第三章 随机向量

3.0. 简介

3.1. 二维随机向量及其分布函数

3.2 边缘分布

3.3. 条件分布

3.4. 随机变量的独立性

3.5. 两个随机变量的函数的分布

Introduction

概率与统计中很多问题都会同事涉及到两个或两个以上的随机变量

例如：随机选择一个人，X，Y 分别表示其身高、体重.

X，Y 为炮弹着地点的位置，横坐标和纵坐标.

X_1, X_2, X_3 为某一天某商店顾客使用 Visa, MasterCard, and American Express 等信用卡的数目

随机向量

定义 3.1 设 E 是一个随机试验，它的样本空间是 $\Omega=\{e\}$. 设 $X(e)$ 与 $Y(e)$ 是定义在同一个样本空间 Ω 上的两个随机变量，则称 $(X(e),Y(e))$ 为 Ω 上的二维随机向量(2-dimensional random vector)或二维随机变量(2-dimensional random variable)，简记为 (X,Y).

设 E 是一个随机试验，它的样本空间是 $\Omega=\{e\}$. 设随机变量 $X_1(e),X_2(e),\cdots,X_n(e)$ 是定义在同一个样本空间 Ω 上的 n 个随机变量，则称向量 $(X_1(e),X_2(e),\cdots,X_n(e))$ 为 Ω 上的 n 维随机变量或 n 维随机向量，简记为 $(X_1,X_2,\cdots,X_n)$.

定义 3.2 设(X,Y)是二维随机向量，对任意实数x和y，称二元函数

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\} \tag{3-1}$$

为二维随机向量(X,Y)的分布函数，或称为随机变量X和Y的联合分布函数.

类似地，可定义n维随机变量$(X_1,X_2,\cdots,X_n)$的分布函数.

设$(X_1,X_2,\cdots,X_n)$是n维随机变量，对任意实数$x_1,x_2,\cdots,x_n$，称n元函数

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

为n维随机变量$(X_1,X_2,\cdots,X_n)$的联合分布函数.

我们容易给出分布函数的几何解释. 如果把二维随机变量(X,Y)看成是平面上随机点的坐标,那么,分布函数$F(x,y)$在(x,y)处的函数值就是随机点(X,Y)落在直线$X=x$的左侧和直线$Y=y$的下方的无穷矩形域内的概率(见图3-1).

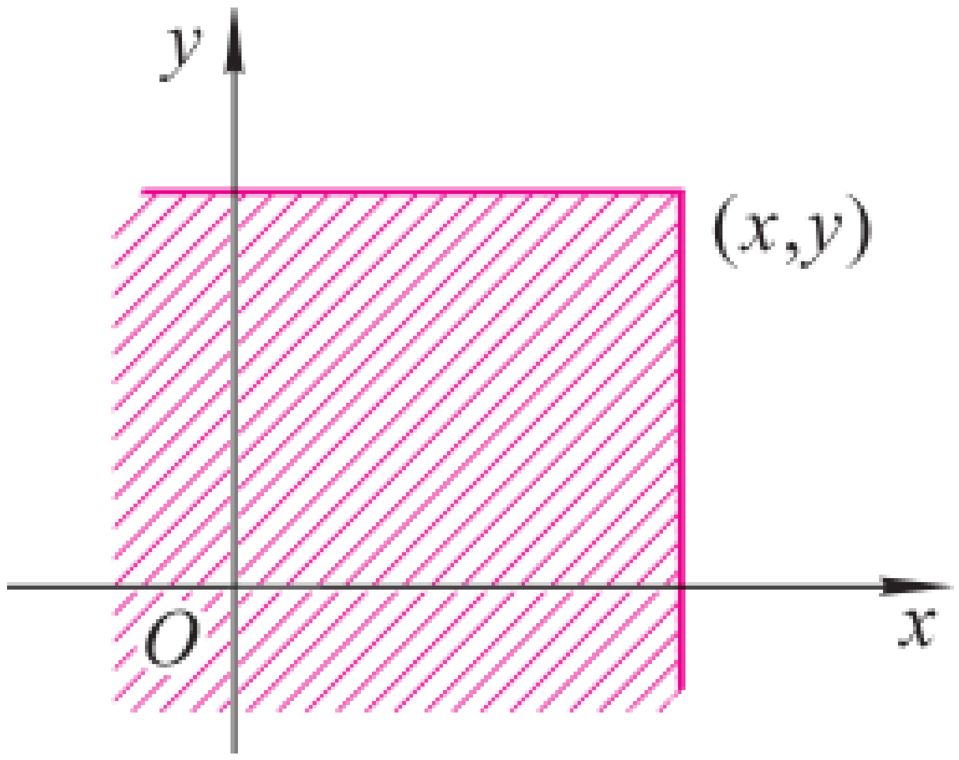


图 3-1

根据以上几何解释并借助于图 3-2，可以算出随机点(X,Y)落在矩形域$\{x_1 < X \leqslant x_2, y_1 < Y \leqslant y_2\}$内的概率为

$$
\begin{aligned}
&P\{x_1 < X \leqslant x_2, y_1 < Y \leqslant y_2\} \\
&\quad = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1).
\end{aligned}
\tag{3-2}
$$

