第二节　边缘分布

- 边缘分布函数
- 离散型随机变量的边缘分布律
- 连续型随机变量的边缘概率密度
- 课堂练习
- 小结 布置作业

二维联合分布全面地反映了二维随机变量(*X*,Y)的取值及其概率规律. 而单个随机变量*X*,Y也具有自己的概率分布. 那么要问:二者之间有什么关系呢?

这一节里,我们就来探求这个问题 .

一、边缘分布函数

二维随机变量 (X,Y)作为一个整体,具有分布函数 $F(x,y)$,而 X 和 Y 都是随机变量,也有各自的分布函数,分别记为 $F_X(x), F_Y(y)$, 依次称为二维随机变量 (X,Y) 关于 X 和 Y的边缘分布函数.

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

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

二、离散型随机变量的边缘分布律

一般地，对离散型 *r.v* (*X,Y*)，

*X*和*Y* 的联合分布律为

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

则 **(X,Y) 关于X 的边缘分布律**为

$$P\{X=x_i\}=\sum_{j=1}^{\infty} P\{X=x_i, Y=y_j\}=\sum_{j=1}^{\infty} p_{ij} \triangleq p_{i\cdot}$$

$$(i=1,2,\cdots)$$

$$\left(\{X=x_i\}=\bigcup_{j=1}^{\infty}\{X=x_i, Y=y_j\}\right)$$

(X,Y) 关于 Y 的边缘分布律为

$$P\{Y=y_j\}=\sum_{i=1}^{\infty} P\{X=x_i, Y=y_j\}=\sum_{i=1}^{\infty} p_{ij} \triangleq p_{\cdot j}$$

$$(j=1,2,\cdots)$$

例1 把一枚均匀硬币抛掷三次，设X为三次抛掷中正面出现的次数，而Y为正面出现次数与反面出现次数之差的绝对值，求(X,Y)的分布律.

解 (X,Y)可取值$(0,3)$，$(1,1)$，$(2,1)$，$(3,3)$

$$P\{X=0,Y=3\}=(1/2)^3=1/8$$

$$P\{X=1,Y=1\}=\binom{3}{1}\cdot\frac{1}{2}\cdot\left(\frac{1}{2}\right)^2=3/8$$

$$P\{X=2,Y=1\}=\binom{3}{2}\cdot\left(\frac{1}{2}\right)^2\cdot\frac{1}{2}=3/8$$

$$P\{X=3,Y=0\}=(1/2)^3=1/8.$$

X \ Y	1	3
0	0	1/8
1	3/8	0
2	3/8	0
3	0	1/8

X \ Y	1	3
0	0	1/8
1	3/8	0
2	3/8	0
3	0	1/8

$$P\{X=0\}=P\{X=0,\ Y=1\}+P\{X=0,\ Y=3\}=1/8,$$

$$P\{X=1\}=P\{X=1,\ Y=1\}+P\{X=1,\ Y=3\}=3/8,$$

$$P\{X=2\}=P\{X=2,\ Y=1\}+P\{X=2,\ Y=3\}=3/8,$$

$$P\{X=3\}=P\{X=3,\ Y=1\}+P\{X=3,\ Y=3\}=1/8.$$

$$P\{Y=1\}=\sum_{k=0}^{3}P\{X=k,Y=1\}=3/8+3/8=6/8,$$

$$P\{Y=3\}=\sum_{k=0}^{3}P\{X=k,Y=3\}=1/8+1/8=2/8.$$

$X \backslash Y$	1	3	$P\{X=x_i\}$
0	0	1/8	1/8
1	3/8	0	3/8
2	3/8	0	3/8
3	0	1/8	1/8
$P\{Y=y_j\}$	6/8	2/8	

我们常将边缘分布律写在联合分布律表格的边缘上，由此得出边缘分布这个名词.

联合分布与边缘分布的关系

$X \backslash Y$	1	3	$P\{X=x_i\}$
0	0	1/8	1/8
1	3/8	0	3/8
2	3/8	0	3/8
3	0	1/8	1/8
$P\{Y=y_j\}$	6/8	2/8	

由联合分布可以确定边缘分布;

但由边缘分布一般不能确定联合分布.

