第三章 多维随机变量

3.1多维随机变量

一. 二维随机变量

二. 二维随机变量的分布函数

三. 二维随机变量及其分布律

为什么要研究多维随机变量?

在实际问题中, 试验结果有时需要同时用两个或两个以上的 *r.v.*来描述.

例如 用温度和风力来描述天气情况.身高体重

例如通过对含碳、硫、磷量的测定来研究钢的成分. 要研究这些 *r.v.*之间的联系, 就需考虑多维 *r.v.*及其取值规律—多维分布.

1. 定义 设 $X_1 = X_1(\omega), X_2 = X_2(\omega), \cdots, X_n = X_n(\omega)$ 是定义在 (Ω, F, P) 上的n个随机变量，由它们构成的一个 n 维向量$(X_1(\omega), X_2(\omega), \cdots, X_n(\omega))$ 叫做 n 维随机向量或n 维随机变量，简记为$(X_1, X_2, \cdots, X_n)$.

其中第i个随机变量X_i 称为第i个分量。$(i = 1,2,\cdots n)$

n 维随机变量$(X_1, X_2, \cdots, X_n)$ 取值范围是R^n或其子集。

请注意与一维情形的对照 .

一维随机变量X——R^1上的随机点坐标

二维随机变量(X,Y)——R^2上的随机点坐标

n维随机变量$(X_1,X_2,\ldots,X_n)$————R^n上的随机点坐标

多维随机变量的研究方法也与一维类似，

用分布函数、概率密度或分布律来描述其统计规律

2. 分布函数

n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 的联合分布函数

$$F(x_1, x_2, \cdots, x_n) = P\{X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n\},$$

其中 $x_1, x_2, \cdots, x_n$ 为任意实数.

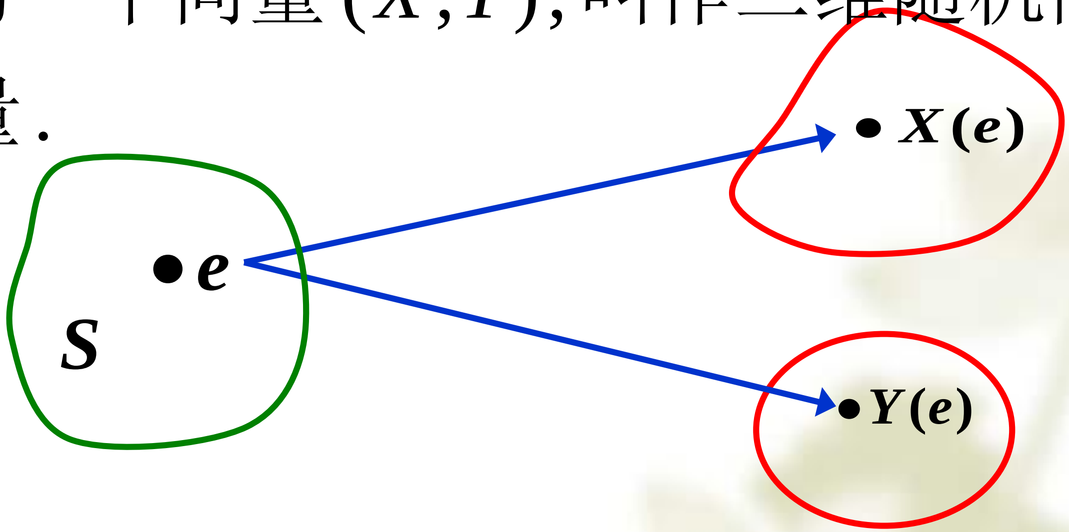
其中 $\{X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n\} = \bigcap_{i=1}^{n} \{ X_i \le x_i \}$

以下重点讨论二维随机变量,n维随机变量情形可平行推广。

一. 二维随机变量

定义 设 E 是一个随机试验，它的样本空间是 $\Omega=\{\omega\}$，设 $X=X(\omega)$ 和 $Y=Y(\omega)$ 是定义在 Ω 上的随机变量，由它们构成的一个向量 (X,Y)，叫作二维随机向量或二维随机变量.