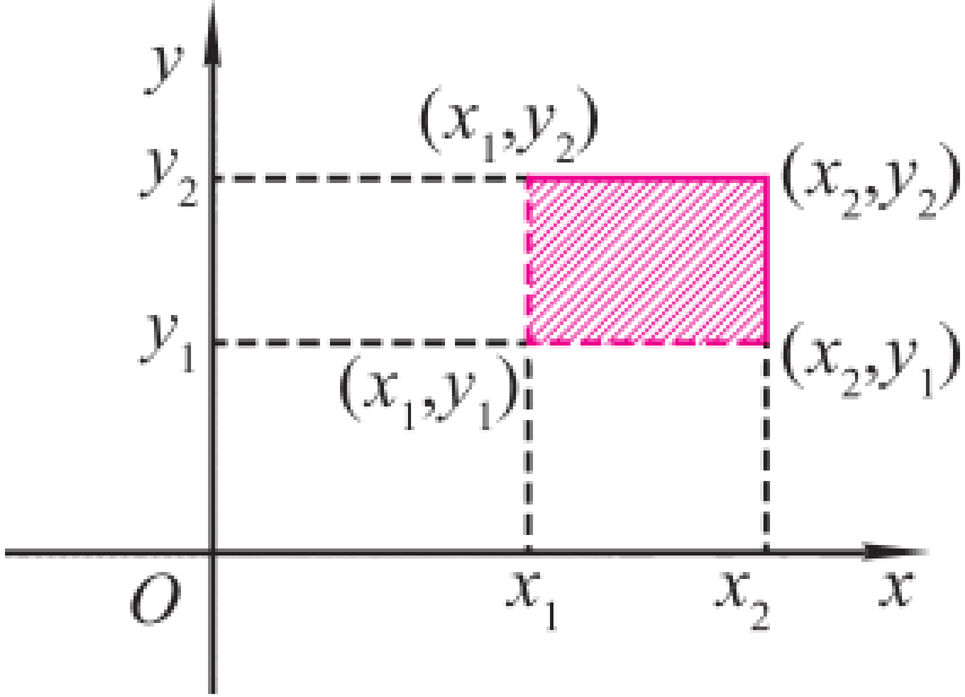
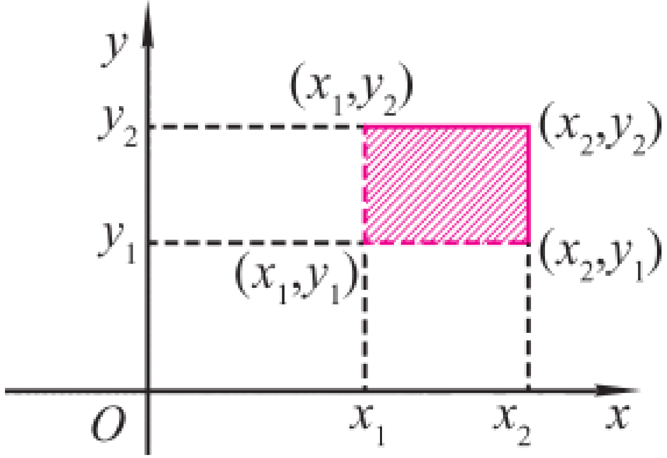


图 3-2

分布函数 $F(x,y)$ 具有以下基本性质:

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,$F(x_2,y)\geqslant F(x_1,y)$;对于任意固定的 x,当 $y_2>y_1$ 时,$F(x,y_2)\geqslant F(x,y_1)$.

2° $0\leqslant F(x,y)\leqslant 1$,且对于任意固定的 y,$F(-\infty,y)=0$,对于任意固定的 x,$F(x,-\infty)=0$,$F(-\infty,-\infty)=0$,$F(+\infty,+\infty)=1$.



3° $F(x,y)$ 关于 x 和 y 是右连续的,即

$$F(x,y)=F(x+0,y),\quad F(x,y)=F(x,y+0).$$

4° 对于任意 (x_1,y_1),(x_2,y_2),$x_1<x_2$,$y_1<y_2$,下述不等式成立:

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

二维离散型随机变量

一维离散型随机变量的分布律，给出了1概率是如何分配给每一个变量值的.

