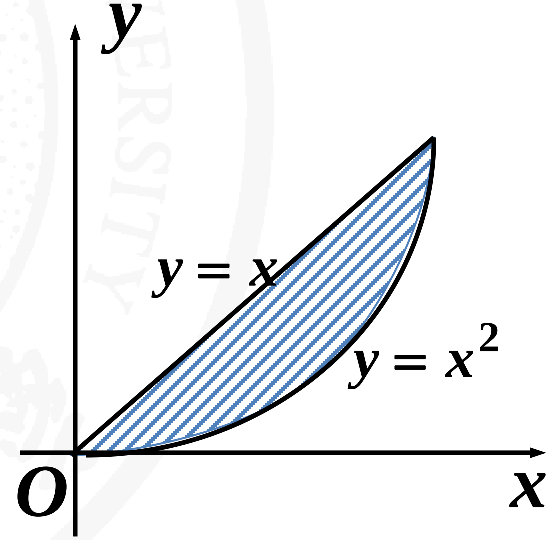


图3-3

例3 设(X,Y)的联合密度函数为

$$p(x,y)=\begin{cases} e^{-(x+y)}, & x>0, y>0, \\ 0, & \text{其他}. \end{cases}$$

求关于X和Y的边缘密度函数.

解 根据定义有

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

$$=\int_{0}^{+\infty} e^{-(x+y)}\mathrm{d}y = e^{-x} \quad (0<x<+\infty)$$

所以关于 X 的边缘密度为

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

同理可得关于 Y 的边缘密度为

$$p_Y(y)=\begin{cases} e^{-y}, & y>0, \\ 0, & y\le 0. \end{cases}$$

注 由联合分布**能推出**边缘分布，

但由边缘分布**推不出**联合分布.

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

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

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

则有

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

$\sqrt{2\pi}$

即

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

同理可得

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

注　通过本题,我们可以得到如下结论

1°　二元正态分布的边缘分布是一元正态分布.

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

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

2°　上述的两个边缘分布中的参数与二元正态分布中的常数ρ无关.

3° 如果 $(X_1,Y_1) \sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho_1)$,

$(X_2,Y_2) \sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho_2)$,

其中 $\rho_1 \neq \rho_2$

则 (X_1,Y_1) 与 (X_2,Y_2) 的联合分布函数不同,

但是 X_1 与 X_2, Y_1与Y_2的边缘分布函数相同.

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

思考题

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

$$p_X(x)=\frac{1}{2\pi}e^{-\frac{x^2}{2}}\int_{-\infty}^{+\infty}e^{-\frac{y^2}{2}}(1+\sin x\sin y)dy=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},$$

同理 $p_Y(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}.$

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.

内容小结

1. 边缘分布函数

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

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

2. 离散型随机变量的边缘分布

(1) 二维离散型随机变量的边缘分布律

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

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

(2) 二维离散型随机变量(X,Y) 的边缘分布函数

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

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

3. 连续型随机变量边缘分布

(1) 二维连续型随机变量(X,Y) 的边缘密度函数

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

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

(2) 二维连续型随机变量(X,Y)的边缘分布函数

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

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

2-2 多维随机变量

Thank You !

备用题

例3-1 如果二维随机变量(X,Y)的联合分布函数为

$$F(x,y)=\begin{cases}1-e^{-\lambda_1 x}-e^{-\lambda_2 y}\\ +e^{-\lambda_1 x-\lambda_2 y-\lambda_{12}\max\{x,y\}}, & x>0, y>0,\\ 0, & \text{其他}.\end{cases}$$

试求X和Y各自的边缘分布函数.

解 因为

$$\lim_{y\to+\infty}\{1-e^{-\lambda_1 x}-e^{-\lambda_2 y}+e^{-\lambda_1 x-\lambda_2 y-\lambda_{12}\max\{x,y\}}\}=1-e^{-\lambda_1 x},$$

$$\lim_{x\to+\infty}\{1-e^{-\lambda_1 x}-e^{-\lambda_2 y}+e^{-\lambda_1 x-\lambda_2 y-\lambda_{12}\max\{x,y\}}\}=1-e^{-\lambda_2 y},$$

所以X和Y各自的边缘分布函数为

$$F_X(x)=F(x,+\infty)=\lim_{y\to+\infty}F(x,y)=\begin{cases}1-\mathrm{e}^{-\lambda_1 x}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

$$F_Y(y)=F(+\infty,y)=\lim_{x\to+\infty}F(x,y)=\begin{cases}1-\mathrm{e}^{-\lambda_2 y}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

可见，这两个边缘分布都是指数分布，但这两个分布对应的随机变量不相互独立.

