第3章 多维随机变量及其分布

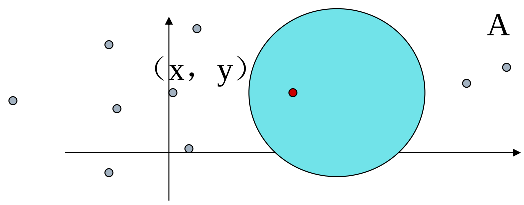
- 二维随机变量及其联合分布
- 边缘分布与独立性

前面我们讨论的是随机实验中单独的一个随机变量，又称为一维随机变量；然而在许多实际问题中，常常需要同时研究一个试验中的两个甚至更多个随机变量。

例如 抽样调查15-18岁青少年的身高 X与体重 Y，以研究当前该年龄段青少年的身体发育情况。

此时我们需要研究的不仅仅是X及Y各自的性质，更需要了解这两个随机变量的相互依赖和制约关系。因此， 我们将二者作为一个整体来进行研究，记为(X, Y),称为二维随机变（向）量。

二维随机变量(X, Y)的取值可看作平面上的点

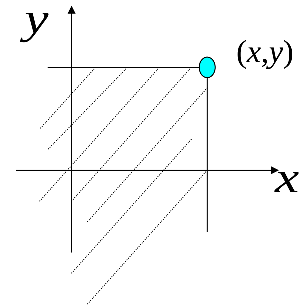


第1节

二维随机变量的联合分布

二维随机变量的联合分布函数

■ **定义** 若（X，Y）是二维随机变量，对于任意的实数x, y.



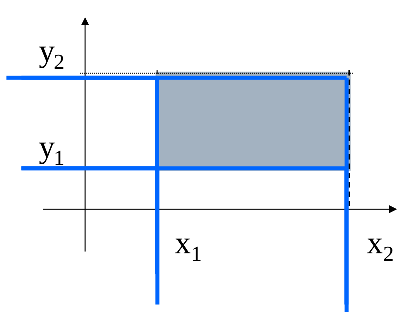
$$F(x,y)=P\{X\le x,Y\le y\}$$

称为二维随机变量的联合分布函数

5

联合分布函数表示矩形域概率

$P（x_1 < X \le x_2，y_1 < Y \le y_2）$



$F(x_2, y_2)$

$-F(x_2, y_1)$

$-F(x_1, y_2)$

$+F(x_1, y_1)$

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

6

联合分布函数的性质

1. $F(x,y)$ 分别关于x和y单调不减；
2. $F(x,y)$ 关于x右连续，关于y右连续；
3. $F(x,y)$ 的值域为[0, 1]，并且

$$F(-\infty, y) = 0$$

$$F(x, -\infty) = 0$$

$$F(+\infty, +\infty) = 1$$

7

第2节

二维离散型随机变量

8

二维离散型随机变量

- 定义

若二维随机变量（X，Y）的所有可能取值只有限个或可列个，则称（X，Y）为二维离散型随机变量。

如何表达（X，Y）的取值规律呢？

9

（X，Y）的联合概率分布（分布律）

- 公式法

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

- 表格法（常见形式）

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

性质

$$0\le p_{ij}\le 1$$

$$\sum_i\sum_j p_{ij}=1$$

10

例1 一个口袋中有三个球，依次标有数字1，2，2，从中任取一个，不放回袋中，再任取一个。设每次取球时，各球被取到的可能性相等.以X、Y分别记第一次和第二次取到的球上标有的数字，求(X,Y)的联合分布律.

解 (X,Y)的可能取值为(1，2)，(2，1)，(2，2).

