§ 3.2 边缘分布

1. 边缘分布函数

2. 二维离散型随机变量的边缘分布

3. 二维连续型随机变量的边缘分布

二维随机变量(X,Y)的分量X和Y是一维随机变量,它们各有其分布，称为(X,Y)分别关于X和Y的边缘分布.

本节主要讨论二维离散型随机变量(X,Y)分别关于X和Y的边缘分布律和二维连续型随机变量(X,Y)分别关于X和Y的边缘概率密度函数.

1. 边缘分布函数

设二维随机变量(X,Y)的分布函数为$F(x,y)$，关于X和Y的边缘分布函数分别记为$F_X(x)$和$F_Y(y)$.

联合分布可以确定边缘分布

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

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

注意： 由联合分布可以决定边缘分布，反过来，由边缘分布决定不了联合分布。但当分量独立时就可以决定。

例: 设(X,Y)的联合分布函数

$$F(x,y)=\begin{cases}1-e^{-0.5x}-e^{-0.5y}-e^{-0.5(x+y)}, & x\ge 0, y\ge 0\\ 0\ , & \text{其它}\end{cases}.$$

求(X,Y)关于 Y 的边缘分布函数$F_Y(y)$.

解 (X,Y)关于Y的边缘分布函数

$$F_Y(y)=F(+\infty,y)=\lim_{x\to+\infty}F(x,y)$$

$$=\begin{cases}\lim\limits_{x\to+\infty}[1-e^{-0.5x}-e^{-0.5y}-e^{-0.5(x+y)}], & y\ge 0\\ 0\ , & y<0\end{cases}=\begin{cases}1-e^{-0.5y}, & y\ge 0\\ 0\ , & y<0\end{cases}.$$

即 Y 服从参数λ =0.5 的指数分布.

2. 二维离散型随机变量的边缘分布

对于二维离散型随机变量(X,Y)，分量X,Y的分布列（律）称为二维随机变量(X,Y)的关于X和Y的边缘概率分布或分布列（律）.

设二维离散型随机变量(X,Y)的概率分布为

$$P(X=x_i, Y=y_j)=P_{ij}, \quad i,j=1,2,\ldots,$$

则 $P(X=x_i)= P((X=x_i)\bigcap[\bigcup_j (Y=y_j)])$

$$=\sum_j P((X=x_i)\bigcap(Y=y_j))$$

$$=\sum_j P(X=x_i, Y=y_j) = \sum_j p_{ij} \qquad (i=1,2,\ldots)$$

同理:

$$P(Y=y_j)=\sum_i p_{ij} \qquad (j=1,2,\ldots)$$

一般地，记:

$P(X=x_i) \Rightarrow \boxed{\boldsymbol{P_{i\cdot}}}$ $\qquad$ $P(Y=y_j) \Rightarrow \boxed{\boldsymbol{P_{\cdot j}}}$

其分布表如下:

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

例(前例)： 令随机变量 X 表示在 1,2,3,4 中等可能地取一个值，随机变量 Y 表示在 1~X 中等可能地取一个整数值. 求（X，Y）分别关于 X 和 Y 的边缘分布律.

解 $P(X=i,Y=j)=P(Y=j|X=i)P(X=i)=1/4i,\ (i\geq j)$

于是(X,Y)的分布律及关于X和Y的边缘分布律为

X \ Y	1	2	3	4	$P(X=i)$
1	1/4	0	0	0	1/4
2	1/8	1/8	0	0	1/4
3	1/12	1/12	1/12	0	1/4
4	1/16	1/16	1/16	1/16	1/4
$P(Y=j)$	25/48	13/48	7/48	3/48	1

例(前例) 袋中有 2 个黑球 3 个白球，从袋中随机取两次，每次取一个球，在不放回的情况下．令

$$X=\begin{cases}1 & \text{第一次取到黑球}\\0 & \text{第一次取到白球}\end{cases},\quad Y=\begin{cases}1 & \text{第二次取到黑球}\\0 & \text{第二次取到白球}\end{cases},$$

求(X,Y)的联合分布律及边缘概率分布

解 在不放回抽样下，列表如下：

X \ Y	0	1	$P_{i\cdot}$
0	6/20	6/20	3/5
1	6/20	2/20	2/5
$P_{\cdot j}$	3/5	2/5	

3. 二维连续型随机变量的边缘分布

对于二维连续型随机变量(X,Y), 设其概率密度函数为$f(x,y)$，分布函数为$F(x,y)$，则有

$$F_X(x) = F(x,+\infty) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f(u,y)dudy$$

$$= \int_{-\infty}^{x}[\underline{\int_{-\infty}^{+\infty} f(u,y)dy}]du$$

$$F_Y(y) = F(+\infty, y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{y} f(x,v)dxdv$$

$$= \int_{-\infty}^{y}[\underline{\int_{-\infty}^{+\infty} f(x,v)dx}]dv$$

记

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

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx$$

分别称$f_X(x)$, $f_Y(y)$为二维连续型随机变量(X,Y)关于X和Y的边缘概率密度函数，简称密度函数。

边缘密度函数完全由联合密度函数所决定.