图示



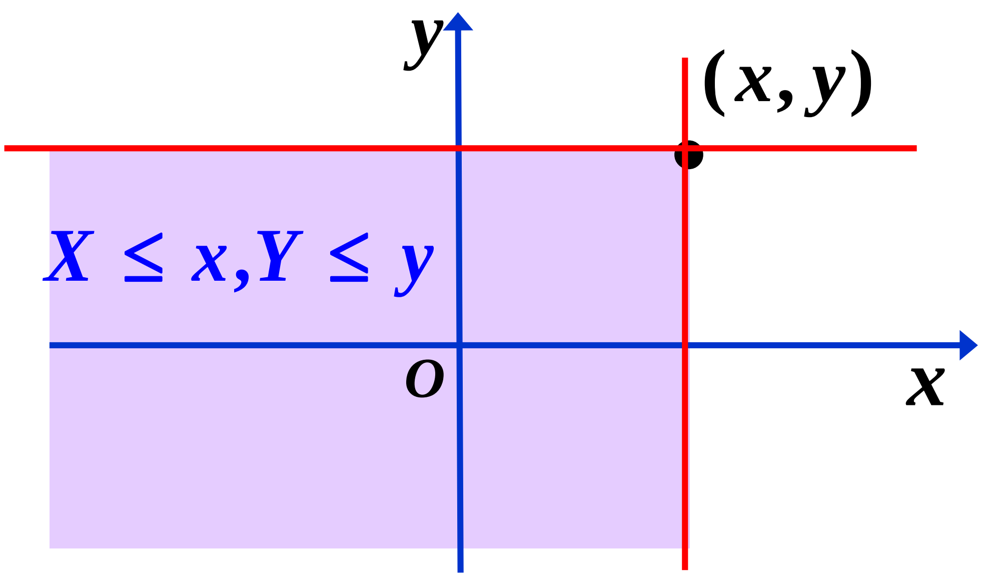
二. 二维随机变量的分布函数

1. 定义 设 (X,Y) 是二维随机变量,对于任意实数 x, y,

二元函数:

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

称为二维随机变量 (X,Y) 的分布函数,或称为随机变量 X 和 Y 的联合分布函数.



几何意义:

分布函数F(x, y)表示随机点(X, Y)落在阴影区域中的概率

2. 分布函数的性质

1° $F(x,y)$ 是变量 x 和 y 的不减函数,即对于任意固定的 y,当 $x_2 > x_1$ 时 $F(x_2,y) \geq F(x_1,y)$,

对于任意固定的x,当$y_2 > y_1$时$F(x,y_2) \geq F(x,y_1)$.

