§ 3.2 边缘分布 Marginal Distribution

- 边缘分布函数
- 边缘分布律
- 边缘概率密度

边缘分布marginal distribution

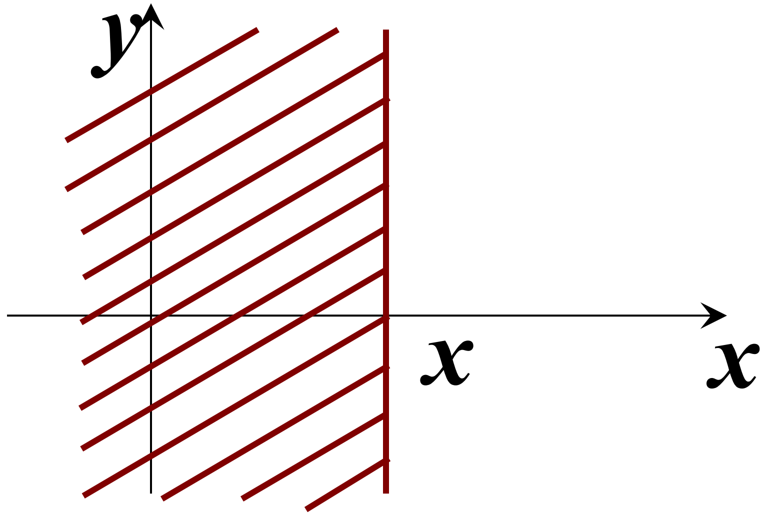
如果（X，Y）是一个二维随机变量，则它的分量X（或者Y）是一维随机变量，因此，分量X（或者Y）也有分布．我们称X（或者Y）的分布为X（或者Y）关于二维随机变量（X，Y）的边缘分布．

边缘分布 也称为 边沿分布 或 边际分布

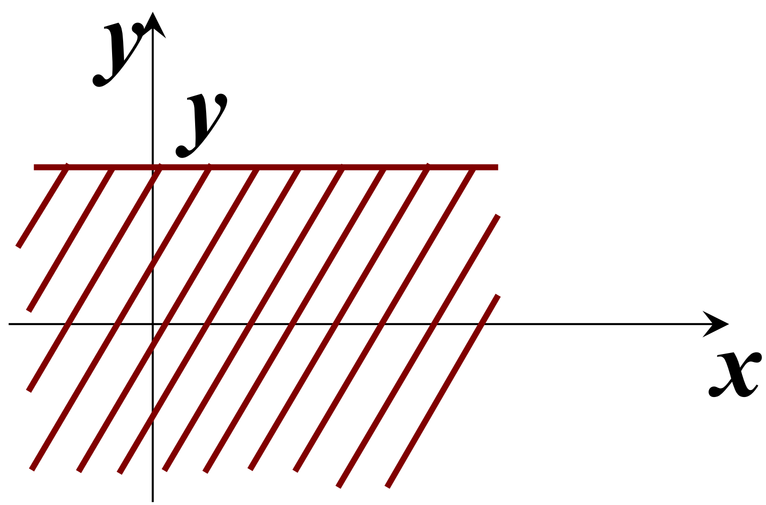
1. 二维随机变量的边缘分布函数 重要

由联合分布函数 $\Rightarrow$ 边缘分布函数, 逆不真.

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



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



例 设随机变量(X,Y)的联合分布函数为

$$F(x,y) = A\left(B + \arctan\frac{x}{2}\right)\left(C + \arctan\frac{y}{2}\right)$$

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

其中A，B，C为常数.

(1) 确定A，B，C；

(2) 求X和Y的边缘分布函数；

(3) 求$P(X > 2)$.

解 **(1)**

$$F(+\infty,+\infty)=A\left(B+\frac{\pi}{2}\right)\left(C+\frac{\pi}{2}\right)=1$$

$$F(-\infty,+\infty)=A\left(B-\frac{\pi}{2}\right)\left(C+\frac{\pi}{2}\right)=0$$

$$F(+\infty,-\infty)=A\left(B+\frac{\pi}{2}\right)\left(C-\frac{\pi}{2}\right)=0$$

$$\Rightarrow\quad B=\frac{\pi}{2},C=\frac{\pi}{2},A=\frac{1}{\pi^2}$$

(2)