X	x_1	x_2	x_3	$\cdots$	x_k	$\cdots$
p_k	p_1	p_2	p_3	$\cdots$	p_k	$\cdots$

1. $p_k \geq 0$, 对所有的 k

2. $\sum_{k=1}^{\infty} p_k = 1$

如何将整体的 1概率 分配给二维离散型随机变量（X,Y）？

联合分布律，或联合概率分布律，joint probability mass function

两个随机变量 X 和 Y 的联合概率分布律描述了每一对 (x,y) 的概率为多少

$Y \backslash X$	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}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	

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

1. 非负性 $p_{ij} \geq 0,\ i,j = 1,2,\ldots$

2. 规范性 $\sum_{i,j} p_{ij} = 1$

联合分布函数 (joint cumulative distribution function)

$Y \backslash X$	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}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	

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

对一切满足 $x_i \leqslant x, y_j \leqslant y$ 的 i, j 来求和的.

例: 若随机向量 (X,Y) 有如下联合分布律，计算如下概率

$Y \backslash X$	1	2	3
1	0.1	0.3	0
2	0	0	0.2
3	0.1	0.1	0
4	0	0.2	0

$P(X > 1, Y \geq 3) = 0.3$

$P(X = 1) = 0.2$

$P(X = 2) = 0.6$

$P(X = 3) = 0.2$

$P(Y = 1) = 0.4$

$P(Y = 2) = 0.2$

$P(Y = 3) = 0.2$

$P(Y = 4) = 0.2$

X，Y 的边缘概率分布律 marginal distribution function

$Y \backslash X$	1	2	3	$P(Y = y_j)$
1	0.1	0.3	0	0.4
2	0	0	0.2	0.2
3	0.1	0.1	0	0.2
4	0	0.2	0	0.2
$P(X = x_i)$	0.2	0.6	0.2	

定义: 给定随机变量 X 和 Y 的联合分布律，则其边缘概率分布律为

$Y \backslash X$	x_1	x_2	$\cdots$	x_i	$\cdots$	$P(Y=y_j)$
y_1	p_{11}	p_{21}	$\cdots$	p_{i1}	$\cdots$	$\sum_i p_{i1}$
y_2	p_{12}	p_{22}	$\cdots$	p_{i2}	$\cdots$	$\sum_i p_{i2}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
y_j	p_{1j}	p_{2j}	$\cdots$	p_{ij}	$\cdots$	$\sum_i p_{ij}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		
$P(X=x_i)$	$\sum_j p_{1j}$	$\sum_j p_{2j}$	$\ldots$	$\sum_j p_{ij}$		$\ldots$

例: 假设随机变量 X 等可能随机在 1, 2, 3, 4, 中取值, 随机变量 Y 等可能 在 $1\sim X$ 中取值. 求 (X,Y) 的联合概率分布

解

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

$$=\frac{1}{i}\frac{1}{4},i=1,2,3,4;j\le i$$

$Y \backslash X$	1	2	3	4
1	1/4	1/8	1/12	1/16
2	0	1/8	1/12	1/16
3	0	0	1/12	1/16
4	0	0	0	1/16

例: 设 X，Y 分别服从以下概率分布律:

X	0	1
p_k	$\frac{1}{3}$	$\frac{2}{3}$

Y	-1	0	1
p_k	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

且 $P(X^2 = Y^2) = 1.$

1. 求 (X,Y) 的联合分布律

2. 求 $Z = XY$ 的分布律

解: 由 $P(X^2 = Y^2) = 1$ 得，$P(X^2 \neq Y^2) = 0$ 即

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

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

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

再由 X，Y 的分布律，得

$X \backslash Y$	-1	0	1	$P(X = x_i)$
0	0	1/3	0	1/3
1	1/3	0	1/3	2/3
$P(Y = y_j)$	1/3	1/3	1/3	

例: 假设一个盒子有 4 个白球和 5 个红球. 现从中随机抽取两次，每次取一个并放回. 定义随机变量 X 和 Y 如下

$$X=\begin{cases}0, & \text{第一次是白球,}\\ 1, & \text{第一次是红球,}\end{cases}\qquad Y=\begin{cases}0, & \text{第二次是白球,}\\ 1, & \text{第二次是红球,}\end{cases}$$

