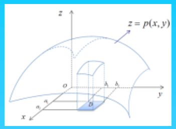


第四章 随机向量及其分布(上)

设n个随机变量$X_1,\cdots,X_n$描述同一个随机现象,一般地它们之间存在一定的联系,因而需要把它们作为一个整体来研究,我们称n个随机变量$X_1,\cdots,X_n$的整体$(X_1,\cdots,X_n)$为n维随机向量（$n-$dimensional random vector）。

本章主要讨论二维随机向量及其分布(包括离散型和连续型),随机变量的独立性,随机向量函数的分布。

第一节 二维随机向量

一.随机向量及其联合分布函数

定义4－1:设随机试验的样本空间为Ω,对每一个$\omega\in\Omega$,有确定的二个实值单值函数$X(\omega),Y(\omega)$与之对应,则称$\left(X(\omega),Y(\omega)\right)$为二维随机向量,简记为$(X,Y)$。

在定义4－1中要注意X和Y是定义在同一个样本空间Ω上的二个随机变量。

例如,有五件产品,其中两件是次品($用a_1, a_2$表示),三件是正品($用b_1, b_2, b_3$表示)。从中依次不放回地任意取出两件,此时随机试验的样本空间为

$$\Omega = \left\{ \begin{array}{l} (a_1, a_2), (a_1, b_1), (a_1, b_2), (a_1, b_3), (a_2, b_1) \\ (a_2, b_2), (a_2, b_3), (b_1, b_2), (b_1, b_3), (b_2, b_3) \\ (a_2, a_1), (b_1, a_1), (b_2, a_1), (b_3, a_1), (b_1, a_2) \\ (b_2, a_2), (b_3, a_2), (b_2, b_1), (b_3, b_1), (b_3, b_2) \end{array} \right\}$$

我们将Ω中的样本点依次为$\omega_1,\omega_2,\cdots,\omega_{20}$。定义随机变量$X$和$Y$如下：

$$X=\begin{cases}0, & \text{如果第一次取出的是正品}\\ 1, & \text{如果第一次取出的是次品}\end{cases},$$

$$Y=\begin{cases}0, & \text{如果第二次取出的是正品}\\ 1, & \text{如果第二次取出的是次品}\end{cases},$$

在一次随机试验中，若出现了样本点(a_1,b_3)（即ω_4），则$X(\omega_4)=1, Y(\omega_4)=0$；若出现了样本点$(b_3,a_1)$（即$\omega_{14}$），则$X(\omega_{14})=0, Y(\omega_{14})=1$。

$$P(X=1,Y=1)=P(X=1)P(Y=1|X=1)$$

$$=\frac{4}{5}\times\frac{1}{4}=\frac{1}{5},$$

$$P(X+Y=1)=P(X=0,Y=1)+P(X=1,Y=0)$$

$$=P(X=0)P(Y=1|X=0)+P(X=1)P(Y=0|X=1)$$

$$=\frac{3}{5}\times\frac{2}{4}+\frac{2}{5}\times\frac{3}{4}=\frac{3}{5},$$

其中$\{X=0,Y=1\}$和$\{X=1,Y=0\}$互不相容。

ω

注意：X和Y都是定义在Ω上的，对Ω中的每一个样本点，X和Y都有一个数与此样本点对应。

现在约定：

$$\left\{X=x_i,Y=y_j\right\}\hat{=}\left\{X=x_i\right\}\bigcap\left\{Y=y_j\right\},$$

其中$\left\{X=x_i\right\}$和$\left\{Y=y_j\right\}$均是样本空间Ω的子集。同理有

$$\left\{X\le x,Y\le y\right\}\hat{=}\{X\le x\}\bigcap\{Y\le y\}。$$

定义4-2:设(X,Y)是一个二维随机向量,且x,y是二个任意实数,则称二元函数

$$F(x,y)=P(X\le x,Y\le y),\forall(x,y)\in\mathbb{R}^2$$

为(X,Y)的联合分布函数(joint distribution function)。

$$F(x,y)=P\left(\left\{\omega\middle|X(\omega)\le x,Y(\omega)\le y\right\}\right)。$$

与一维的情形一样,掌握了联合分布函数也就掌握了二维随机向量的统计规律。

联合分布函数具有下列五个基本性质:

(1) $0 \le F(x,y) \le 1, \forall (x,y) \in \mathbb{R}^2$;

(2) $F(x,y)$对每个自变量都是单调非降的;

(3) 对一切实数x和y,则有

$$F(-\infty, y) = F(x, -\infty) = 0, F(+\infty, +\infty) = 1;$$

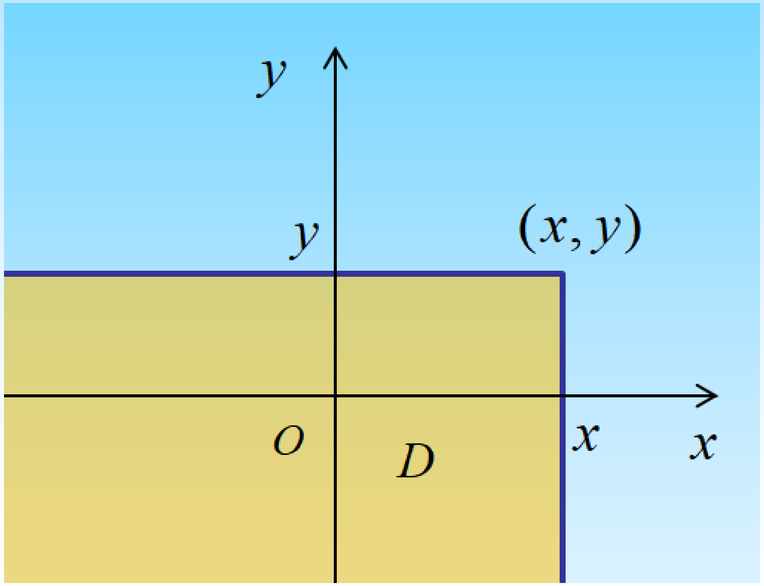
(4)$F(x,y)$对每个自变量都是右连续的;

(5)对一切实数$x_1 < x_2, y_1 < y_2$,则有

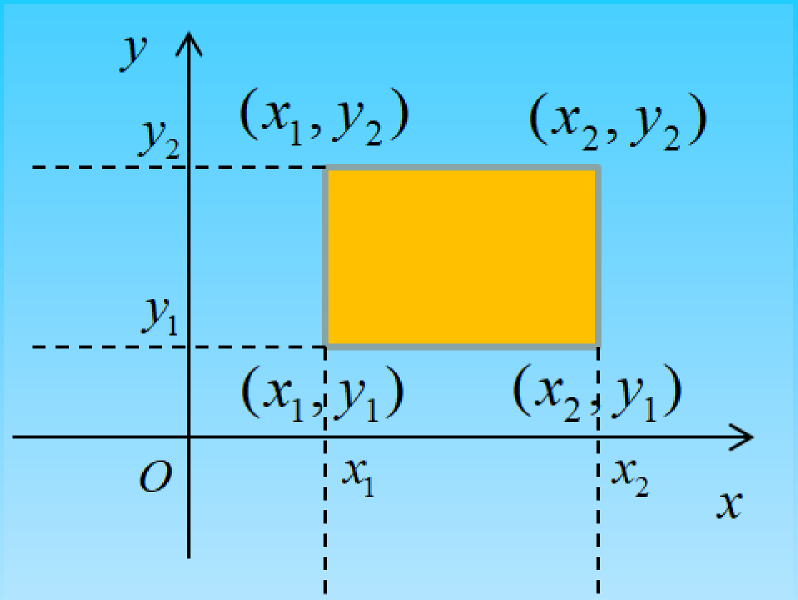
$$F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \ge 0。$$

证明:(1) ~ (4), 类似一维随机变量分布函数的四个基本性质, 下面我们只证(5),

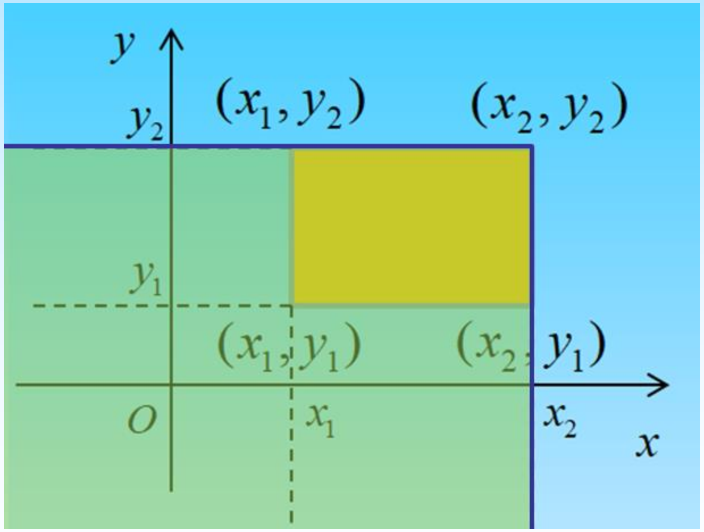
由定义2知, $F(x, y) = P(X \le x, Y \le y)$是$(X, Y)$落在区域$D$内的概率,