P{X＝1，Y＝2}＝(1/3)×(2/2)＝1/3

P{X＝2，Y＝1}＝(2/3)×(1/2)＝1/3

P{X＝2，Y＝2}= (2/3)×(1/2)＝1/3

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

11

离散型随机变量分布函数的计算

$$F(x,y)=P\{X\le x,Y\le y\}=\sum_{x_i\le x}\sum_{y_j\le y}p_{ij}$$

12

作业

P91 习题3
4

13

第3节

二维连续型随机变量

14

二维连续型随机变量的联合概率密度

- 定义　设二维随机变量（X，Y）的分布函数F（X，Y），若存在非负函数 $f(x, y)$，使对任意实数x，y，都有

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

则称(X,Y)是二维连续型随机变量，$f(x, y)$ 称为二维随机变量(X,Y)的联合概率密度函数.

15

联合概率密度函数的性质

- 非负性　$f(x,y)\geq 0$
- $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=F(+\infty,+\infty)=1$
- 若F（x，y）存在二阶连续偏导数，则

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

- 设D为平面区域

$$P\{(x,y)\in D\}=\iint_D f(x,y)d\sigma$$

16

例 设二维随机变量 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases} ke^{-(2x+3y)} & x>0, y>0 \\ 0 & \text{其它} \end{cases}$$

(1) 确定常数 **k**;

(2) 求 (X,Y) 的分布函数;

(3) 求 $P\{0<X\le 4, 0<Y\le 1\}$

(4) 求 $P\{X<Y\}$

17

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

$$=\int_0^{+\infty}\int_0^{+\infty} ke^{-(2x+3y)}dxdy$$

$$=k\int_0^{+\infty} e^{-2x}dx\int_0^{+\infty} e^{-3y}dy$$

$$=k[-\frac{1}{2}e^{-2x}]_0^{+\infty}[-\frac{1}{3}e^{-3y}]_0^{+\infty}$$

$$=k\cdot\frac{1}{6}=1$$

所以 $k=6$

18

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

当 $x\le 0,$或$y\le 0$ 时， $F(x,y)=0$

当 $x>0,$且$y>0$ 时，

$$F(x,y)=\int_0^x\int_0^y 6e^{-(2u+3v)}dudv \;=(1-e^{-2x})(1-e^{-3y})$$

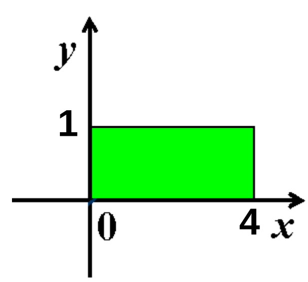
所以， $F(x,y)=\begin{cases}(1-e^{-2x})(1-e^{-3y}), & (x>0,y>0)\\ 0 & \text{其他}\end{cases}$

19

(3) $P\{0<X\le 4,\ \ 0<Y\le 1\}$

$$=\int_0^1\int_0^4 6e^{-(2x+3y)}dxdy$$

$$=(1-e^{-8})(1-e^{-3})\approx 0.95$$



解法2

$P\{0<X\le 4,\ \ 0<Y\le 1\}$

$$=F(4,1)+F(0,0)-F(4,0)-F(0,1)$$

$$=F(4,1)=(1-e^{-8})(1-e^{-3})\approx 0.95$$

20

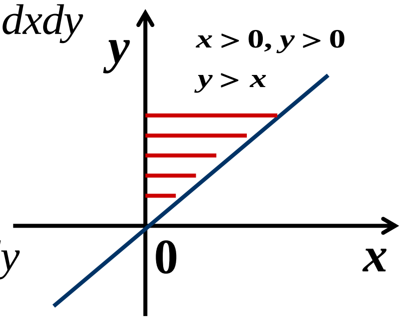
(4) $P\{X<Y\}=\iint_D f(x,y)dxdy$

$$=\iint_{x<y} f(x,y)dxdy$$

$$=\int_0^{+\infty}\left[\int_0^{y}6e^{-(2x+3y)}dx\right]dy$$

$$=\int_0^{+\infty}3e^{-3y}[1-e^{-2y}]dy$$

$$=\int_0^{+\infty}3e^{-3y}dy-\int_0^{+\infty}3e^{-5y}dy=1-\frac{3}{5}=\frac{2}{5}$$