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

$$=\frac{1}{2}+\frac{1}{\pi}\arctan\frac{x}{2},\qquad -\infty<x<+\infty.$$

arctanx在(-∞，+-∞) 上的值域为(-π/2,π/2)

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

$$= \frac{1}{2} + \frac{1}{\pi}\arctan\frac{y}{2}, \quad -\infty < y < +\infty.$$

(3) $P(X > 2) = 1 - P(X \le 2) = 1 - F_X(2)$

$$= 1 - \left(\frac{1}{2} + \frac{1}{\pi}\arctan\frac{2}{2}\right)$$

$$= 1/4.$$

arctan 1 = π /4

2. 二维离散型随机变量的边缘分布 重要

$$P(X = x_i) = \sum_{j=1}^{+\infty} p_{ij} \overset{\text{记作}}{=} p_{i\cdot}, \quad i = 1,2,\cdots$$

$$P(Y = y_j) = \sum_{i=1}^{+\infty} p_{ij} \overset{\text{记作}}{=} p_{\cdot j}, \quad j = 1,2,\cdots$$

由联合分布律可确定边缘分布律

联合分布律 及边缘分布律

$Y \backslash X$	x_1	$\cdots$	x_i	$\cdots$	$p_{\bullet j}$
y_1	p_{11}	$\cdots$	p_{i1}	$\cdots$	$p_{\bullet 1}$
$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$
y_j	p_{1j}	$\cdots$	p_{ij}	$\cdots$	$p_{\bullet j}$
$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$
$p_{i\bullet}$	$p_{1\bullet}$	$\cdots$	$p_{i\bullet}$	$\cdots$	1

例(P61) 设随机变量 X 在 1,2,3三个数中等可能地取值，另一个随机变量 Y 在1~X 中等可能地取一整数值，试求 X, Y 的边缘分布律。

(X,Y)的联合分布律为

$X \backslash Y$	1	2	3
1	$\frac{1}{3}$	0	0
2	$\frac{1}{6}$	$\frac{1}{6}$	0
3	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

(X,Y)的联合与边缘分布律

$X \backslash Y$	1	2	3	$p_{i\cdot}$
1	$\frac{1}{3}$	0	0	$\frac{1}{3}$
2	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{3}$
3	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{3}$
$p_{\cdot j}$	$\frac{11}{18}$	$\frac{5}{18}$	$\frac{2}{18}$	1

例 箱子里装有4只白球和2只黑球，在其中随机地取两次，每次取一只。考虑两种试验：

(1) 有放回抽样，(2) 不放回抽样。

定义随机变量 *X*，*Y* 如下，写出*X*和*Y*的联合分布律和边缘分布律。

$$X = \begin{cases} 0, \text{若第一次取出的是黑球}, \\ 1, \text{若第一次取出的是白球}. \end{cases}$$

$$Y = \begin{cases} 0, \text{若第二次取出的是黑球}, \\ 1, \text{若第二次取出的是白球}. \end{cases}$$

(1) 有放回抽样

P(X=0)=2/6 P(X=1)=4/6
P(Y=0)=2/6 P(Y=1)=4/6

$X \backslash Y$	0	1	$p_{i\cdot}$
0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$
1	$\frac{2}{9}$	$\frac{4}{9}$	$\frac{2}{3}$
$p_{\cdot j}$	$\frac{1}{3}$	$\frac{2}{3}$	1

P(X=0,Y=0)=P(X=0)P(Y=0)

(2) 不放回抽样

$X \backslash Y$	0	1	$p_{i\cdot}$
0	$\frac{1}{15}$	$\frac{4}{15}$	$\frac{1}{3}$
1	$\frac{4}{15}$	$\frac{6}{15}$	$\frac{2}{3}$
$p_{\cdot j}$	$\frac{1}{3}$	$\frac{2}{3}$	1

P(X=0,Y=0)
=P(X=0)P(Y=0|X=0)
=2/6 * 1/5 = 1/15

3. 二维连续型随机变量的边缘分布 重要

$$F_X(x) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f(u,v)\,dv\,du$$

$$F_Y(y) = \int_{-\infty}^{y}\int_{-\infty}^{+\infty} f(u,v)\,du\,dv$$

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

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

已知联合密度可以求得边缘密度

例 设随机变量(X,Y)服从区域D上的均匀分布.