2° $0 \leq F(x,y) \leq 1$, 且有

对于任意固定的y, $F(-\infty,y) = \lim\limits_{x\to-\infty} F(x,y) = 0,$

对于任意固定的x, $F(x,-\infty) = \lim\limits_{y\to-\infty} F(x,y) = 0,$

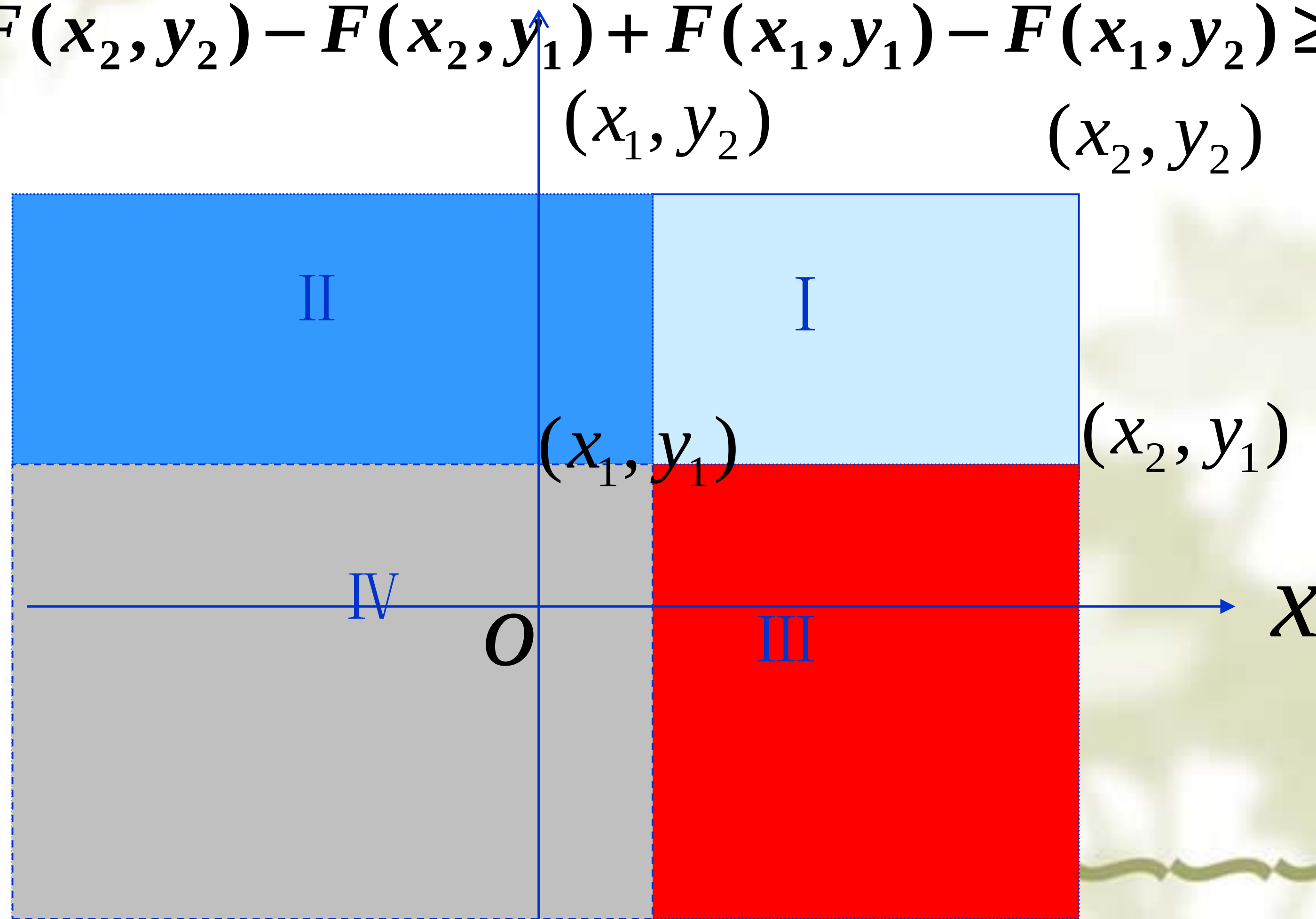
$$F(-\infty,-\infty) = \lim_{\substack{x\to-\infty \\ y\to-\infty}} F(x,y) = 0,\ F(+\infty,+\infty) = \lim_{\substack{x\to+\infty \\ y\to+\infty}} F(x,y) = 1.$$

$3^{\circ}\quad F(x,y)=F(x+0,y),\ F(x,y)=F(x,y+0),$

即 $F(x,y)$ 关于 x 右连续，关于 y 也右连续.

4° 对于任意 $(x_1,y_1),(x_2,y_2)$, $x_1<x_2$, $y_1<y_2$,

有 $F(x_2,y_2)-F(x_2,y_1)+F(x_1,y_1)-F(x_1,y_2)\geq 0.$



证明 $P\{x_1 < X \le x_2, y_1 < Y \le y_2\}$

$$= P\{X \le x_2, y_1 < Y \le y_2\} - P\{X \le x_1, y_1 < Y \le y_2\}$$

$$= P\{X \le x_2, Y \le y_2\} - P\{X \le x_2, Y \le y_1\}$$

$$- P\{X \le x_1, Y \le y_2\} + P\{X \le x_1, Y \le y_1\} \ge 0,$$

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

反之，任一满足上述四个性质的二元函数F(x, y)都可以作为某个二维随机变量(X, Y)的分布函数。

例 1．已知二维随机变量(X, Y)的分布函数为

$$F(x,y)=A[B+\arctan(\frac{x}{2})][C+\arctan(\frac{y}{3})]$$

1)求常数A，B，C。　2)求$P\{0<X\le 2,0<Y\le 3\}$

解：(1) $\because F(+\infty,+\infty)=A[B+\frac{\pi}{2}][C+\frac{\pi}{2}]=1$

$$F(-\infty,y)=A[B-\frac{\pi}{2}][C+arctg(\frac{y}{3})]=0$$

$$F(x,-\infty)=A[B+arctg(\frac{x}{2})][C-\frac{\pi}{2}]=0$$

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

$$(2)P\{0<X\le 2,0<Y\le 3\}=F(2,3)-F(0,3)-F(2,0)+F(0,0)=\frac{1}{16}$$

三. 二维随机变量及其分布律

1.二维离散型随机变量定义 若二维随机变量 (X, Y) 所取的可能值是有限对或无限可列多对, 则称 (X, Y) 为二维离散型随机变量.