21

作业

P91 习题3
7，8

22

第4节

常见多维随机变量

23

1. 多项分布

在独立重复试验中，设每次实验必有$A_1, A_2, \cdots, A_r$之一发生，且事件A_i在每次实验中发生的概率为p_i，记X_i为A_i出现的次数，则 $X_1, X_2, \cdots, X_r$ 的分布律为

$$P\{X_1 = n_1, X_2 = n_2, \cdots, X_r = n_r\}$$

$$= \frac{n!}{n_1! n_2! \cdots n_r!} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$$

其中 $p_1 + p_2 + \cdots + p_r = 1, n_1 + n_2 + \cdots + n_r = n$

这种联合分布律称为多项分布，记为

$$M(n, p_1, p_2, \cdots, p_r)$$

24

2. 二维均匀分布

设二维随机变量 (X,Y)的概率密度为

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

则称 (X,Y)在D上服从均匀分布，记为

$$(X,Y)\sim U(D)$$

25

3. 二维正态分布

设二维随机变量 (X,Y) 的概率密度为

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

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

其中 $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 均为参数 $\sigma_1>0,\sigma_2>0,-1<\rho<1$

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

$$N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$$

26

第5节

边缘分布

27

边缘分布 marginal distribution

二维随机变量 (X,Y) ,是两个随机变量视为一个整体，来讨论其取值规律的，我们可用分布函数来描述其取值规律。

$$F(x,y)=P\{X\le x,Y\le y\}$$

问题：能否由二维随机变量的分布来确定两个一维随机变量的取值规律呢？如何确定呢？

——边缘分布问题

28

边缘分布 marginal distribution

设二维随机变量 (X,Y) 的分布函数为 $F(x,y)$，

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

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

依次称为二维随机变量 (X,Y)关于 X 和关于 Y 的边缘分布函数.

$$\boxed{F_X(x)=F(x,+\infty)} \qquad \boxed{F_Y(y)=F(+\infty,y)}$$

29

二维离散型随机变量的边缘分布

如果二维离散型随机变量（**X,Y**）的联合分布律为

$$P\{X=x_i,Y=y_j\}=p_{ij} \qquad i,j=1,2,3,\cdots$$

即

X \ Y	y_1	y_2	y_3	…
x_1	p_{11}	p_{12}	p_{13}	…
x_2	p_{21}	p_{22}	p_{23}	…
x_3	p_{31}	p_{32}	p_{33}	…
…	…	…	…	…

30

二维离散型随机变量的边缘分布

Y / X	y_1	y_2	y_3	…	$P_{i\cdot}$
x_1	p_{11}	p_{12}	p_{13}	…	$P_{1\cdot}$
x_2	p_{21}	p_{22}	p_{23}	…	$P_{2\cdot}$
x_3	p_{31}	p_{32}	p_{33}	…	$P_{3\cdot}$
…	…	…	…	…	…
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	$p_{\cdot 3}$	…	

关于X的边缘分布律 $p_{i\bullet} = P\{X = x_i\} = \sum_j p_{ij}$

关于Y的边缘分布律 $p_{\bullet j} = P\{Y = y_j\} = \sum_i p_{ij}$

31

二维离散型随机变量的边缘分布

关于X的边缘分布列

X	x_1	x_2	x_3	…
概率	$P_{1\cdot}$	$P_{2\cdot}$	$P_{3\cdot}$	…

$$p_{i\bullet} = P\{X = x_i\} = \sum_j p_{ij}$$

关于Y的边缘分布列

Y	y_1	y_2	y_3	…
概率	$P_{\cdot 1}$	$P_{\cdot 2}$	$P_{\cdot 3}$	…

$$p_{\bullet j} = P\{Y = y_j\} = \sum_i p_{ij}$$

32

例1　设二维离散型随机变量（X,Y）的联合分布律为

X \ Y	0	1	1/3
-1	0	1/3	1/12
0	1/6	0	0
2	5/12	0	0

求关于X、Y的边缘分布

33

解　（X，Y）的联合分布列

X \ Y	0	1	1/3	$P_{i\bullet}$
-1	0	1/3	1/12	5/12
0	1/6	0	0	1/6
2	5/12	0	0	5/12
$P_{\bullet j}$	7/12	1/3	1/12	