例 设随机变量(X,Y)服从区域D上的均匀分布，其中

D={(x,y) , $x^2+y^2\le1$}，求X ,Y的边缘密度函数$f_X(x)$和$f_Y(y)$.

解 (1)由题意得:

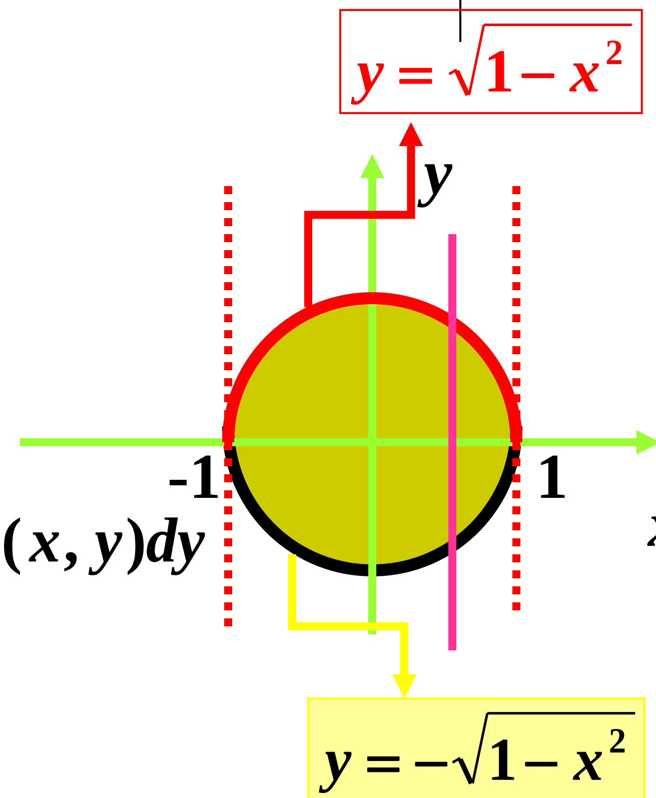
$$f(x,y)=\begin{cases}\dfrac{1}{\pi} & ,\ x^2+y^2\le 1\\ 0 & ,\ 其它\end{cases}$$

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

当$|x|>1$时，$f(x,y)=0$，所以，$f_X(x)=0$

当$|x|\le1$时，
$$\boldsymbol{f_X(x)=[\int_{-\infty}^{-\sqrt{1-x^2}}+\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}+\int_{\sqrt{1-x^2}}^{+\infty}]f(x,y)dy}$$

$$\boldsymbol{=\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}dy=\frac{2}{\pi}\sqrt{1-x^2}}$$



所以，
$$f_X(x)=\begin{cases}\dfrac{2}{\pi}\sqrt{1-x^2} & ,|x|\le 1\\ 0 & ,\qquad |x|>1\end{cases}$$

同理,

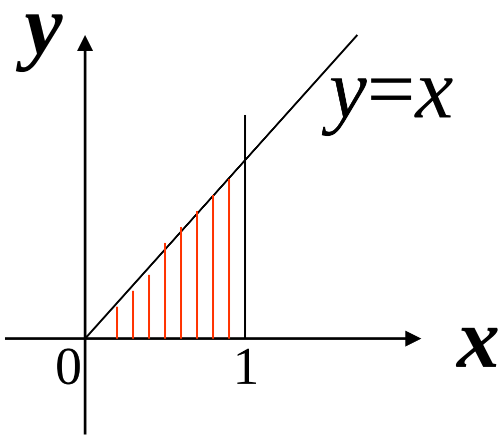
$$f_Y(y)=\begin{cases}\dfrac{2}{\pi}\sqrt{1-y^2}, & |y|\le 1\\ 0, & |y|>1\end{cases}$$

注意: 均匀分布的边缘密度函数不再是一维均匀分布

例 设(X,Y)的概率密度函数是

$$f(x,y)=\begin{cases} cy(2-x), & 0\le x\le 1,\ 0\le y\le x \\ 0\ , & \text{其它} \end{cases}$$

求 (1) c的值； （2）两个边缘概率密度函数.



解 (1) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy$

$$=\int_0^1[\int_0^x cy(2-x)dy]dx$$

$$=c\int_0^1[x^2(2-x)/2]dx=\frac{5c}{24}=1$$

所以，c = 24/5

$$(2)\quad f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_0^x \frac{24}{5}y(2-x)dy$$

注意积分限

$$=\frac{12}{5}\boldsymbol{x}^2(2-\boldsymbol{x}),\qquad 0\le x\le 1$$

注意取值范围

同理 $$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_y^1 \frac{24}{5}y(2-x)dx$$

$$=\frac{24}{5}y\left(\frac{3}{2}-2y+\frac{y^2}{2}\right),\qquad 0\le y\le 1$$

即

$$f_X(x)=\begin{cases}\dfrac{12}{5}x^2(2-x), & 0\le x\le 1\\ 0\quad, & \text{其它}\end{cases}$$