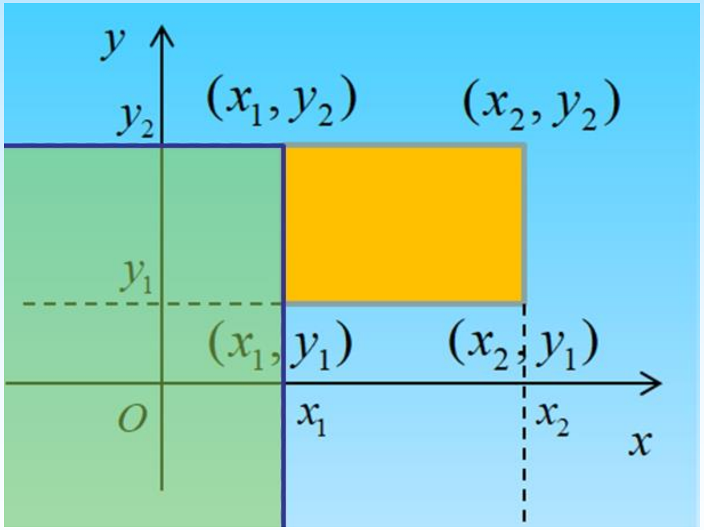
杨勇制作



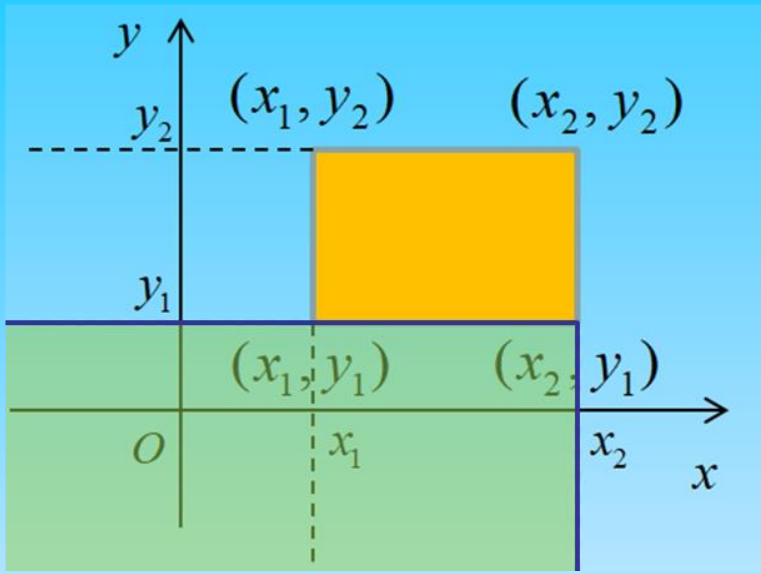
y
y_2
(x_1, y_2)
(x_2, y_2)
y_1
(x_1, y_1)
(x_2, y_1)
O
x_1
x_2
x

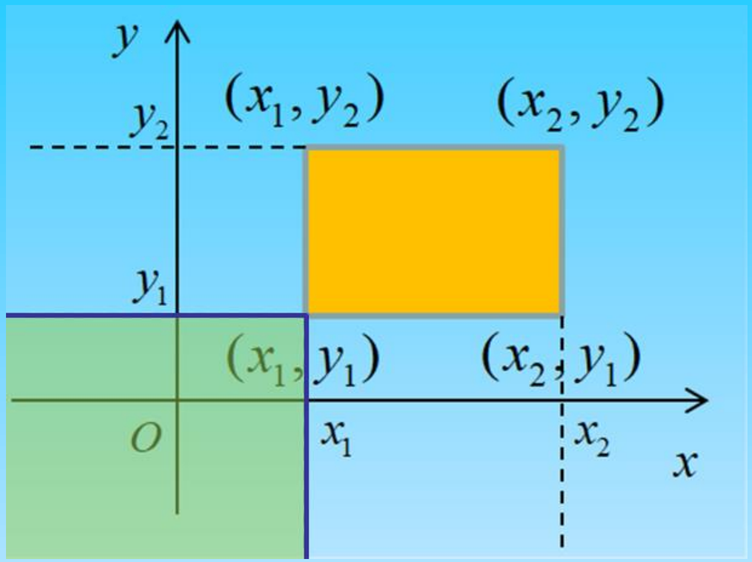


y
y_2
(x_1, y_2)
(x_2, y_2)
y_1
(x_1, y_1)
(x_2, y_1)
O
x_1
x_2
x



y
y_2
(x_1, y_2)
(x_2, y_2)
y_1
(x_1, y_1)
(x_2, y_1)
O
x_1
x_2
x





$$P(x_1 < X \le x_2, y_1 < Y \le y_2)$$
$$= P(X \le x_2, Y \le y_2) - P(X \le x_1, Y \le y_2)$$
$$- P(X \le x_2, Y \le y_1) + P(X \le x_1, Y \le y_1)$$
$$= F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1)$$

这是用$F(x,y)$来计算(X,Y)落在矩形区域$\{x_1 < X \le x_2, y_1 < Y \le y_2\}$概率的公式,

再由概率的非负性,即知(5)成立,即

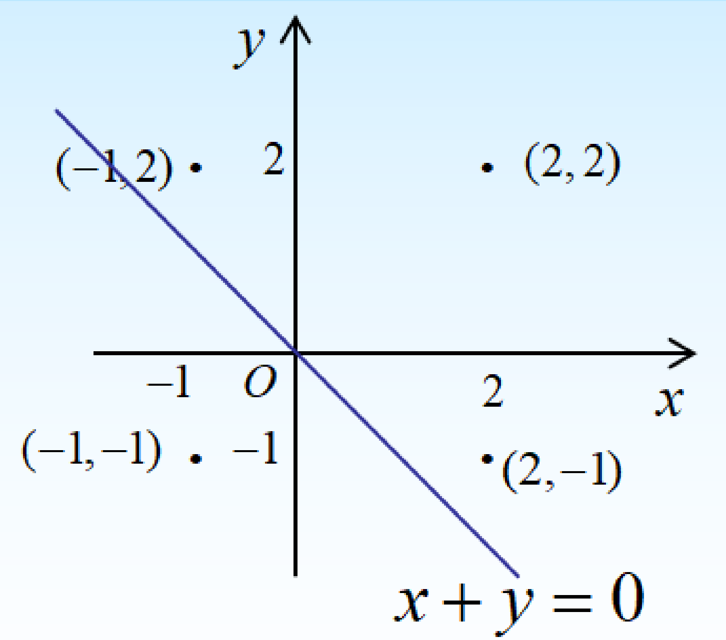
$$F(x_2,y_2)-F(x_1,y_2)-F(x_2,y_1)+F(x_1,y_1)\ge 0。$$

任何一个联合分布函数$F(x,y)$一定具有以上五个基本性质;反之,任何具有以上五个基本性质的二元函数$F(x,y)$必可作为某一二维随机向量(X,Y)的联合分布函数。

习题集77页典型例题

例1:试问二元函数 $F(x,y)=\begin{cases}1, x+y\geq 0\\ 0, x+y<0\end{cases}$

能否成为某二维随机向量的联合分布函数?



解：此二元函数$F(x,y)$满足基本性质(1)～(4)，但因为

$$F(2,2)-F(2,-1)-F(-1,2)+F(-1,-1)$$
$$=1-1-1+0=-1<0,$$

即$F(x,y)$不满足联合分布函数的基本性质(5)，

故$F(x,y)$不能作为某二维随机向量的联合分布函数。

由于联合分布函数$F(x,y)$全面描述了随机向量(X,Y)的统计规律，显然由(X,Y)的联合分布函数$F(x,y)$，我们可以得到随机变量X和Y各自的分布函数。即

$$F_X(x) = P(X \le x) = P(X \le x, \Omega)$$

$$= P(X \le x, Y < +\infty)$$

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

其中 $F(x,+\infty) = \lim\limits_{y \to +\infty} F(x,y)$。

同样，$F_Y(y)=F(+\infty, y)$，其中

$F(+\infty, y)=\lim\limits_{x\to+\infty} F(x, y)$。

我们把$F_X(x), F_Y(y)$分别称为(X, Y)关于X, Y的边际分布函数（marginal distribution function）。

2016 – 2107第二学期试题，选择题

2018-2019学年第1学期，填空题

3.设随机变量X和Y的联合分布函数为$F(x,y)$，而$F_1(x)$和$F_2(x)$分别为X和Y的分布函数，则对任意a,b，概率$P(X>a,Y>b)=$ ______________。

$(A)1-F(a,b)$，$(B)F(a,b)+1-\left[F_1(a)+F_2(b)\right]$，

$(C)1-F_1(a)+F_2(b)$，$(D)F(a,b)-1+\left[F_1(a)+F_2(b)\right]$。

解：$P(X>a,Y>b)=P(a<X<+\infty,b<Y<+\infty)$

$$=F(+\infty,+\infty)-F(a,+\infty)-F(+\infty,b)+F(a,b)$$

$$=1-F_X(a)-F_Y(b)+F(a,b)$$

$$=1-F_1(a)-F_2(b)+F(a,b)\text{。}$$

(B)