关于X的边缘分布为

X	-1	0	2
$P_{i\cdot}$	5/12	1/6	5/12

关于Y的边缘分布

Y	0	1	1/3
$P_{\cdot j}$	7/12	1/3	1/12

34

二维连续型随机变量的边缘分布

关于 X 的边缘分布函数为

$$F_X(x)=F(x,+\infty)=\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty} f(x,y)dy\right]dx$$

- 关于X的边缘分布密度为 $f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy$

关于 Y 的边缘分布函数为

$$F_Y(y)=F(+\infty,y)=\int_{-\infty}^{y}\left[\int_{-\infty}^{+\infty} f(x,y)dx\right]dy$$

- 关于Y的边缘分布密度为 $f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx$

35

例2 设（X, Y) 的联合分布密度为

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

（1）求k值

(2) 求关于X和Y的边缘密度

（3）求概率P(X+Y< -1) 和 P(X>1/2)

36

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

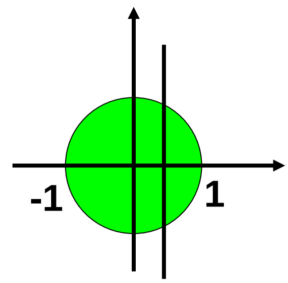
得 $\iint\limits_{x^2+y^2\le 1} kdxdy = k\pi = 1 \Longrightarrow k = \frac{1}{\pi}$　均匀分布

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

当 $x \in [-1,1]$ 时

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy$$

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

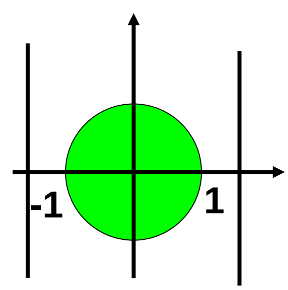


37

当 $x \notin [-1,1]$ 时

$$f_X(x) = 0$$

所以，关于X的边缘分布密度函数为



$$f_X(x) = \begin{cases} \frac{2}{\pi}\sqrt{1-x^2} & x \in [-1,1] \\ 0 & \text{其它} \end{cases}$$

38

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

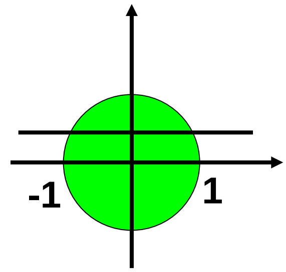
当 $y\in[-1,1]$ 时

$$f_Y(y)=\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\frac{1}{\pi}dx$$

$$=\frac{2}{\pi}\sqrt{1-y^2}$$

当 $y\notin[-1,1]$ 时

$$f_Y(y)=0$$



所以，关于Y的边缘分布密度函数为

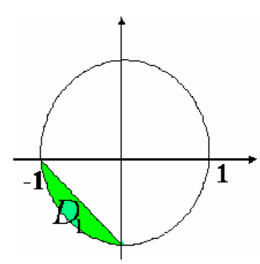
$$f_Y(y)=\begin{cases}\dfrac{2}{\pi}\sqrt{1-y^2} & y\in[-1,1]\\ 0 & \text{其它}\end{cases}$$

39

（3） $P(X+Y<-1)=\iint\limits_D f(x,y)dxdy$

$$=\iint\limits_{D_1}\frac{1}{\pi}dxdy=\int_{-1}^{0}dx\int_{-\sqrt{1-x^2}}^{-1-x}\frac{1}{\pi}dy$$