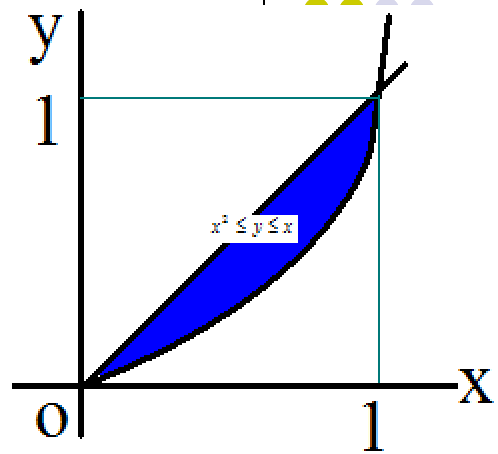
$$f_Y(y)=\begin{cases}\dfrac{24}{5}y(\dfrac{3}{2}-2y+\dfrac{y^2}{2}), & 0\le y\le 1\\ 0\quad, & \text{其它}\end{cases}$$

注意： 在求二维连续型随机变量的边缘概率密度时，往往要对联合概率密度在一个变量取值范围上进行积分．当联合密度函数是分段函数的时候，在计算积分时应特别注意积分限．

例 设随机变量X和Y具有联合概率密度

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

求边缘概率密度 $f_X(x)$和 $f_Y(y)$.



解

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\begin{cases}\int_{x^2}^{x}6dy=6(x-x^2), & 0\le x\le 1\\ 0\ , & 其它\end{cases}$$

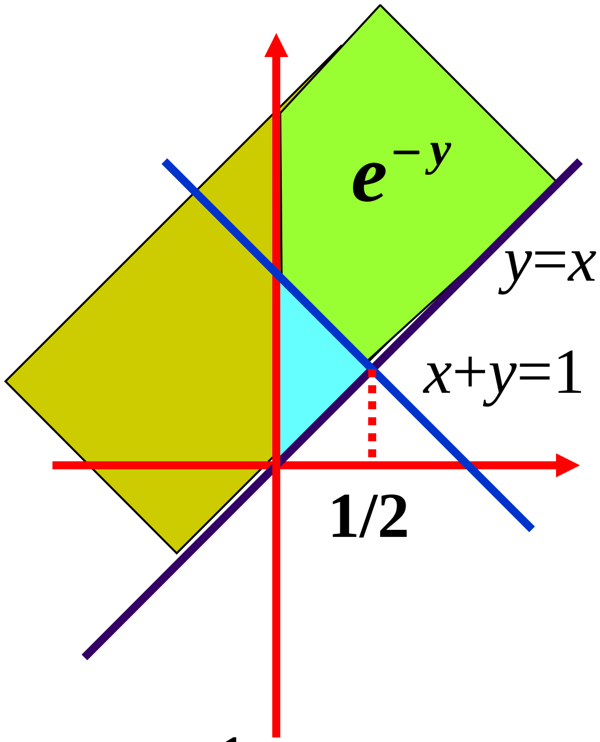
$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx=\begin{cases}\int_{y}^{\sqrt{y}}6dx=6(\sqrt{y}-y), & 0\le y\le 1\\ 0\ , & 其它\end{cases}$$

例 设二维随机变量（X，Y）的概率密度为

$$f(x,y)=\begin{cases} e^{-y}, & 0<x<y \\ 0, & 其他 \end{cases}$$

(1) 求随机变量X的边缘密度函数；

(2) 求概率P(X+Y≤1).



解 (1)$x\le 0$时, $f_X(x)=0$;

$x>0$时，$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy$

$$=\int_{x}^{+\infty} e^{-y}dy = e^{-x}$$

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

(2) P(X+Y≤1)= $\int_0^{1/2} dx \int_x^{1-x} e^{-y} dy = 1+e^{-1}-2e^{-\frac{1}{2}}$

例 求二维正态随机变量的边缘密度函数.

解 已知

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

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

为了计算方便，设

$$\frac{x-\mu_1}{\sigma_1}=u \qquad \frac{y-\mu_2}{\sigma_2}=v$$

则(X,Y)关于X的边缘密度函数为

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

$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\exp[-\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)]\,\sigma_2 dv$$

$$=\frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\exp\{-\frac{1}{2(1-\rho^2)}[(u^2-\rho^2u^2)+(\rho^2u^2-2\rho uv+v^2)]\}dv$$

$$=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{u^2}{2}}\underline{\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}e^{-\frac{(v-\rho u)^2}{2(1-\rho^2)}}dv}$$

积分中的被积函数恰好是服从正态分布 $N(\rho u,(\sqrt{1-\rho^2})^2)$ 的随机变量的密度函数

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

同理，(X,Y)关于Y的边缘密度函数为

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}\cdot(\frac{y-\mu_2}{\sigma_2})^2},-\infty<y<+\infty,$$

由此可见：二维正态分布的两个边缘分布都是一维正态分布，而且这两个边缘分布与其中的参数ρ无关。即 $X\sim N(\mu_1,\sigma_1^2)$ $Y\sim N(\mu_2,\sigma_2^2)$

这表明，仅仅由X和Y的边缘分布，一般不能完全确定二维随机变量(X,Y)的联合分布