我们经常讨论的随机向量有两种类型:离散型和连续型。

二. 二维离散型随机向量及其联合概率分布

定义4－3:若二维随机向量(X,Y)的所有可能取值是有限对或可列无限多对,则称(X,Y)为<u>二维离散型随机向量</u>

（2－dimensionat discrete random vector）。

定义4－4:设(X,Y)的所有可能取值为$(x_i,y_j),i=1,\cdots,m,\cdots,j=1,\cdots,n,\cdots$,则称

$$p_{ij}=P\left(X=x_i,Y=y_j\right),\begin{matrix} i=1,\cdots,m,\cdots \\ j=1,\cdots,n,\cdots \end{matrix}$$

为二维离散型随机向量(X,Y)的联合概率分布(joint probability distribution）。

也常用表格列出:

$X \backslash Y$	y_1	y_2	$\cdots$	y_n	$\cdots$
x_1	p_{11}	p_{12}	$\cdots$	p_{1n}	$\cdots$
x_2	p_{21}	p_{22}	$\cdots$	p_{2n}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_m	p_{m1}	p_{m2}	$\cdots$	p_{mn}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\ddots$

D是一个区域

$$P((X,Y)\in D)=\sum_{(i,j):(x_i,y_j)\in D} p_{ij}$$

$$=\sum_{i:(x_i,y_j)\in D}\ \sum_{j:(x_i,y_j)\in D} p_{ij}$$

联合概率分布完整地描述了离散型随机向量的统计规律。

联合概率分布具有下列两个基本性质：

(1) $p_{ij} \geq 0, \begin{matrix} i = 1, 2, \cdots, m, \cdots \\ j = 1, 2, \cdots, n, \cdots \end{matrix};$

(2) $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1$。

证明：(1) 显然;

(2) 由于

杨勇制作

$$\sum_{i=1}^{\infty}\{X=x_i\}=\sum_{j=1}^{\infty}\{Y=y_j\}=\Omega,$$

$$\left(\sum_{i=1}^{\infty}\{X=x_i\}\right)\cap\left(\sum_{j=1}^{\infty}\{Y=y_j\}\right)=$$

$$\sum_{i=1}^{\infty}\left(\{X=x_i\}\cap\left(\sum_{j=1}^{\infty}\{Y=y_j\}\right)\right)$$

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

$$=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\left\{X=x_i,Y=y_j\right\}=\Omega,$$

再由概率的可列可加性知,

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

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

$$=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}p_{ij}=P(\Omega)=1,$$

即 $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} p_{ij} = 1$。

联合概率分布一定具有以上二个基本性质。

反之,若一串$p_{ij}\,(i=1,\cdots,m,\cdots,\ j=1,\cdots,n,\cdots)$,
具有以上二个性质,则
$p_{ij}\,(i=1,\cdots,m,\cdots,\ j=1,\cdots,n,\cdots)$一定可作为
某一二维离散型随机向量的联合概率分布。

例：设$x_1 < x_2, y_1 < y_2$，且(X,Y)的联合概率分布为

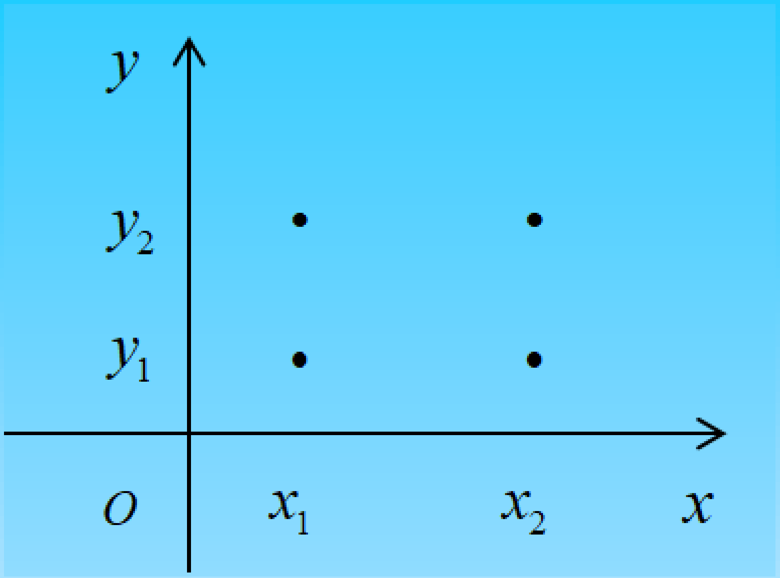
$X \backslash Y$	y_1	y_2
x_1	p_{11}	p_{12}
x_2	p_{21}	p_{22}

求(X,Y)的联合分布函数$F(x,y)$。

解：由题设条件知，$p_{ij} \geq 0, i, j = 1, 2,$

$$\sum_{i=1}^{2}\sum_{j=1}^{2} p_{ij} = 1$$ 。二维离散型随机向量

(X,Y)的联合分布函数为



$$F(x,y)=\sum_{\substack{i:x_i\le x\\ j:y_j\le y}} p_{ij}\ 。$$

当$x<x_1$或者$y<y_1$时, $F(x,y)=0$,

当$x_1\le x<x_2, y_1\le y<y_2$时, $F(x,y)=p_{11}$,

当$x_1 \le x < x_2, y \ge y_2$时，$F(x,y) = p_{11} + p_{12}$,

当$x \ge x_2, y_1 \le y < y_2$时，$F(x,y) = p_{11} + p_{21}$,

当$x \ge x_2, y \ge y_2$时，$F(x,y) = p_{11} + p_{12} + p_{21} + p_{22} = 1$,

所以，

$$F(x,y) = \begin{cases} 0 \quad , \quad x < x_1 \ or \ y < y_1 \\ p_{11}, x_1 \le x < x_2, y_1 \le y < y_2 \\ p_{11} + p_{12}, x_1 \le x < x_2, y \ge y_2 \\ p_{11} + p_{21}, x \ge x_2, y_1 \le y < y_2 \\ 1 \quad , x \ge x_2, y \ge y_2 \end{cases}$$ 。

注1: 用联合分布函数描述离散型随机向量相当麻烦,不如用联合概率分布描述方便。

注2: 离散型随机向量: $F(x, y) \leftrightarrow p_{ij}$ 。

下面举例说明如何求二维离散型随机向量的联合概率分布。

70页例4-1

例$4-1$:箱子里装有a件正品和b件次品。每次从箱子中任取一件产品,共取两次。设随机变量X和Y的定义如下:

$$X=\begin{cases}0, \text{如果第一次取出的是正品}\\ 1, \text{如果第一次取出的是次品}\end{cases},$$

$$Y=\begin{cases}0, 如果第二次取出的是正品\\1, 如果第二次取出的是次品\end{cases},$$

(1) 第一次取出的产品仍放回去;

(2)第一次取出的产品不放回去。

在上述两种情况下分别求出二维随机向量(X,Y)的联合概率分布。

解:(1) $X:0,1,Y:0,1,$

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

$$=\frac{a}{a+b}\cdot\frac{a}{a+b}=\frac{a^2}{(a+b)^2},$$

$$P(X=0,Y=1)=P(X=0)P(Y=1)$$

$$=\frac{a}{a+b}\cdot\frac{b}{a+b}=\frac{ab}{(a+b)^2},$$

$$P(X=1,Y=0)=P(X=1)P(Y=0)$$

$$=\frac{b}{a+b}\cdot\frac{a}{a+b}=\frac{ba}{(a+b)^2},$$

$$P(X=1,Y=1)=P(X=1)P(Y=1)$$

$$=\frac{b}{a+b}\cdot\frac{b}{a+b}=\frac{b^2}{(a+b)^2},$$

即联合概率分布为

$X \backslash Y$	0	1
0	$\frac{a^2}{(a+b)^2}$	$\frac{ab}{(a+b)^2}$
1	$\frac{ba}{(a+b)^2}$	$\frac{b^2}{(a+b)^2}$

(2) $X:0,1, Y:0,1,$

$$P(X=0,Y=0)=P(X=0)P(Y=0\mid X=0)$$

$$=\frac{a}{a+b}\cdot\frac{a-1}{a+b-1}=\frac{a(a-1)}{(a+b)(a+b-1)},$$

$$P(X=0,Y=1)=P(X=0)P(Y=1\mid X=0)$$

$$=\frac{a}{a+b}\cdot\frac{b}{a+b-1}=\frac{ab}{(a+b)(a+b-1)},$$

$$P(X=1,Y=0)=P(X=1)P(Y=0\mid X=1)$$

$$=\frac{b}{a+b}\cdot\frac{a}{a+b-1}=\frac{ba}{(a+b)(a+b-1)},$$

$$P(X=1,Y=1)=P(X=1)P(Y=1\mid X=1)$$

$$=\frac{b}{a+b}\cdot\frac{b-1}{a+b-1}=\frac{b(b-1)}{(a+b)(a+b-1)},$$

即联合概率分布为

$X \backslash Y$	0	1
0	$\dfrac{a(a-1)}{(a+b)(a+b+1)}$	$\dfrac{ab}{(a+b)(a+b+1)}$
1	$\dfrac{ba}{(a+b)(a+b+1)}$	$\dfrac{b(b-1)}{(a+b)(a+b+1)}$

