第二节 边缘分布

一、边缘分布函数

定义 1 设(X,Y)为二维随机变量，分别称X和Y的分布函数为(X,Y)关于X和关于Y的边缘分布函数，记为$F_X(x)$和$F_Y(y)$.

定理 1 设二维随机变量(X,Y)的分布函数为$F(x,y)$，

则有
$$F_X(x)=F(x,+\infty)=\lim_{y\to+\infty}F(x,y),\ -\infty<x<+\infty;$$

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

例 1 设二维随机变量(X,Y)的分布函数为

$$F(x,y)=\frac{1}{\pi^2}(\frac{\pi}{2}+\arctan x)(\frac{\pi}{2}+\arctan y)，\quad -\infty<x,y<+\infty，$$

试求出$F_X(x)$和$F_Y(y)$，并问是否有$F(x,y)=F_X(x)F_Y(y)$？

解 $F_X(x)=F(x,+\infty)=\frac{1}{\pi}(\frac{\pi}{2}+\arctan x)，\quad -\infty<x<+\infty；$

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

并易得，$F(x,y)=F_X(x)F_Y(y)，\quad -\infty<x,y<+\infty．$

二、二维离散型随机变量的边缘分布律

定义 2 设(X,Y)为二维离散型随机变量，分别称X和Y的分布律为(X,Y)关于X和关于Y的边缘分布律.

定理 2 设二维离散型随机变量(X,Y)的分布律为

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

则(X,Y)关于X和关于Y的边缘分布律分别为

$$P\{X=x_i\}=\sum_j p_{ij} \overset{\text{记为}}{=} p_{i\bullet}, \quad i=1,2,\cdots;$$

$$P\{Y=y_j\}=\sum_i p_{ij} \overset{\text{记为}}{=} p_{\bullet j}, \quad j=1,2,\cdots.$$

(X,Y) 关于 X 和关于 Y 的边缘分布律分别为

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots \\ p_{1\bullet} & p_{2\bullet} & \cdots & p_{i\bullet} & \cdots \end{pmatrix}, \quad Y \sim \begin{pmatrix} y_1 & y_2 & \cdots & y_j & \cdots \\ p_{\bullet 1} & p_{\bullet 2} & \cdots & p_{\bullet j} & \cdots \end{pmatrix}.$$

将 X 和 Y 的边缘分布律添加到 (X,Y) 分布律的列表得：

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

X 的边缘分布律可对表中的 p_{ij} 进行横向求和即得；关于 Y 的边缘分布律可对表中的 p_{ij} 进行纵向求和.

例 2 设同一品种的五个产品中，有二个次品，每次从中取一个检验，连续二次．设 X 表示第一次取到的次品个数；Y 表示第二次取到的次品个数．试分别就 (1) 不放回；(2)有放回两种情况，求出 (X,Y) 关于 X 和关于 Y 的边缘分布律．

解 (1)不放回；(2)有放回两种情况时，(X,Y) 关于 X 和关于 Y 的边缘分布律如下

$X \backslash Y$	0	1	$p_{i\bullet}$
0	$\frac{3}{10}$	$\frac{3}{10}$	$\frac{3}{5}$
1	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{2}{5}$
$p_{\bullet j}$	$\frac{3}{5}$	$\frac{2}{5}$	1

$X \backslash Y$	0	1	$p_{i\bullet}$
0	$\frac{9}{25}$	$\frac{6}{25}$	$\frac{3}{5}$
1	$\frac{6}{25}$	$\frac{4}{25}$	$\frac{2}{5}$
$p_{\bullet j}$	$\frac{3}{5}$	$\frac{2}{5}$	1

例 3 已知随机变量 X_1 和 X_2 的概率分布为

$$X_1 \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}, \quad X_2 \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

且 $P\{X_1X_2=0\}=1$．求 X_1 和 X_2 的联合分布．

解 由 $P\{X_1X_2=0\}=1$ 知，$P\{X_1X_2\neq 0\}=0$，所以

$$P\{X_1=-1,X_2=1\}=P\{X_1=1,X_2=1\}=0,$$

根据 X_1 的分布律得 $P\{X_1=-1,X_2=0\}=P\{X_1=1,X_2=0\}=\frac{1}{4}$.

根据 X_2 的分布律得 $P\{X_1=0,X_2=0\}=0, P\{X_1=0,X_2=1\}=\frac{1}{4}$.

（续解） 综上，X_1 和 X_2 的联合分布律为

$X_2 \backslash X_1$	-1	0	1	$P\{X_2 = y_j\}$
0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
1	0	$\frac{1}{2}$	0	$\frac{1}{2}$
$P\{X_1 = x_i\}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

三、二维连续型随机变量的边缘概率密度

定义 3 设(X,Y)为二维连续型随机变量，分别称 X 和 Y 的密度函数为(X,Y)关于 X 和关于 Y 的**边缘密度函数**，记为 $f_X(x)$ 和 $f_Y(y)$.