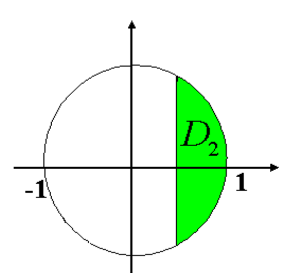
$$=\frac{1}{\pi}(\frac{\pi}{4}-\frac{1}{2})$$



$$P(X>\frac{1}{2})=\iint\limits_D f(x,y)dxdy$$

$$=\iint\limits_{D_2}\frac{1}{\pi}dxdy=\int_{\frac{1}{2}}^{1}dx\int_{-\sqrt{1-x^2}}^{+\sqrt{1-x^2}}\frac{1}{\pi}dy$$

$$=\frac{1}{\pi}(\frac{\pi}{3}-\frac{\sqrt{3}}{4})$$



40

如果二维随机变量（X，Y）服从正态分布

$$N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)$$

则两个边缘分布分别服从正态分布

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

与相关系数 ρ 无关

可见，联合分布可以确定边缘分布，但边缘分布不能确定联合分布

41

例3　设（X，Y）的联合分布密度函数为

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

求关于X，Y的边缘分布密度函数

解　关于X的分布密度函数为

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

$$=\int_{-\infty}^{+\infty}\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}(1+\sin x\sin y)dy$$

$$=\int_{-\infty}^{+\infty}\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}dy+\int_{-\infty}^{+\infty}\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}\sin x\sin y dy$$

42

$$=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\cdot\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy$$

$$+\frac{1}{2\pi}e^{-\frac{x^2}{2}}\sin x\cdot\int_{-\infty}^{+\infty}e^{-\frac{y^2}{2}}\sin y dy$$

$$=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

所以， $X\sim N(0,1)$

不同的联合分布，可有相同的边缘分布。

同理可得 $Y\sim N(0,1)$

可见, 边缘分布不能确定联合分布。

43

作业

P91 习题3
12

44

第7节

随机变量的独立性

45

随机变量的相互独立性

■ 定义　设（X，Y）的联合分布函数为F(x, y)，两个边缘分布函数分别为$F_X(x)$, $F_Y(y)$，如果对于任意的x, y都有$F(x,y)= F_X(x)\ F_Y(y)$，则称随机变量X，Y相互独立。

■ 特别，对于离散型和连续型的随机变量，该定义分别等价于

★ $$\boxed{p_{ij} = p_{i\bullet} \times p_{\bullet j}}$$ 对任意i,j

★ $$\boxed{f(x,y) = f_X(x) \cdot f_Y(y)}$$ 在$f(x,y)$连续处

46

■ 实际意义

在实际问题或应用中，当X的取值与Y的取值互不影响时，我们就认为X与Y是相互独立的，进而把上述定义式当公式运用.

■ 补充说明

在X与Y是相互独立的前提下,

♥边缘分布可确定联合分布!

$$F(x, y) = F_X(x) \cdot F_Y(y)$$

47

例1　设（X，Y）的概率分布（律）为

X \ Y	-1	0	2	$P_{i\bullet}$
1/2	2/20	1/20	2/20	1/4
1	2/20	1/20	2/20	1/4
2	4/20	2/20	4/20	2/4
$P_{\bullet j}$	2/5	1/5	2/5	

证明：X、Y相互独立。

逐个验证等式 $p_{ij} = p_{i\bullet} \times p_{\bullet j}$

48

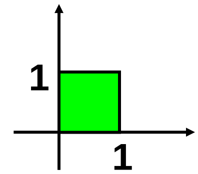
例2 设（X，Y)的概率密度为

$$\varphi(x,y)=\begin{cases}6e^{-(2x+3y)} & x\geq 0\ ,y\geq 0\\ 0 & 其他\end{cases}$$