99页习题2

2. 一正整数X随机地在1, 2, 3, 4四个数字中取一个值, 另一个正整数Y随机地在$1 \sim X$中取一个值, 试求(X, Y)的联合概率分布。

解： $X:1,2,3,4, Y:1,2,3,4,$

$$P(X=i,Y=j)=P(X=i)P(Y=j\,|\,X=i)$$

$$=\begin{cases} 0, i<j \\ \dfrac{1}{4}\times\dfrac{1}{i}, i\geq j \end{cases}$$

$$=\begin{cases} 0, i<j \\ \dfrac{1}{4i}, i\geq j \end{cases}, i,j=1,2,3,4。$$

由于(X,Y)的联合概率分布全面地描述了二维离散型随机向量(X,Y)取值的统计规律,因此,当(X,Y)的联合概率分布已知时,我们就可以求出随机变量X和Y的概率分布。

具体地说,已知(X,Y)的联合概率分布为

$$P(X=x_i,Y=y_j)=p_{ij},\begin{matrix} i=1,\cdots,m,\cdots \\ j=1,\cdots,n,\cdots \end{matrix},$$

则X的概率分布为

杨勇制作

$$P(X = x_i) = P(\{X = x_i\} \cap \Omega)$$

$$= P\left(\{X = x_i\} \cap \sum_{j=1}^{\infty} \{Y = y_j\} \right)$$

$$= P\left(\sum_{j=1}^{\infty} \left[\{X = x_i\} \cap \{Y = y_j\} \right] \right)$$

$$= \sum_{j=1}^{\infty} P(\{X = x_i\} \cap \{Y = y_j\})$$

$$= \sum_{j=1}^{\infty} P(X = x_i, Y = y_j)$$

$$=\sum_{j=1}^{\infty} p_{ij}, i=1,\cdots,m,\cdots。$$

同理可求得Y的概率分布为

$$P\left(Y=y_j\right)=\sum_{i=1}^{\infty} p_{ij}, j=1,\cdots,n,\cdots。$$

我们记

$$p_{i\bullet} \triangleq \sum_{j=1}^{\infty} p_{ij}, \quad p_{\bullet j} \triangleq \sum_{i=1}^{\infty} p_{ij}。$$

所以, X的概率分布(又称边际概率分布)为

$$P(X=x_i)=p_{i\bullet}, i=1,\cdots,m,\cdots,$$

Y的概率分布(又称边际概率分布)为

$$P(Y=y_j)=p_{\bullet j}, j=1,\cdots,n,\cdots。$$

我们也可以直接将二个边际概率分布写在(X,Y)的联合概率分布表中:

$X \backslash Y$	y_1	y_2	$\cdots$	y_n	$\cdots$	$p_{i\bullet}$
x_1	p_{11}	p_{12}	$\cdots$	p_{1n}	$\cdots$	$p_{1\bullet}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2n}	$\cdots$	$p_{2\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
x_m	p_{m1}	p_{m2}	$\cdots$	p_{mn}	$\cdots$	$p_{m\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\ddots$	$\vdots$
$p_{\bullet j}$	$p_{\bullet 1}$	$p_{\bullet 2}$	$\cdots$	$p_{\bullet n}$	$\cdots$	1

例如,例4－1中的二个联合概率分布的边际概率分布分别为

$X \backslash Y$	0	1	$p_{i\cdot}$
0	$\dfrac{a^2}{(a+b)^2}$	$\dfrac{ab}{(a+b)^2}$	$\dfrac{a}{a+b}$
1	$\dfrac{ba}{(a+b)^2}$	$\dfrac{b^2}{(a+b)^2}$	$\dfrac{b}{a+b}$
$p_{\cdot j}$	$\dfrac{a}{a+b}$	$\dfrac{b}{a+b}$	

$X \backslash Y$	0	1	$p_{i\cdot}$
0	$\dfrac{a(a-1)}{(a+b)(a+b+1)}$	$\dfrac{ab}{(a+b)(a+b+1)}$	$\dfrac{a}{a+b}$
1	$\dfrac{ba}{(a+b)(a+b+1)}$	$\dfrac{b(b-1)}{(a+b)(a+b+1)}$	$\dfrac{b}{a+b}$
$p_{\cdot j}$	$\dfrac{a}{a+b}$	$\dfrac{b}{a+b}$	

从例4–1中我们还可以发现,(1),(2)两者有完全相同的边际概率分布,而联合概率分布却是不相同的。由此可知,由边际概率分布并不能唯一地确定联合概率分布。

事实上,(X,Y)的联合概率分布还包含有X与Y之间的相互关系的信息,它是边际概率分布所不能提供的。

因而对单个随机变量X与Y的研究并不能代替对二维随机向量(X,Y)整体的研究。

99页习题4

4. 设随机变量X与Y同分布, 且

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

又 $P(XY=0)=1$, 试求$P(X=Y)$。

解:

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

由$P(XY=0)=1$得, $P(XY\neq 0)=0$,

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

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

$$P(X=Y)=P(X=-1,Y=-1)+P(X=0,Y=0)$$
$$+P(X=1,Y=1)$$

$$=0\,。$$

三. 二维连续型随机向量及其联合密度函数

定义4－6:设二维随机向量(X,Y)的联合分布函数为$F(x,y)$,若存在非负可积二元函数$p(x,y)$,使得对任意实数x,y,有

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)dudv,$$

则称(X,Y)为二维连续型随机向量

（2－dimensionat continuous random vector）,

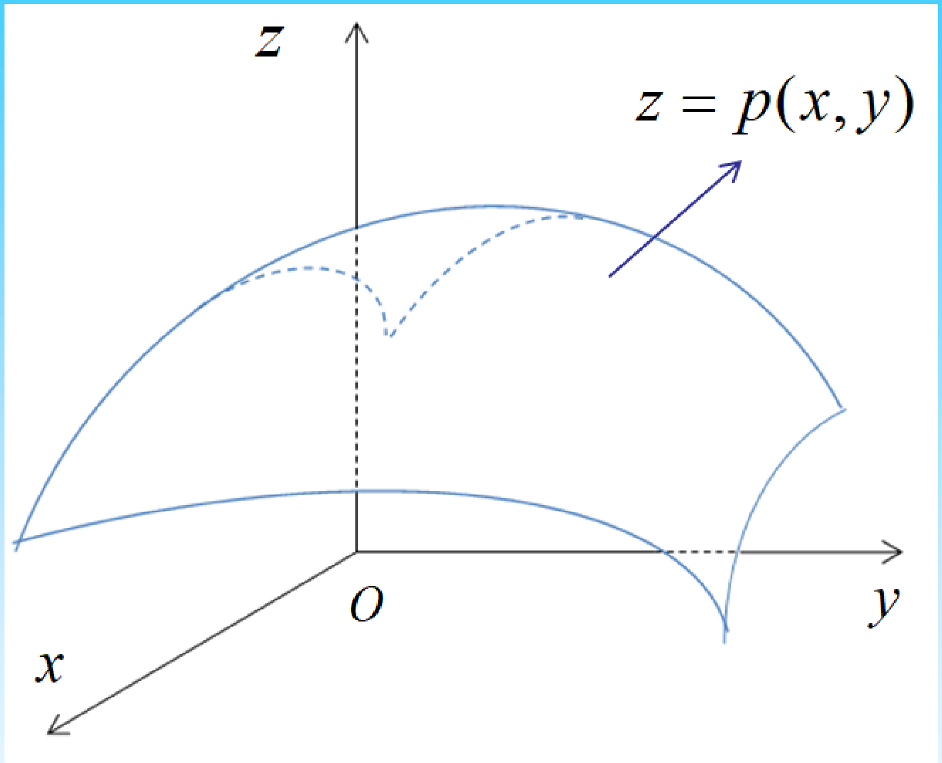
而称$p(x,y)$为二维连续型随机向量(X,Y)的联合密度函数（joint density function）。

联合密度函数完整地描述了二维连续型随机向量的统计规律。

联合密度函数具有下列二个基本性质：

(1) $p(x,y)\geq 0, \forall (x,y)\in \mathbb{R}^2;$

(2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy = 1$。


z
$z = p(x, y)$
O
y
x

证明:(1)显然,(2)

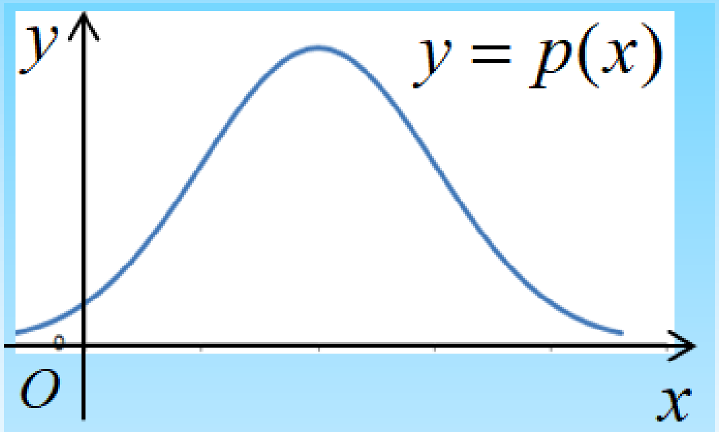
$$1=F(+\infty,+\infty)=\lim_{\substack{x\to+\infty\\ y\to+\infty}}F(x,y)$$