求 X 和 Y 的联合概率分布律

解

$X \backslash Y$	0	1	$P(X=x_i)$
0	$\frac{4}{9}\times\frac{4}{9}$	$\frac{4}{9}\times\frac{5}{9}$	$\frac{4}{9}$
1	$\frac{5}{9}\times\frac{4}{9}$	$\frac{5}{9}\times\frac{5}{9}$	$\frac{5}{9}$
$P(Y=y_j)$	$\frac{4}{9}$	$\frac{5}{9}$	

定义: 独立随机变量（离散型）

两个离散型随机变量 X 和 Y 是独立的，如果对每一对 x_i 和 y_j 有

$$P\left(X = x_i, Y = y_j\right) = P(X = x_i) \cdot P(Y = y_j)$$

注1: 事件 A 与 B 相互独立，如果 $P(AB) = P(A) \cdot P(B)$

注2：事件 $\left(X = x_i, Y = y_j\right)$ 与 事件 AB 表达式是否相同?

以上都是指事件的交，或积，或同时发生

例: 假设随机变量 X 和 Y 是独立的，其联合概率分布和边缘概率分布如下所示，试完成以下空白处概率.

$X \backslash Y$	y_1	y_2	y_3	$P(X=x_i)$
x_1	1/24	1/8	1/12	1/4
x_2	1/8	3/8	1/4	3/4
$P(Y=y_j)$	1/6	1/2	1/3	

由 $P(Y=y_1)=P(X=x_1,Y=y_1)+P(X=x_2,Y=y_1)$

$\Rightarrow P(X=x_1,Y=y_1)=P(Y=y_1)-P(X=x_2,Y=y_1)=\dfrac{1}{6}-\dfrac{1}{8}=\dfrac{1}{24}$

由 $P(X=x_1,Y=y_1)=P(X=x_1)\times P(Y=y_1)$ $\quad\Rightarrow P(X=x_1)=\dfrac{1}{4}$

条件概率 $P(B|A)=\dfrac{P(AB)}{P(A)}$, 如果 $P(A)>0$.

定义: 条件分布，设 (X,Y) 是二维离散型随机变量，对某整数 j, 如果$P\left(Y=y_j\right)>0$, 那么

$$P\left(X=x_i\middle|Y=y_j\right)=\frac{P\left(X=x_i,Y=y_j\right)}{P\left(Y=y_j\right)},\ i=1,2,\ldots$$

称为在 $Y=y_j$ 的条件下，随机变量 X 的概率分布.

类似地，对某一整数 i, 如果 $P(X=x_i)>0$, 那么

$$P\left(Y=y_j\middle|X=x_i\right)=\frac{P\left(X=x_i,Y=y_j\right)}{P(X=x_i)},\ j=1,2,\ldots$$

称为在 $X=x_i$ 的条件下，随机变量 Y 的概率分布.

例: 随机向量 (X,Y) 的联合概率分布如下，求：（见教材P70）

$Y \backslash X$	1	2	3	4	$P(Y=y_j)$
1	1/4	1/8	1/12	1/16	25/48
2	0	1/8	1/12	1/16	13/48
3	0	0	1/12	2/16	10/48
$P(X=x_i)$	1/4	1/4	1/4	1/4	

(1) X 在 $Y=1$ 的条件下的概率分布.

X	1	2	3	4
p_k	$\frac{12}{25}$	$\frac{6}{25}$	$\frac{4}{25}$	$\frac{3}{25}$

(2) Y 在 $X=2$ 的条件下的概率分布

Y	1	2	3
p_k	$\frac{1}{2}$	$\frac{1}{2}$	0

例 3.10 一射手进行射击，击中的概率为 p，$0 < p < 1$，射击到击中目标两次为止. 记 X 表示首次击中目标时的射击次数，Y 表示射击的总次数. 试求 X，Y 的联合分布律与条件分布律.

解 依题意，$\{X = m, Y = n\}$ 表示"前 $m-1$ 次不中，第 m 次击中，接着又 $n-1-m$ 次不中，第 n 次击中". 因各次射击是独立的，故 X，Y 的联合分布律为

$$P\{X = m, Y = n\} = p^2(1-p)^{n-2}, \quad m = 1,2,\cdots,n-1; n = 2,3,\cdots.$$

又因

$$P\{X = m\} = \sum_{n=m+1}^{\infty} P\{X = m, Y = n\} = \sum_{n=m+1}^{\infty} p^2(1-p)^{n-2}$$

$$= p^2 \sum_{n=m+1}^{\infty} (1-p)^{n-2} = p(1-p)^{m-1}, \quad m = 1,2,\cdots;$$

$$P\{Y = n\} = (n-1)p^2(1-p)^{n-2}, \quad n = 2,3,\cdots,$$

因此，所求的条件分布律分别为

当 $n = 2,3,\cdots$ 时，

$$P\{X = m \mid Y = n\} = \frac{P\{X = m, Y = n\}}{P\{Y = n\}} = \frac{1}{n-1}, \quad m = 1,2,\cdots,n-1;$$