定理 3 设二维连续型随机变量(X,Y)的密度函数为 $f(x,y)$，则 X 和 Y 也均为连续型随机变量，且

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

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

注 1: $f_X(x)$ 可通过在给定点 x 处， $f(x,y)$ 的对 y 从 $-\infty$ 到 $+\infty$ （纵向）积分求得， $f_Y(y)$ 可通过在给定点 y 处， $f(x,y)$ 的对 x 从 $-\infty$ 到 $+\infty$ （横向）积分求得.

注 2: 由于 $f(x,y)$ 通常以分块函数的形式给出，因此经常需要对 x 或 y 进行分段讨论，以计算 $f_X(x)$ 或 $f_Y(y)$.

例 4 设二维随机变量(X,Y)的密度函数为 $f(x,y)=\begin{cases} e^{-x}, 0<y<x, \\ 0 \quad 其它. \end{cases}$

试分别计算 $f_X(x)$ 和 $f_Y(y)$.

解 当$x\le 0$时，对任意的$y\in(-\infty,+\infty)$， $f(x,y)=0$，进而

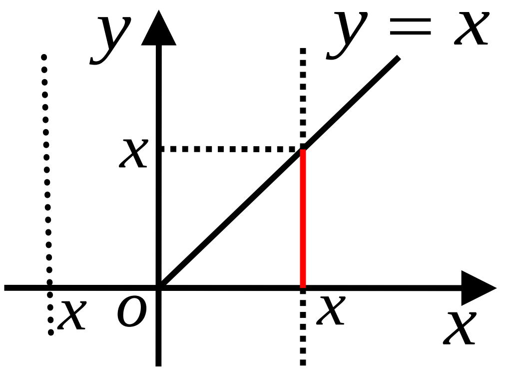
$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=0.$$

当$x>0$时，对任意的$y\in(-\infty,0]\bigcup[x,+\infty)$， $f(x,y)=0$，故

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_0^x e^{-x}dy=xe^{-x}。$$

见右图，所以

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



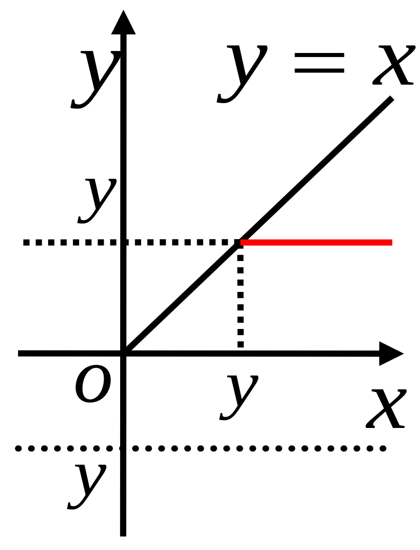
（续解） $f(x,y)=\begin{cases} e^{-x}, 0<y<x, \\ 0 \quad 其它. \end{cases}$

当 $y \le 0$ 时，对任意的 $x \in (-\infty,+\infty)$， $f(x,y)=0$， 得

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx = 0.$$

当 $y > 0$ 时，对任意的 $x \in (-\infty, y]$， $f(x,y)=0$， 得

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



见右图，所以

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

定理 4 设二维随机变量 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则

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

定理 4 表明二维正态分布的边缘分布为一维正态分布，且其边缘分布只分别依赖于 μ_1,σ_1 和 μ_2,σ_2，而不依赖于参数 ρ $(-1<\rho<1)$．因此对于给定 $\mu_1,\mu_2,\sigma_1,\sigma_2$，不同的 ρ 对应了不同的二维正态分 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，但其却有相同的边缘分布 $N(\mu_1,\sigma_1^2)$ 和 $N(\mu_2,\sigma_2^2)$，所以**边缘密度函数不能惟一地确定联合密度函数.**

练习：

1. 设二维随机变量(X,Y)的密度函数为

$$f(x,y)=\begin{cases} x^2+\frac{1}{3}xy, 0\le x\le 1, 0\le y\le 2 \\ 0, \quad 其它. \end{cases}$$

试分别计算 $f_X(x)$ 和 $f_Y(y)$.

答案：

$$f_X(x)=\begin{cases} 2x^2+\frac{2}{3}x, 0\le x\le 1, \\ 0, \quad 其它. \end{cases} \qquad f_Y(y)=\begin{cases} \frac{1}{3}+\frac{y}{6}, 0\le y\le 2, \\ 0, \quad 其它. \end{cases}$$