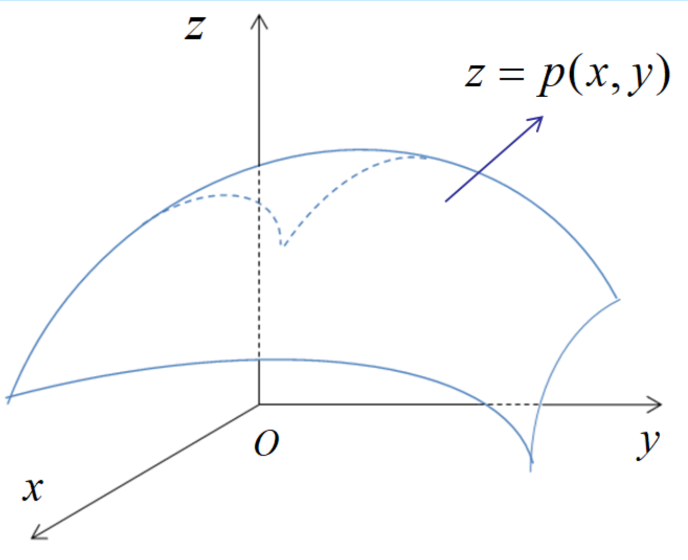
$$=\lim_{\substack{x\to+\infty\\ y\to+\infty}}\int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)dudv$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)dxdy\text{。}$$

注：一维连续型随机变量的密度函数，

二维连续型随机向量的联合密度函数。





联合密度函数一定具有以上二个基本性质;反之,具有以上二个性质的二元函数$p(x,y)$必可作为某一二维连续型随机向量的联合密度函数。

性质4–1:设二维连续型随机向量(X,Y)的联合密度函数为$p(x,y)$,且D为xOy平面上的一个区域,则

$$P\left(\left(X,Y\right)\in D\right)=\iint_{D} p(x,y)dxdy$$

证明:先讨论D为有界区域。

如果D为矩形区域,即

$$D=\{(x,y)\mid a_1<x\le a_2, b_1<y\le b_2\}$$

$$P((X,Y)\in D)=P(a_1<X\le a_2, b_1<Y\le b_2)$$

$$=F(a_2,b_2)-F(a_1,b_2)-F(a_2,b_1)+F(a_1,b_1)$$

$$=\int_{-\infty}^{a_2}\left(\int_{-\infty}^{b_2}p(x,y)dy\right)dx-\int_{-\infty}^{a_1}\left(\int_{-\infty}^{b_2}p(x,y)dy\right)dx$$

$$-\int_{-\infty}^{a_2}\left(\int_{-\infty}^{b_1}p(x,y)dy\right)dx+\int_{-\infty}^{a_1}\left(\int_{-\infty}^{b_1}p(x,y)dy\right)dx$$

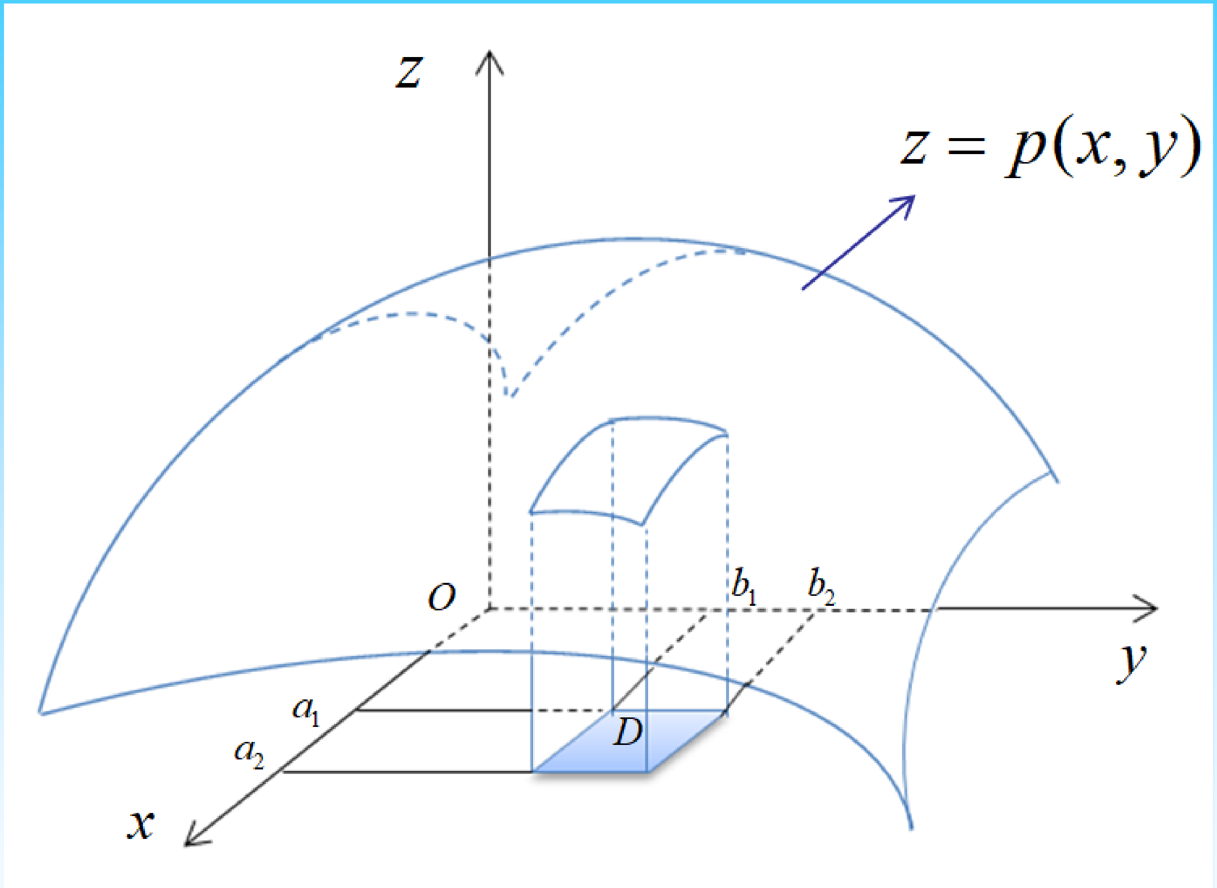
$$=\int_{a_1}^{a_2}\left(\int_{-\infty}^{b_2} p(x,y)dy\right)dx-\int_{a_1}^{a_2}\left(\int_{-\infty}^{b_1} p(x,y)dy\right)dx$$

$$=\int_{-\infty}^{b_2}\left(\int_{a_1}^{a_2} p(x,y)dx\right)dy-\int_{-\infty}^{b_1}\left(\int_{a_1}^{a_2} p(x,y)dx\right)dy$$