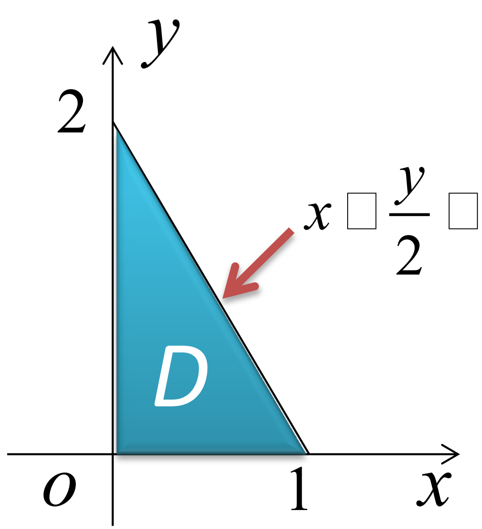
其中$D=\{(x,y)\,|\,x\ge 0, y\ge 0, x+\dfrac{y}{2}\le 1\}$,

试求随机变量(X,Y)的边缘密度函数.

解: 区域D的面积为 $A=1$

(X,Y)的联合密度函数为

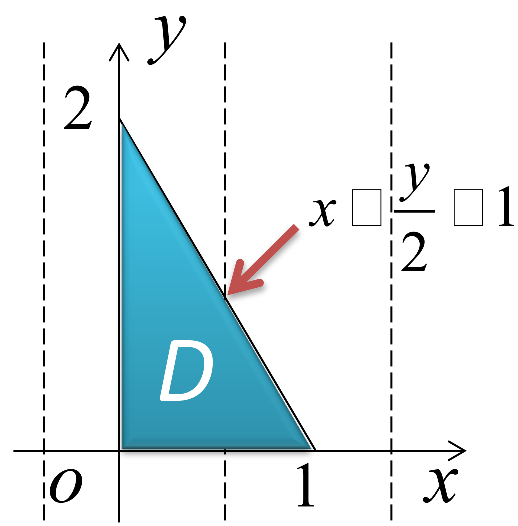
$$f(x,y)=\begin{cases}1, & (x,y)\in D\\ 0, & (x,y)\notin D\end{cases}$$



当 $0 \le x \le 1$ 时，

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_{-\infty}^{0}0dy+\int_{0}^{2(1-x)}1dy+\int_{2(1-x)}^{+\infty}0dy$$

$$=2(1-x)$$



当 $x<0$ 或 $x>1$ 时，$f_X(x)=0$

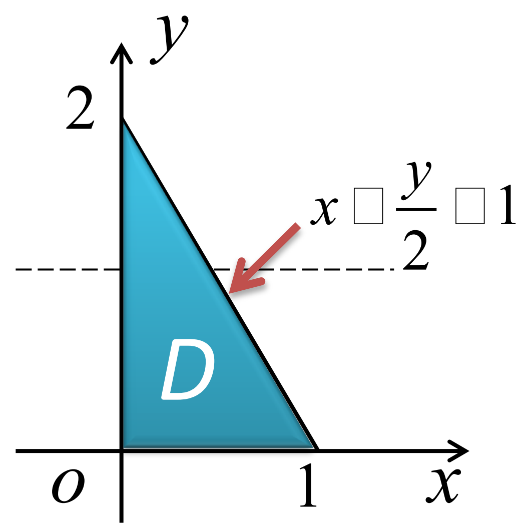
随机变量 X 的边缘密度函数为

$$f_X(x)=\begin{cases}2(1-x) & 0\le x\le 1\\ 0 & \text{其它}\end{cases}$$

同理，随机变量Y的边缘密度函数为

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

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



例 设二维随机变量 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

试求 X 及 Y 的边缘密度函数.

解：(X,Y) 的联合密度函数为

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

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

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

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

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

这表明，$X \sim N(\mu_1, \sigma_1^2)$

$$Y \sim N(\mu_2, \sigma_2^2)$$

通过本题，我们有如下几条结论：

结论（一）

二维正态分布的边缘分布是一维正态分布

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

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

结论（二）

上述的两个边缘分布中的参数与二维正态分布中的常数ρ无关

总结

1. 边缘分布
   a. 分布函数边缘化
   b. 分布律边缘化
   c. 密度函数边缘化
      画图确定密度非0区域
      确定定义域
      负无穷到正无穷边缘化掉某一分量