三、连续型随机变量的边缘概率密度

对连续型 $r.v\ (X,Y)$，

X 和 Y 的联合概率密度为 $f(x,y)$

则 (X,Y) 关于 X 的边缘概率密度为

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

事实上， $F_X(x)=F(x,+\infty)=\int_{-\infty}^{x} dx \int_{-\infty}^{+\infty} f(x,y)dy$

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

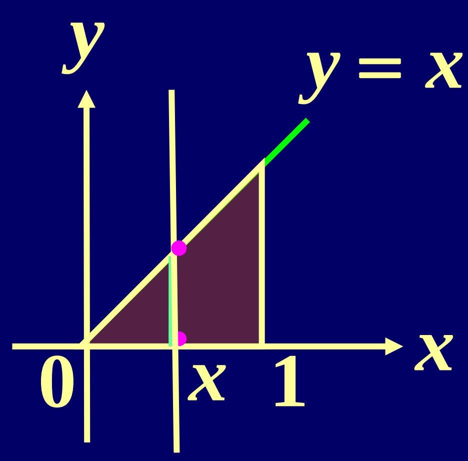
(X,Y)关于Y的边缘概率密度为

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

例2 设(X,Y)的概率密度是

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

求 (1) c的值； （2）两个边缘密度。



解 (1) $$1=\iint_{R^2} f(x,y)\,dxdy$$

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

$$=\frac{c}{2}\int_0^1\left(2x^2-x^3\right)dx$$

$$=5c/24,$$

故 $c=24/5.$

例2 设 (X,Y) 的概率密度是

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

求 (1) c 的值; (2) 两个边缘密度

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

暂时固定

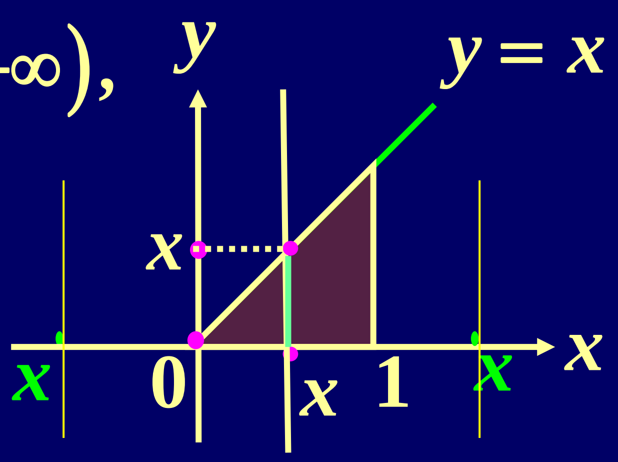
当 $x>1$ 或 $x<0$ 时， $\forall y\in(-\infty,+\infty)$,

都有 $f(x,y)=0$, 故 $f_X(x)=0$.

当 $0\le x\le 1$ 时,

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

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



当 $0 \le x \le 1$时,

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

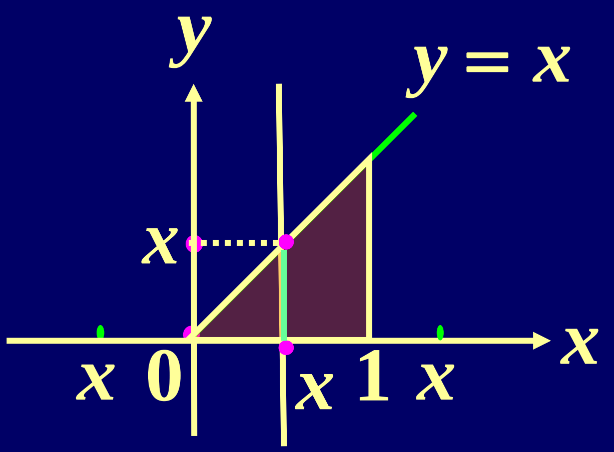
$$+\int_{0}^{x} f(x,y)dy + \int_{x}^{+\infty} f(x,y)dy.$$

$$= \int_0^x \frac{24}{5} y(2-x)dy$$

$$= \frac{12}{5} x^2 (2-x),$$

综上,

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



注意取值范围

例 2 设(X,Y)的概率密度是

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

暂时固定

求 (1) c的值; (2) 两个边缘密度.