2. 二维离散型随机变量的分布律

设二维离散型随机变量(X,Y)所有可能取的值为(x_i, y_j), $i, j = 1, 2, \cdots$, 记

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

称此为二维离散型随机变量 (X,Y) 的分布律, 或随机变量 X 和 Y 的联合分布律.

其中 $p_{ij} \geq 0,\ \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} p_{ij} = 1.$

二维随机变量（X,Y）的分布律也可表示为

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

例2 设随机变量 X 在1, 2, 3, 4四个整数中等可能地取值, 另一个随机变量 Y 在 $1\sim X$ 中等可能地取一整数值. 求 (X,Y) 分布律, 及 $P(X=2), P(X\le Y)$的值.

解

$\{X=i, Y=j\}$ 的取值情况是: $i=1,2,3,4,$ j 取不大于 i 的正整数.

且由乘法公式得

$$P\{X=i,Y=j\}$$

$$=P\{Y=j\mid X=i\}P\{X=i\}=\frac{1}{i}\cdot\frac{1}{4},$$

$$i=1,2,3,4,\quad j\le i.$$

于是 (X,Y) 的分布律为

$Y \backslash X$	1	2	3	4
1	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{16}$
2	0	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{16}$
3	0	0	$\frac{1}{12}$	$\frac{1}{16}$
4	0	0	0	$\frac{1}{16}$

说明

离散型随机变量 (X,Y) 的分布函数归纳为

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

其中和式是对一切满足 $x_i \le x, y_j \le y$ 的 i, j 求和.

3.二维连续型随机变量:

1. 定义

对于二维随机变量(X,Y)的分布函数$F(x,y)$,如果存在非负的函数$f(x,y)$使对于任意x,y有

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

则称(X,Y)是连续型的二维随机变量,函数$f(x,y)$称为二维随机变量(X,Y)的概率密度,或称为随机变量X和Y的联合概率密度.

2. 性质

(1) $f(x,y) \geq 0.$

反之，具有这两个性质的二元函数f (x, y)，必是某个二维连续型随机变量的密度函数。

(2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = F(\infty,\infty) = 1.$

(3) 设 G 是 xOy 平面上的一个区域，点 (X,Y) 落在 G 内的概率为

$$P\{(X,Y)\in G\} = \iint_G f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

(4)若 $f(x,y)$ 在 (x,y) 连续，则有 $\dfrac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y).$

3. 说明

几何上, $z = f(x, y)$ 表示空间的一个曲面.

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = 1,$$

表示介于 $f(x, y)$和 xOy 平面之间的空间区域的全部体积等于1.

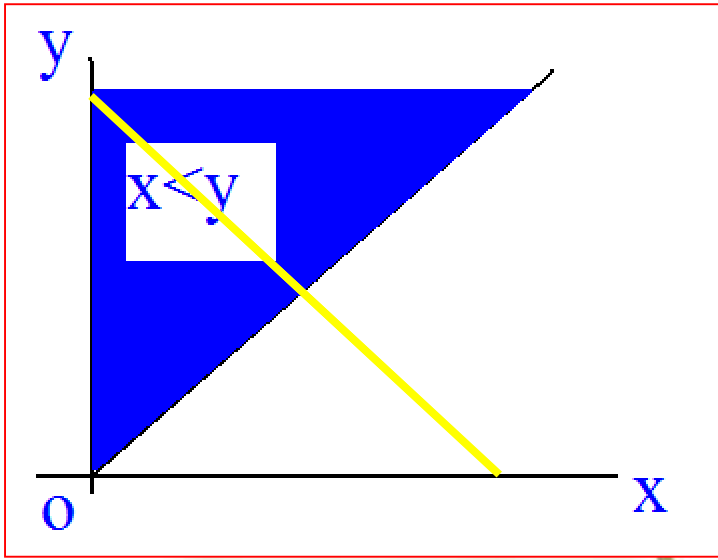
$$P\{(X, Y) \in G\} = \iint_G f(x, y)\,\mathrm{d}x\,\mathrm{d}y,$$

$P\{(X, Y) \in G\}$的值等于以G为底, 以曲面$z = f(x, y)$为顶面的柱体体积.