$$=\int_{b_1}^{b_2}\left(\int_{a_1}^{a_2} p(x,y)dx\right)dy$$

$$=\int_{a_1}^{a_2}\int_{b_1}^{b_2} p(x,y)dxdy$$

$$=\iint_D p(x,y)dxdy 。$$


z
z = p(x, y)
O
b1
b2
y
a1
a2
D
x

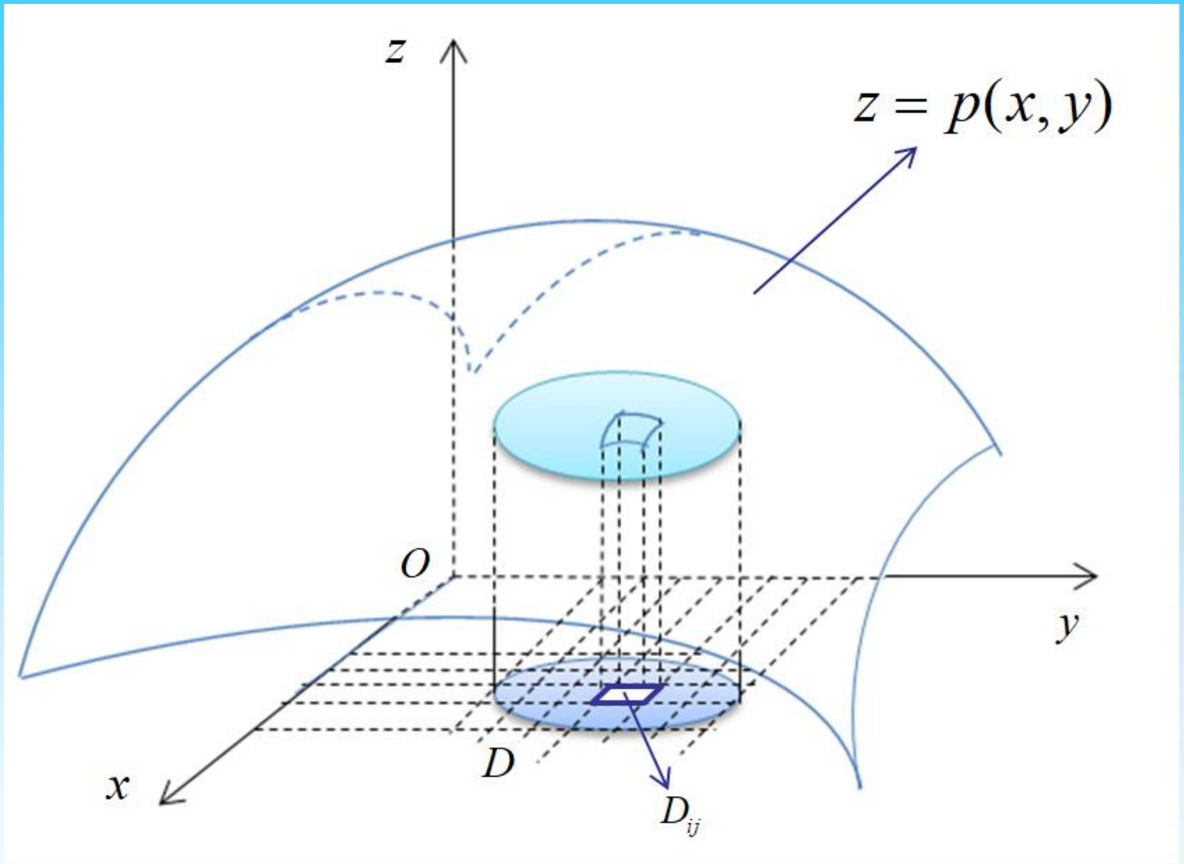
如果D不是矩形区域,可用平行于坐标轴的直线将D分成若干小矩形,即在x轴取分点$x_0, x_1, \cdots, x_m$,在y轴取分点$y_0, y_1, \cdots, y_n$,对位于D内每个小矩形D_{ij},则

$$P\left(\left(X, Y\right) \in D_{ij}\right) = P\left(x_i < X \le x_i + \Delta x_i, y_j < Y \le y_j + \Delta y_j\right)$$

$$= \iint_{D_{ij}} p(x, y) dx dy$$

$$\approx p(x_i, y_j) \Delta x_i \Delta y_j,$$

其中小矩形D_{ij}的直径很小,$\Delta x_i = x_i - x_{i-1}, \Delta y_j = y_j - y_{j-1}$。



z
$z = p(x, y)$
O
y
x
D
D_{ij}

把这些体积(概率)加起来,再令这些小矩形的最大直径λ趋于零得

$$P\left(\left(X,Y\right)\in D\right)=\lim_{\lambda\to 0}\sum_{i=1}^{m}\sum_{j=1}^{n}p(x_i,y_j)\Delta x_i\Delta y_j$$

$$=\iint_D p(x,y)dxdy$$

此性质的几何意义是:(X,Y)落入区域D的概率等于以区域D为底,D的边界为准线,母线平行于z轴,曲面$z=p(x,y)$为顶的曲顶柱体的体积。

再讨论D无界区域。

由于$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy=1$，在$x$轴取分点$x_0,x_1,\cdots,x_m$，在$y$轴取分点$y_0,y_1,\cdots,y_n$，使得

$P(X<x_0)\approx 0, P(X>x_m)\approx 0,$

$P(Y<y_0)\approx 0, P(Y>y_n)\approx 0$。

这样，再利用有界区域的证明，所以对D为无界区域也成立，即

$$P\left(\left(X,Y\right)\in D\right)=\iint_D p(x,y)dxdy。$$

注1: 基本性质$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x, y)dxdy = 1$

的几何意义。

注2: (X, Y)落在一个点上或者落在一条曲线上的概率为0。

性质4－2:设$F(x,y)$为二维连续型随机向量的联合分布函数,则$F(x,y)$处处连续。

证明:任意实数x,y,

$$\lim_{(\Delta x,\Delta y)\to(0,0)} F\left(x+\Delta x,y+\Delta y\right)$$

$$=\lim_{(\Delta x,\Delta y)\to(0,0)} \int_{-\infty}^{x+\Delta x}\int_{-\infty}^{y+\Delta y} p(u,v)dudv$$

$$=\int_{-\infty}^{x}\int_{-\infty}^{y} p(u,v)dudv$$

$$=F(x,y),$$

即$F(x,y)$在(x,y)处连续，又(x,y)的任意性可知，$F(x,y)$处处连续。

二维连续型随机向量的名称也由此性质得。

性质4－3：设$F(x,y)$和$p(x,y)$分别是二维连续型随机向量的联合分布函数和联合密度函数，则在$p(x,y)$的连续点上，有

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = p(x,y)。$$

证明：
$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)dudv$$
$$=\int_{-\infty}^{x}\left(\int_{-\infty}^{y}p(u,v)dv\right)du$$
$$\frac{\partial F(x,y)}{\partial x}=\int_{-\infty}^{y}p(x,v)dv$$
$$\frac{\partial^2 F(x,y)}{\partial x\partial y}=p(x,y)\text{。}$$

由性质4−1的几何意义还可知,连续型随机向量(X,Y)取任何一对数的概率等于零。所以,

$$P(a_1 < X \le a_2, b_1 < Y \le b_2)$$

$$= P(a_1 < X < a_2, b_1 < Y < b_2)$$

$$= P(a_1 \le X < a_2, b_1 \le Y < b_2)$$

$$= P(a_1 \le X \le a_2, b_1 \le Y \le b_2)$$

$$= \int_{a_1}^{a_2} \int_{b_1}^{b_2} p(x, y) dxdy 。$$

性质4－3表明，对于二维连续型随机向量(X,Y)而言，当已知联合分布函数$F(x,y)$时，用求混合二阶偏导数可得其联合密度函数。

在$p(x,y)$不连续点上，即$F(x,y)$的混合偏导数不存在点上，$p(x,y)$的值可任意用一个非负常数给出，这不会影响以后有关事件概率的计算结果。

另外,当已知二维连续型随机向量(X,Y)的联合密度函数$p(x,y)$,则由定义

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)dudv$$

即可求得联合分布函数 。

注:

和一维连续型随机变量一样,连续型随机向量(X,Y)的联合密度函数$p(x,y)$不唯一,但是,它们都对应唯一的联合分布函数$F(x,y)$。在这种意义上说,它们一一对应。

$$F(x,y) \leftrightarrow p(x,y)$$

75页例4-3

例$4-3$:设随机向量(X,Y)的联合密度函数为

$$p(x,y)=\begin{cases} x^2+cxy, 0\le x\le 1, 0\le y\le 2 \\ 0 \quad , \qquad\qquad 其他 \end{cases},$$

求 (1) 常数; (2)$P\left(X+Y\ge 1\right)$;

(3)联合分布函数$F(x,y)$。

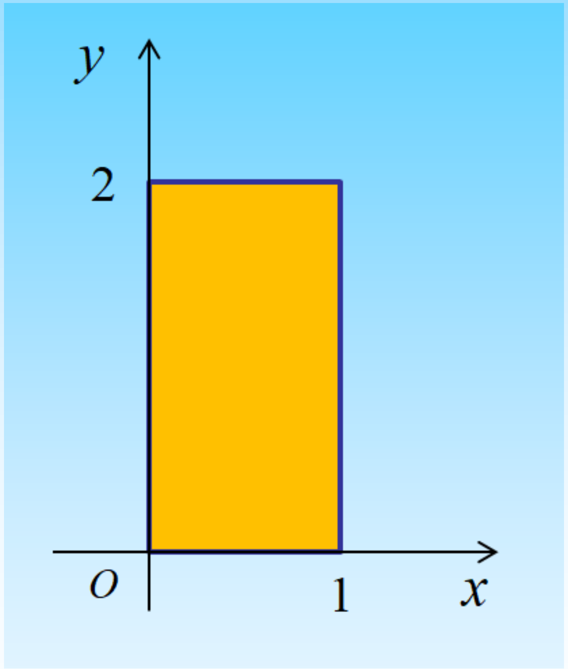
解：(1) 由联合密度函数的基本性质知：

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy = \int_0^1\int_0^2 (x^2+cxy)dxdy$$