解 (2) $f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx$

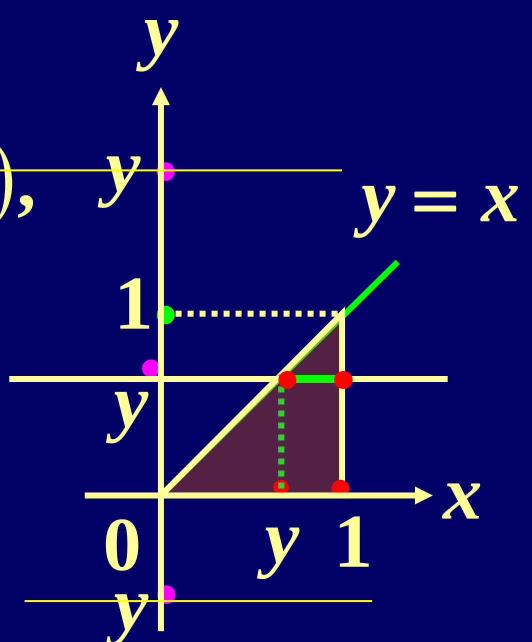
当$y>1$或$y<0$时, 对$\forall x\in(-\infty,+\infty)$,

都有 $f(x,y)=0$, 故 $f_Y(y)=0$.

当$0\le y\le 1$时,

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

$$+\int_{y}^{1} f(x,y)dx+\int_{1}^{+\infty} f(x,y)dx.$$



$$= \int_y^1 \frac{24}{5} y(2-x)dx$$

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

综上,

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

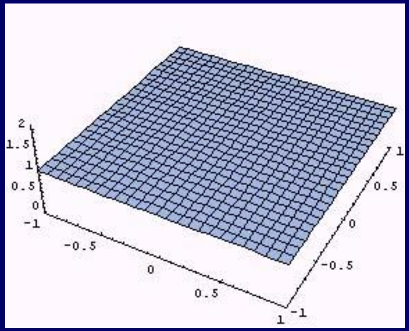
注意取值范围

在求连续型 *r.v* 的边缘密度时，往往要求联合密度在某区域上的积分. 当联合密度函数是分片表示的时候，在计算积分时应特别注意积分限 .

下面我们介绍两个常见的二维分布.

设G是平面上的有界区域，其面积为A.若二维随机变量（X,Y）具有概率密度

$$f(x,y)=\begin{cases}\dfrac{1}{A}, & (x,y)\in G\\ 0, & \text{其它}\end{cases}$$



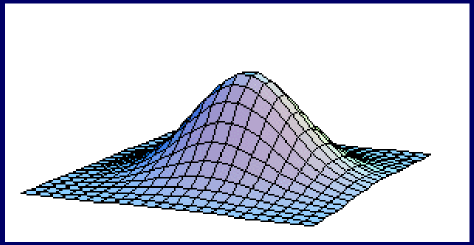
则称（X,Y）在G上服从均匀分布.

例

向平面上有界区域G上任投一质点，若质点落在G内任一小区域B的概率与小区域的面积成正比，而与B的形状及位置无关. 则质点的坐标 (X,Y)在G上服从均匀分布.

若二维随机变量（X,Y）具有概率密度

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



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

其中 $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 均为常数，且 $\sigma_1>0,\sigma_2>0$, $|\rho|<1$. 则称（X,Y）服从参数为 $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 的二维正态分布.记作（X,Y）～$N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$.

例 3 试求二维正态随机变量的边缘概率密度.

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

因为 $$\frac{(y-\mu_2)^2}{\sigma_2^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}$$

$$=\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right)^2-\rho^2\frac{(x-\mu_1)^2}{\sigma_1^1}$$

所以

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

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

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

$$f_X(x)=\frac{1}{2\pi\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt=\frac{1}{2\pi\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\cdot\sqrt{2\pi}$$

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

同理

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

可见

二维正态分布的两个边缘分布都是一维正态分布，并且不依赖于参数 ρ .

也就是说,对于给定的 $\mu_1,\mu_2,\sigma_1,\sigma_2$, 不同的 ρ 对应不同的二维正态分布,但它们的边缘分布却都是一样的.

此例表明 由边缘分布一般不能确定联合分布.

四、课堂练习

设(X,Y)的概率密度是

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

求(X,Y)关于 X 和 Y 的边缘概率密度.