例 5: 设二维随机变量(X,Y)的概率密度

$$f(x,y)=\begin{cases} kx & 0\le x\le y\le 1 \\ 0 & \text{其他} \end{cases}.$$

求:(1)常数 k;　(2) $P(X+Y\le 1)$.



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

得 $1=\int_0^1 dx\int_x^1 kxdy=\frac{k}{6}$

所以 k=6

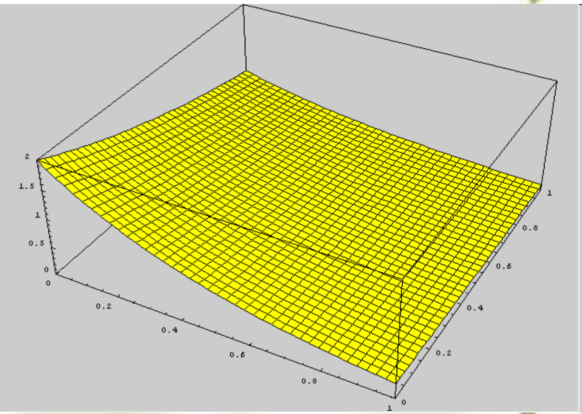
(2) $P(X+Y\le 1)=\iint\limits_{x+y\le 1} 6xdxdy=6\int_0^{\frac{1}{2}} xdx\int_x^{1-x} dy=\frac{1}{4}$

例7 设二维随机变量(X,Y)具有概率密度

$$f(x,y)=\begin{cases}2\mathrm{e}^{-(2x+y)}, & x>0,\ y>0,\\ 0, & \text{其他}.\end{cases}$$

(1) 求分布函数$F(x,y)$; (2) 求概率$P\{Y\le X\}$.

解 (1) $F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)\,\mathrm{d}u\,\mathrm{d}v$



$$=\begin{cases}\int_0^y\int_0^x 2\mathrm{e}^{-(2u+v)}\,\mathrm{d}u\,\mathrm{d}v, & x>0,\ y>0,\\ 0, & \text{其他}.\end{cases}$$

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

(2) 将 (X,Y)看作是平面上随机点的坐标,

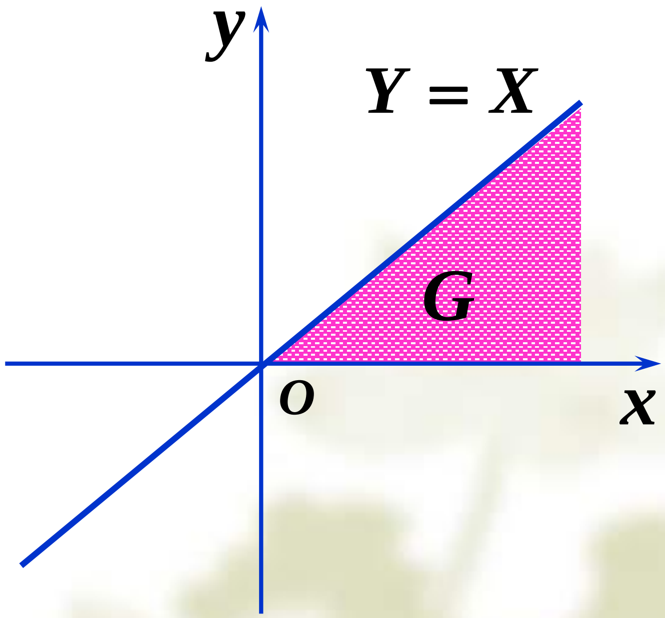
即有 $\{Y \le X\} = \{(X,Y) \in G\},$

$$P\{Y \le X\} = P\{(X,Y) \in G\}$$

$$= \iint_G f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

$$= \int_0^{+\infty}\int_y^{+\infty} 2\mathrm{e}^{-(2x+y)}\,\mathrm{d}x\,\mathrm{d}y$$