例3-2 设二维随机变量(ε,η)有密度函数

$$p(x,y)=\frac{20}{\pi^2(x^2+16)(y^2+25)}$$

求(ε,η)的联合分布函数及关于ε和η的边缘密度函数.

解

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}P(u,v)\mathrm{d}u\mathrm{d}v$$

$$=\frac{20}{\pi^2}\int_{-\infty}^{x}\int_{-\infty}^{y}\frac{\mathrm{d}u\mathrm{d}v}{(u^2+16)(v^2+25)}$$

$$=\frac{20}{\pi^2}\left(\int_{-\infty}^{x}\frac{\mathrm{d}u}{u^2+16}\right)\left(\int_{-\infty}^{y}\frac{\mathrm{d}v}{v^2+25}\right)$$

$$= \frac{1}{\pi^2}(\arctan\frac{x}{4}+\frac{\pi}{2})(\arctan\frac{y}{5}+\frac{\pi}{2})$$

$$故 F(x,y) = \frac{1}{\pi^2}(\arctan\frac{x}{4}+\frac{\pi}{2})(\arctan\frac{y}{5}+\frac{\pi}{2})$$

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

即为所求的联合分布函数.

$$F_\varepsilon(x) = \lim_{y\to+\infty} F(x,y) = \frac{1}{\pi}(\arctan\frac{x}{4}+\frac{\pi}{2})$$

$$F_\eta(y) = \lim_{x\to+\infty} F(x,y) = \frac{1}{\pi}(\arctan\frac{y}{5}+\frac{\pi}{2})$$

即为所求的边缘分布函数.

例3-4 设二维随机变量(X,Y)具如下联合概率密度，求边缘分布.

$$(1)\ p(x,y)=\begin{cases}\dfrac{2e^{-y+1}}{x^3}, & x>1, y>1,\\ 0, & \text{其他}.\end{cases}$$

$$(2)\ p(x,y)=\begin{cases}\dfrac{1}{\pi}e^{-\frac{1}{2}(x^2+y^2)}, & x>0, y\le 0, \text{或} x\le 0, y>0,\\ 0, & \text{其他}.\end{cases}$$

解 (1)对$x>1$

$$p_X(x)=\int_1^{\infty}\frac{2e^{-y+1}}{x^3}dy=\frac{2}{x^3},$$

所以 $p_X(x)=\begin{cases}\dfrac{2}{x^3}, & x>1,\\ 0, & \text{其他}.\end{cases}$

对$y>1$

$$p_Y(y)=\int_1^{\infty}\frac{2e^{-y+1}}{x^3}\mathrm{d}x=e^{-y+1},$$

所以 $p_Y(y)=\begin{cases}e^{-y+1}, & y>1,\\ 0, & \text{其他}.\end{cases}$

(2)对$x>0$

$$p_X(x)=\int_{-\infty}^{0}\frac{1}{\pi}\mathrm{e}^{-\frac{1}{2}(x^2+y^2)}\mathrm{d}y=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}};$$

对$x\le 0$

$$p_X(x)=\int_{0}^{\infty}\frac{1}{\pi}\mathrm{e}^{-\frac{1}{2}(x^2+y^2)}\mathrm{d}y=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}},$$

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

同理可求出 $p_Y(y)$.

例3-5 设二维随机变量(X,Y)在$x^2+y^2\le r^2 (r>0)$内服从均匀分布，求 X, Y 的边缘概率密度.

解 应先求出联合概率密度，为此求

$$G=\{(x,y): x^2+y^2\le r^2, r>0\}$$

的面积，显然G的面积为πr^2.

所以(X,Y)的联合概率密度为

$$p(x,y)=\begin{cases}\dfrac{1}{\pi r^2}, & x^2+y^2\le r^2,\\ 0, & 其他.\end{cases}$$

由 $p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y$，当$|x|<r$时，

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}}\frac{1}{\pi r^2}\mathrm{d}y=\frac{2\sqrt{r^2-x^2}}{\pi r^2},$$

所以 $p_X(x)=\begin{cases}\dfrac{2\sqrt{r^2-x^2}}{\pi r^2}, & |x|<r,\\ 0, & |x|\geq r.\end{cases}$

同理可求得

$$p_Y(y)=\begin{cases}\dfrac{2\sqrt{r^2-y^2}}{\pi r^2}, & |y|<r,\\ 0, & |y|\geq r.\end{cases}$$

例4-1 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) $P\{X+Y\le 1\}$

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

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

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