暂时固定

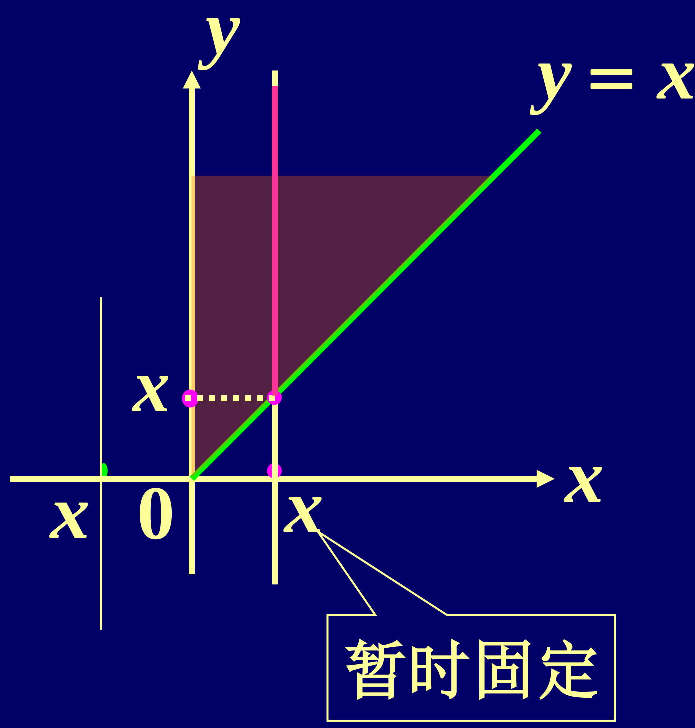
解 $f_X(x) = \int_{-\infty}^{+\infty} f(x, y)\,dy$

当 $x \le 0$ 时，$f_X(x) = \int_{-\infty}^{+\infty} 0\,dy = 0$

当 $x > 0$ 时，

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

$$= e^{-y}\Big|_{+\infty}^{x} = e^{-x}$$



故

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

暂时固定

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

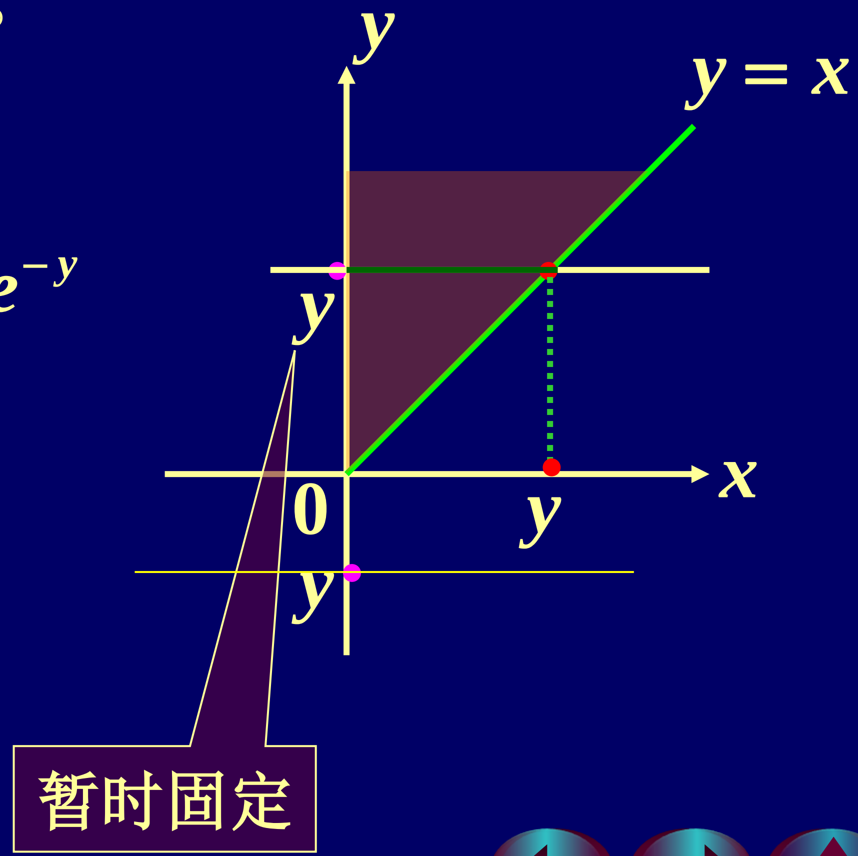
当 $y\le 0$ 时，$f_Y(y)=\int_{-\infty}^{+\infty} 0\,dx=0$

当 $y>0$ 时，

$$f_Y(y)=\int_0^y e^{-y}\,dx=ye^{-y}$$

故

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



五、小结

1. 在这一讲中，我们与一维情形相对照，介绍了二维随机变量的边缘分布.

2. 请注意联合分布和边缘分布的关系:

由联合分布可以确定边缘分布;

但由边缘分布一般不能确定联合分布.

六、布置作业

《概率统计》标准化作业 (三)