$$= \frac{1}{3}.$$

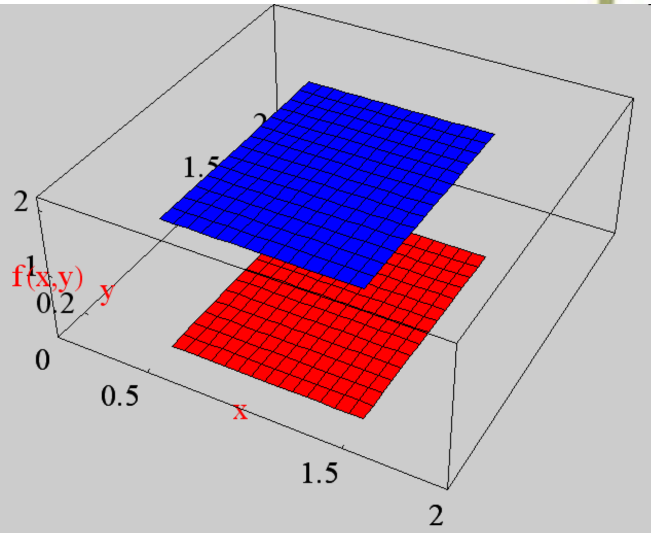


4．二维均匀分布和二维正态分布

1. 二维均匀分布

定义 设 D 是平面上的有界区域, 其面积为 S, 若二维随机变量（X, Y）具有概率密度

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

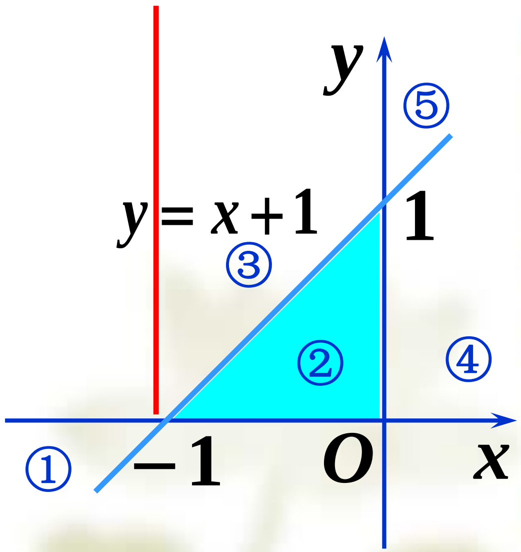


则称（X, Y）在 D 上服从均匀分布.

易见，若（X，Y）在区域D上(内）服从均匀分布，对D内任意区域G，有 $P\{(X,Y\}\in G\}=\dfrac{S_G}{S_D}$

例8 已知随机变量 (X,Y) 在 D 上服从均匀分布,试求 (X,Y) 的概率密度及分布函数,其中 D 为 x 轴,y 轴及直线 $y=x+1$ 所围成的三角形区域 .

解 由 $f(x,y)=\begin{cases}1/S, & (x,y)\in D,\\ 0, & \text{其他}.\end{cases}$



得 $f(x,y)=\begin{cases}2, & (x,y)\in D,\\ 0, & \text{其他}.\end{cases}$

当 $x<-1$ 或 $y<0$ 时, $f(x,y)=0$

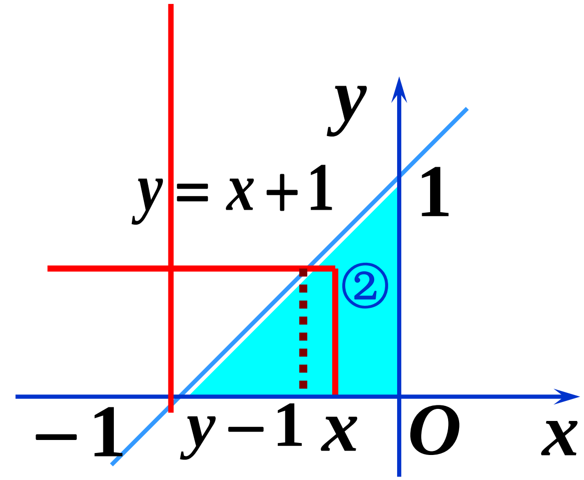
$$\Rightarrow F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,\mathrm{d}u\,\mathrm{d}v=0;$$