求（1）P{0≤X≤1，0≤Y≤1}

（2）(X,Y)的边缘密度，

（3）判断X、Y是否独立。



解 ① 设A={（x，y）：0≤x≤1 ，0≤y≤1）}

$$P\{0\leq x\leq 1,0\leq y\leq 1\}=\iint\limits_{(x,y)\in A}\varphi(x,y)dxdy$$

$$=\int_0^1 dx\int_0^1 6e^{-2x-3y}dy=(1-e^{-2})(1-e^{-3})$$

49

② 边缘密度函数分别为

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

当 $x\geq 0$ 时 $\varphi_X(x)=\int_0^{+\infty}6e^{-2x-3y}dy=2e^{-2x}$

当 $x<0$ 时 $\varphi_X(x)=0$

所以， $\varphi_X(x)=\begin{cases}2e^{-2x}, & (x\geq 0)\\ 0, & (x<0)\end{cases}$

同理可得 $\varphi_Y(y)=\begin{cases}3e^{-3y}, & (y\geq 0)\\ 0, & (y<0)\end{cases}$

50

③

$$\varphi_X(x)=\begin{cases}2e^{-2x}, & x\ge 0\\ 0\ , & x<0\end{cases},\qquad \varphi_Y(y)=\begin{cases}3e^{-3y}\ , & y\ge 0\\ 0\ , & y<0\end{cases}$$

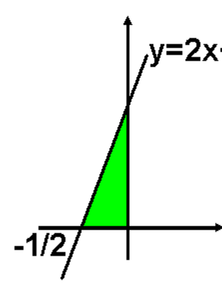
$$\varphi_X(x)\cdot\varphi_Y(y)=\begin{cases}6e^{-(2x+3y)}, & (x\ge 0, y\ge 0)\\ 0, & \text{其它}\end{cases}$$

$$=\varphi(x\ ,y)$$

所以 X 与 Y 相互独立。

51

例3 已知二维随机变量（X，Y）服从区域D上的均匀分布，D为x轴，y轴及直线y=2x+1所围成的三角形区域。判断X，Y是否独立。



解 （X,Y）的密度函数为

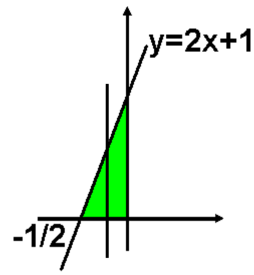
$$f(x,y)=\begin{cases}4, & (-\frac{1}{2}<x<0, 0<y<2x+1)\\ 0, & \text{其他}\end{cases}$$

52

关于**X**的边缘分布密度为 $f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy$

当 $x\le -\frac{1}{2}$ 或 $x\ge 0$ 时 $f_X(x)=0$

当 $-\frac{1}{2}<x<0$ 时，



$$f_X(x)=\int_0^{2x+1} 4dy \;=4(2x+1)$$

所以，关于**X**的边缘分布密度为

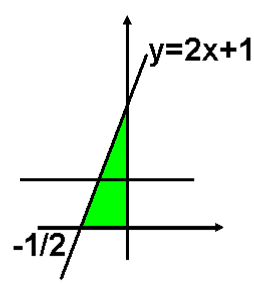
$$f_X(x)=\begin{cases}4(2x+1), & (-\frac{1}{2}<x<0)\\ 0, & \text{其它}\end{cases}$$

53

关于**Y**的边缘分布密度为 $f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx$

当 $y\le 0$ 或 $y\ge 1$ 时 $f_Y(y)=0$

当 $0<y<1$ 时，



$$f_Y(y)=\int_{\frac{y-1}{2}}^{0} 4dx \;=2(1-y)$$

所以，关于**Y**的边缘分布密度为

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

54

所以

$$f_X(x)\cdot f_Y(y)=\begin{cases}8(2x+1)(1-y),(-\frac{1}{2}<x<0,0<y<1\ ) \\ 0, \qquad\qquad \text{其它}\end{cases}$$

$$\neq f(x,y)$$

所以，**X与Y不独立**。

55

作业

P91 习题3
5， 19， 20

56

本章作业

4，5，7，8，12，19，20

下节课交作业，请同学们做好作业，下次课不要忘了带过来

57