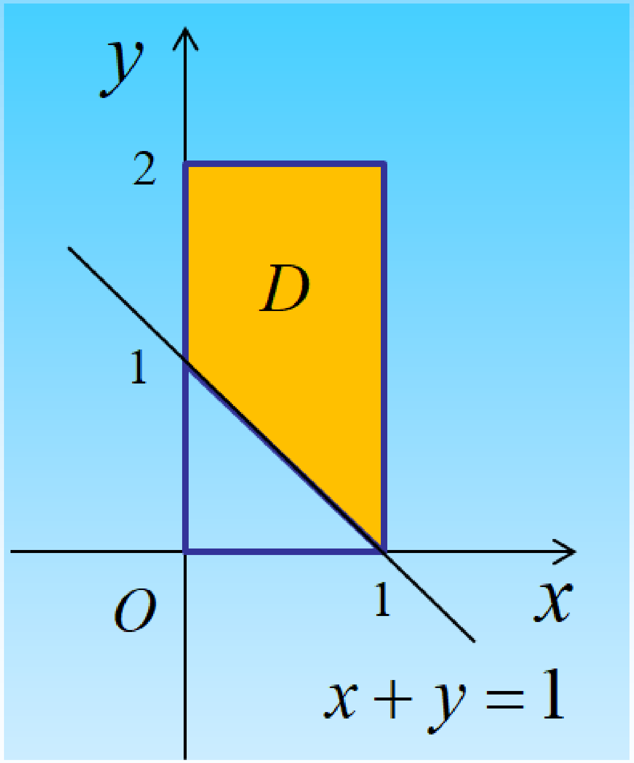
$$= \int_0^1 (2x^2+2cx)dx$$

$$= \frac{2}{3}+c = 1,$$

从而得　$c=\dfrac{1}{3}$;



(2) 由于在区域$\{(x,y)|0\le x\le 1, 0\le y\le 2\}$外, $p(x,y)=0$,



所以,在区域$\left\{(x,y)\middle|x+y\geq 1\right\}$上积分等价于在区域$D$上的积分,

$$\{X+Y\geq 1\}=\left\{(X,Y)\in\left\{(x,y)\middle|x+y\geq 1\right\}\right\}$$

$$P(X+Y\geq 1)=\iint\limits_{x+y\geq 1}p(x,y)dxdy$$

$$=\iint\limits_{D}p(x,y)dxdy$$

$$=\int_0^1 dx\int_{1-x}^2\left(x^2+\frac{xy}{3}\right)dy$$

$$= \int_0^1 \left(\frac{5}{6} x^3 + \frac{4}{3} x^2 + \frac{1}{2} x \right) dx$$