当$-1 \le x < 0,\ 0 \le y < x+1$时,

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

$$=\int_{-1}^{y-1}\mathrm{d}u\int_{0}^{u+1}2\mathrm{d}v+\int_{y-1}^{x}\mathrm{d}u\int_{0}^{y}2\mathrm{d}v$$

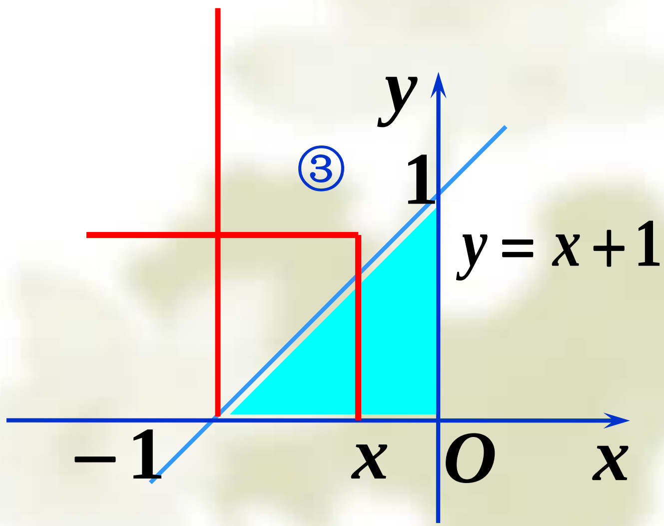
$$=(2x-y+2)y;$$



当$-1 \le x < 0,\ y \ge x+1$时,

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

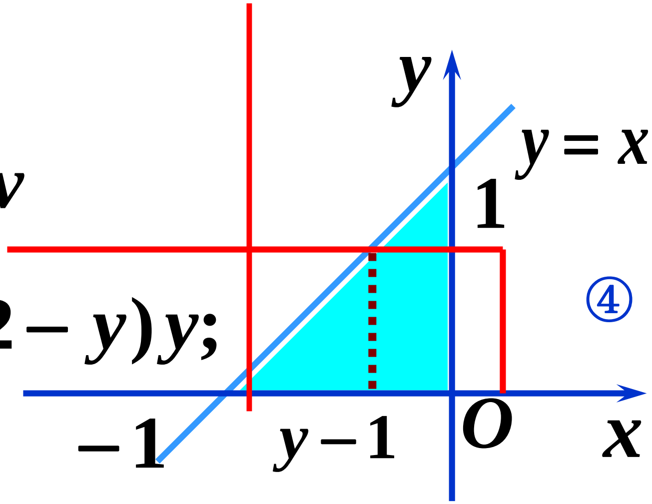
$$=\int_{-1}^{x}\mathrm{d}u\int_{0}^{u+1}2\mathrm{d}v=(x+1)^2;$$



当 $x \ge 0, 0 \le y < 1$ 时,

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

$$= \int_{-1}^{y-1}\mathrm{d}u\int_{0}^{u+1}2\,\mathrm{d}v + \int_{y-1}^{0}\mathrm{d}u\int_{0}^{y}2\,\mathrm{d}v = (2-y)y;$$



当 $x \ge 0, y \ge 1$ 时,

$$F(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)\,\mathrm{d}u\,\mathrm{d}v = \int_{-1}^{0}\mathrm{d}u\int_{0}^{u+1}2\,\mathrm{d}v = 1.$$

所以 (X, Y) 分布函数为

$$F(x,y) = \begin{cases} 0, & x < -1, 或\ y < 0, \\ (2x - y + 2)y, & -1 \le x < 0, 0 \le y < x+1, \\ (x+1)^2, & -1 \le x < 0, y \ge x+1, \\ (2-y)y, & x \ge 0, 0 \le y < 1, \\ 1, & x \ge 0, y \ge 1. \end{cases}$$

2. 二维正态分布

若二维随机变量（X,Y）具有概率密度

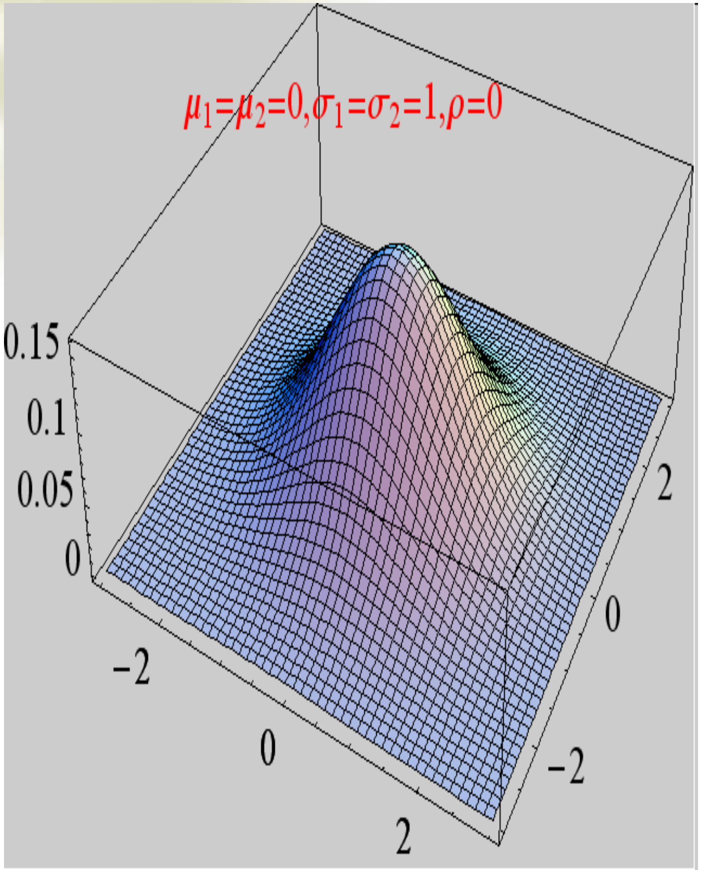
$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}$$

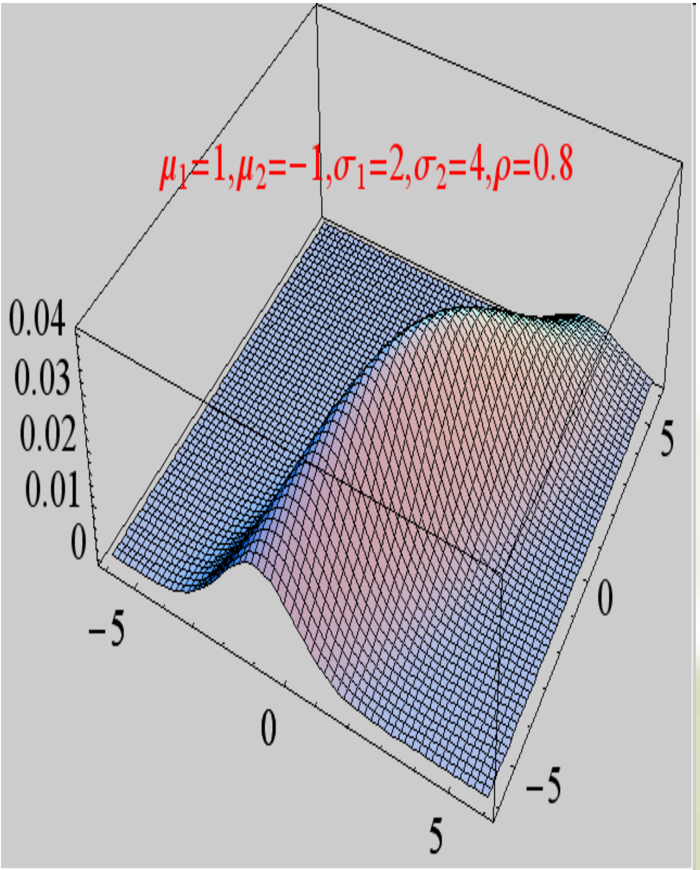
$$(-\infty<x<\infty,-\infty<y<\infty),$$

其中$\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$均为常数,且$\sigma_1>0,\sigma_2>0,-1<\rho<1.$

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

二维正态分布的图形


$\mu_1=\mu_2=0,\sigma_1=\sigma_2=1,\rho=0$
0.15
0.1
0.05
0
−2
0
2
2
0
−2


$\mu_1=1,\mu_2=-1,\sigma_1=2,\sigma_2=4,\rho=0.8$
0.04
0.03
0.02
0.01
0
−5
0
5
5
0
−5