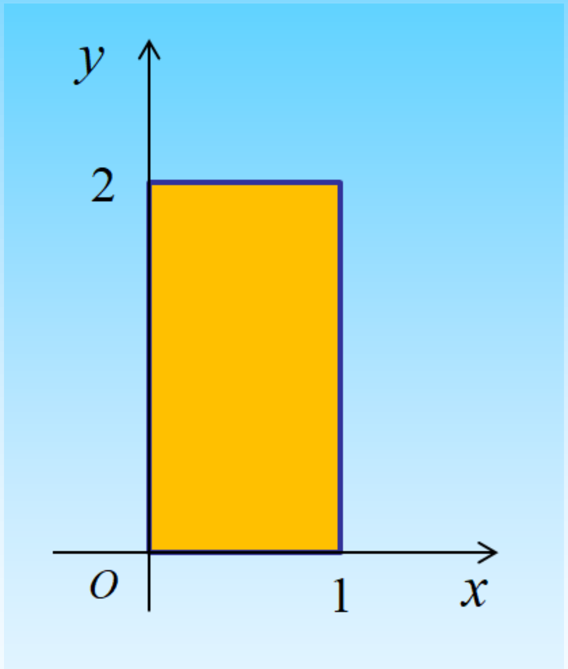
$$= \frac{65}{72};$$

（3）由于

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} p(u, v) du dv$$

所以, 当$x < 0$或$y < 0$时,

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)dudv=0$$

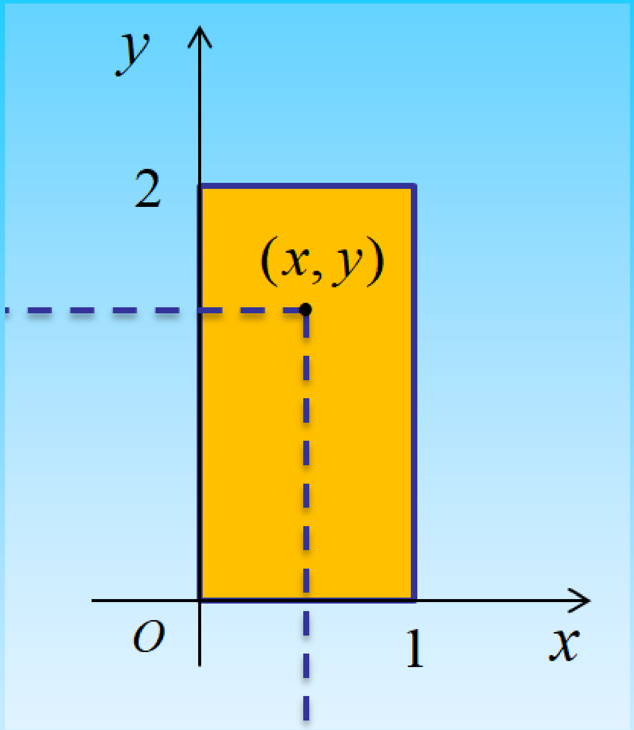


当 $0 \le x < 1, 0 \le y < 2$ 时,

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(u,v)dudv$$

$$= \int_{0}^{x}\int_{0}^{y} (u^2 + \frac{1}{3}uv)dudv$$

$$= \frac{x^3 y}{3} + \frac{x^2 y^2}{12},$$

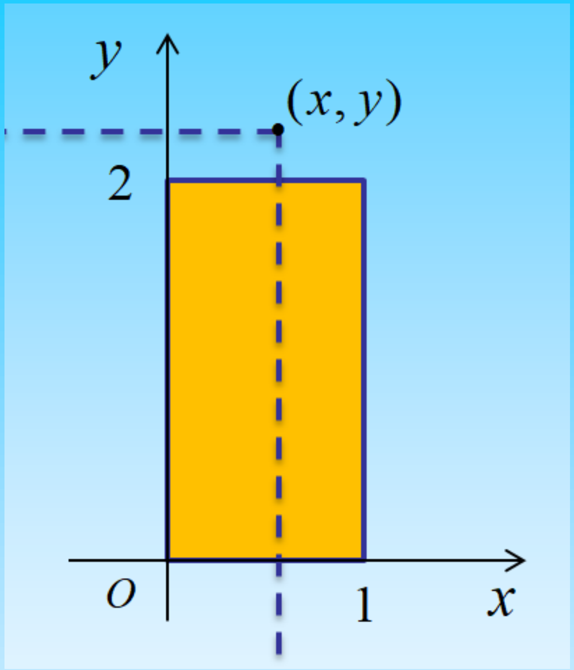


当$0 \le x < 1, y \ge 2$时,

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(u,v)dudv$$

$$= \int_{0}^{x}\int_{0}^{2} (u^2 + \frac{1}{3}uv)dudv$$

$$= \frac{2}{3}x^3 + \frac{1}{3}x^2 ,$$

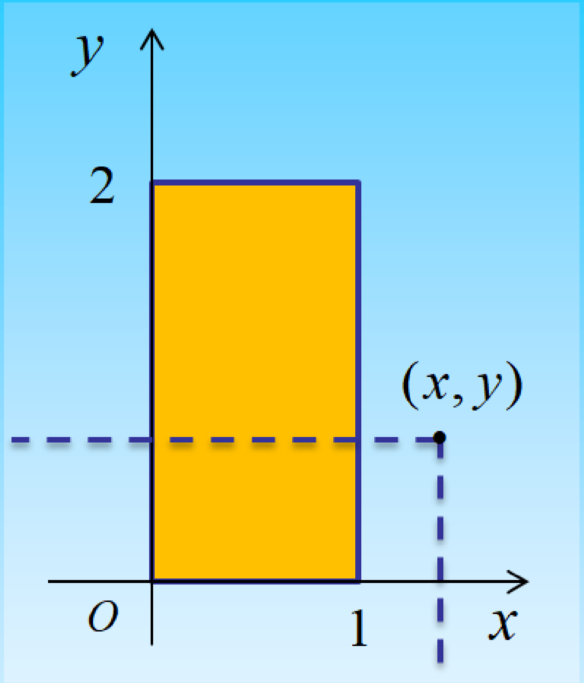


当 $x \geq 1, 0 \leq y < 2$ 时,

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(u,v)dudv$$

$$= \int_{0}^{1}\int_{0}^{y} (u^2 + \frac{1}{3}uv)dudv$$

$$= \frac{y}{3} + \frac{y^2}{12},$$

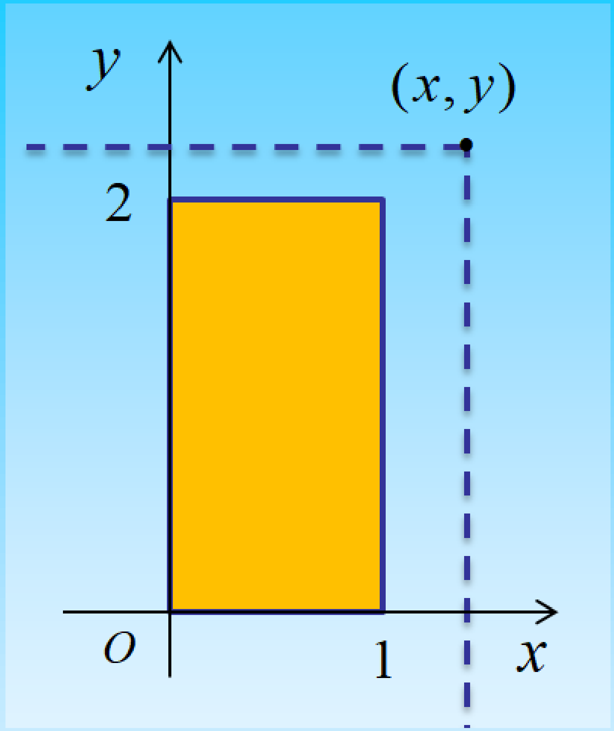


当 $x \geq 1, y \geq 2$ 时，

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(u,v)\,du\,dv$$

$$= \int_{0}^{1}\int_{0}^{2} (u^2 + \frac{1}{3}uv)\,du\,dv$$

$$= 1,$$



杨勇制作

综上所述,

$$F(x,y)=\begin{cases} 0\quad, & x<0\text{或}y<0 \\ \dfrac{x^3y}{3}+\dfrac{x^2y^2}{12}, & 0\le x<1, 0\le y<2 \\ \dfrac{x^2}{3}+\dfrac{2x^3}{3}, & 0\le x<1, y\ge 2 \\ \dfrac{y}{3}+\dfrac{y^2}{12}, & x\ge 1, 0\le y<2 \\ 1\quad, & x\ge 1, y\ge 2 \end{cases}$$ 。

注:就这题而言,已知$F(x,y)$,求$p(x,y)$。

注1：用联合分布函数描述连续型随机向量相当麻烦，不如用联合密度函数描述方便。

注2：连续型随机向量：$F(x, y) \leftrightarrow p(x, y)$。

由于(X,Y)的联合密度函数全面地描述了二维连续型随机向量(X,Y)的统计规律，因此，当(X,Y)的联合密度函数已知时，我们就可以求出随机变量X和Y的密度函数。

具体地说，已知(X,Y)的联合密度函数为$p(x,y)$，则

$$F_X(x) = P(X \le x)$$

$$= P(X \le x, Y < +\infty)$$

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

$$= \int_{-\infty}^{x} \left(\int_{-\infty}^{+\infty} p(u,v)\,dv \right) du,$$

$$p_X(x) = \frac{dF_X(x)}{dx} = \int_{-\infty}^{+\infty} p(x,v)\,dv$$

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

同样可求得Y的密度函数为

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

定义4－7:我们称

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

为二维连续型随机向量(X,Y)关于X的边际密度函数(marginal density function)。

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

为二维连续型随机向量(X,Y)关于Y的边际密度函数。

78页例4-4

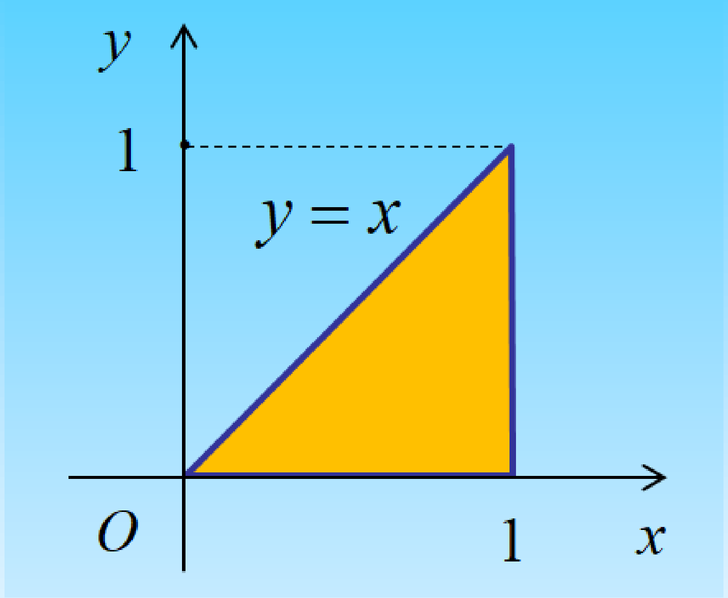
例$4-4$:设(X,Y)的联合密度函数为

$$p(x,y)=\begin{cases}3x, 0<x<1, 0<y<x \\ 0 \quad , \quad \text{其他}\end{cases},$$

求(X,Y)的边际密度函数$p_X(x), p_Y(y)$。

解：先画出区域$\{(x,y)|0<x<1, 0<y<x\}^{\uparrow}$

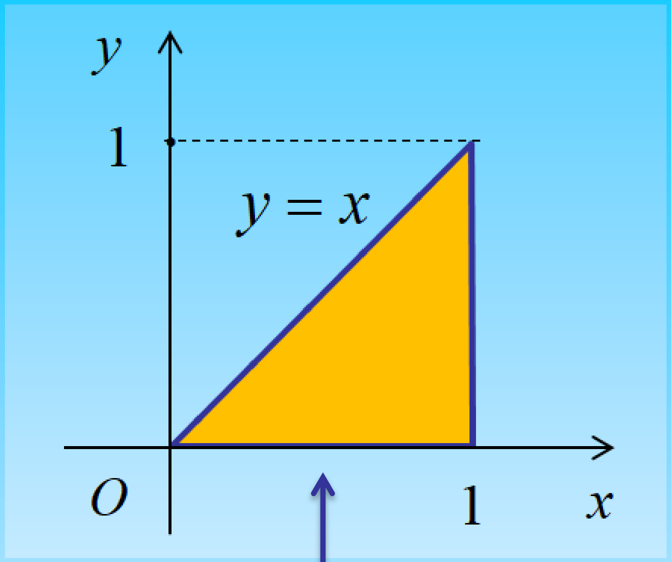
的图形，



另一个表示$\{(x,y)|0<y<1, y<x<1\}^{\rightarrow}$

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

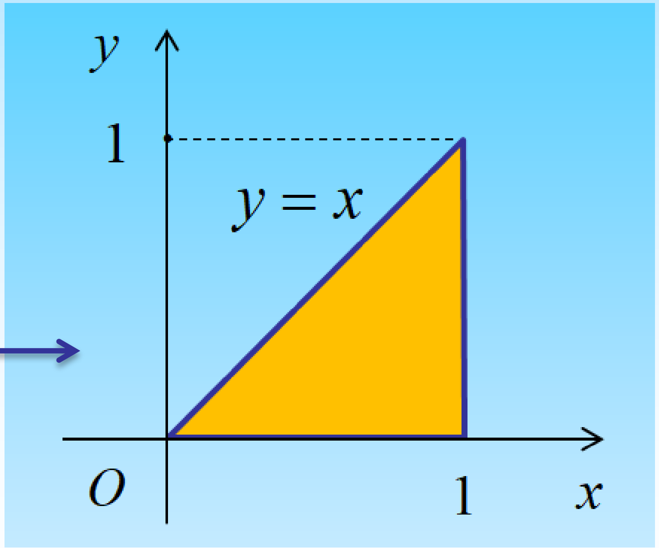
$$=\begin{cases}\int_0^x 3x\,dy, & 0<x<1\\ 0\ , & \text{其他}\end{cases}$$



$$=\begin{cases}3x^2, & 0<x<1\\ 0\ , & \text{其他}\end{cases},$$

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

$$=\begin{cases}\int_y^1 3x\,dx, & 0<y<1\\ 0\ , & \text{其他}\end{cases}$$



$$=\begin{cases}\frac{3}{2}-\frac{3}{2}y^2, & 0<y<1\\ 0, & \text{其他}\end{cases}。$$

与二维离散型随机向量一样,对于二维连续型随机向量(X,Y)而言,二个边际密度函数也不能唯一地确定联合密度函数。

99页习题7

杨勇制作

7. 设$g(x)$为某随机变量的密度函数，且$g(x)=0(x\le 0)$，问二元函数

$$p(x,y)=\begin{cases}\dfrac{2g\left(\sqrt{x^2+y^2}\right)}{\pi\sqrt{x^2+y^2}}, & x>0, y>0\\ 0, & 其它\end{cases}$$

能否作为某二维连续型随机向量的联合密度函数？

解:$g(x)$是密度函数,则

$$\int_{-\infty}^{+\infty} g(x)dx = \int_{0}^{+\infty} g(x)dx = 1。$$

由$g(x)$非负,可知$p(x,y)$也非负,又

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy = \int_{0}^{+\infty}\int_{0}^{+\infty} \frac{2g\left(\sqrt{x^2+y^2}\right)}{\pi\sqrt{x^2+y^2}}dxdy$$

$$\overset{\begin{cases} x=r\cos\theta \\ y=r\sin\theta \end{cases}}{=} \int_{0}^{\frac{\pi}{2}} d\theta \int_{0}^{+\infty} \frac{2g(r)}{\pi r} r dr$$

$$= \int_0^{+\infty} g(r)dr = 1,$$

二元函数$p(x, y)$满足联合密度函数的二个基本性质,所以可以作为一个二维连续性随机向量的联合密度函数。