当 $m = 1,2,\cdots$ 时，

$$P\{Y = n \mid X = m\} = \frac{P\{X = m, Y = n\}}{P\{X = m\}} = p(1-p)^{n-m-1}, \quad n = m+1, m+2,\cdots.$$

二维连续型随机向量

令 X 和 Y 为连续型随机变量. 其联合概率密度 $f(x,y)$ 需满足

(1) $f(x,y) \geq 0$, 且 (2) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy = 1$

(X,Y) 联合分布函数

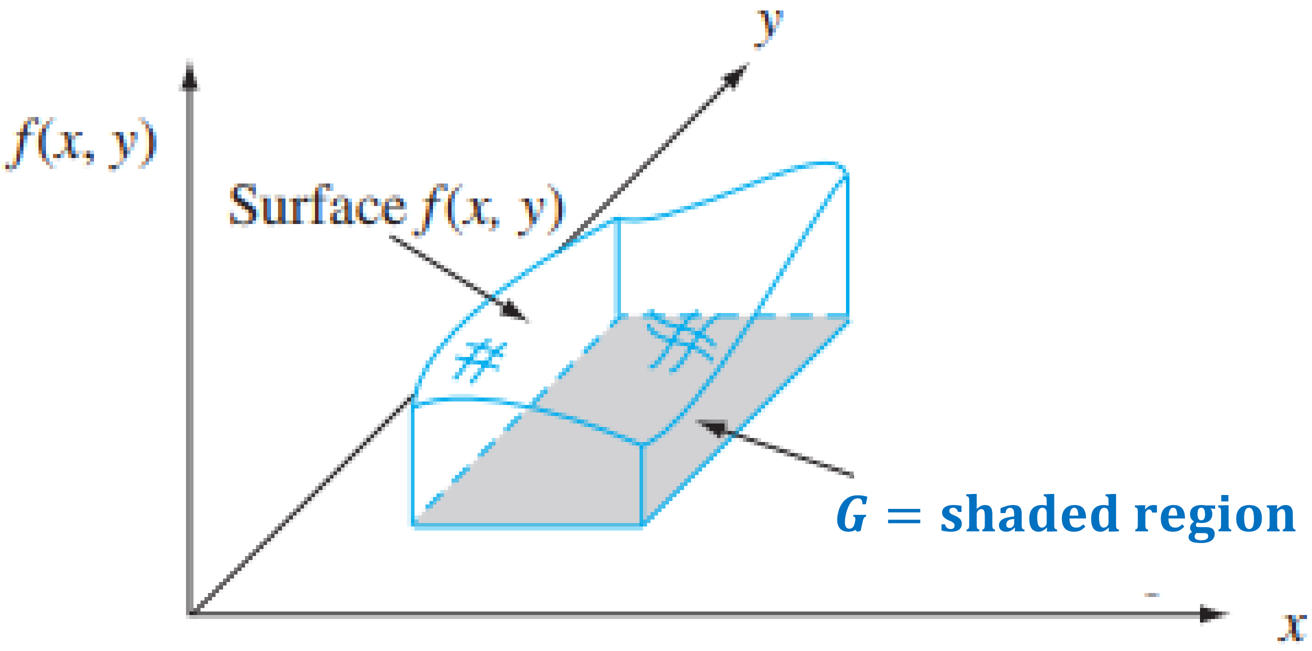
$$F(x,y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,dvdu.$$

(3) 对任意二维区域 G,

$$P\big((X,Y) \in G\big) = \iint_G f(x,y)dxdy.$$

(4) 如果 $f(x,y)$ 在点 (x,y) 连续, 则有 $\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y)$.

在几何上，概率密度 $z=f(x,y)$ 表示空间一曲面，介于它和 XOY 面的空间区域的立体体积等于1，$P((X,Y)\in G)$ 的值等于以G为底，以曲面 $z=f(x,y)$ 为顶的曲顶柱体的体积



例: 验证函数 $f(x,y)=\begin{cases}\frac{6}{5}(x+y^2),\ 0\le x\le 1,\ 0\le y\le 1\\ 0, \qquad\qquad\qquad 其他\end{cases}$ 是二维连续型随机变量 (X,Y) 的联合概率密度函数，并计算概率$P\left(0\le X\le\frac{1}{4},0\le Y\le\frac{1}{4}\right)$.

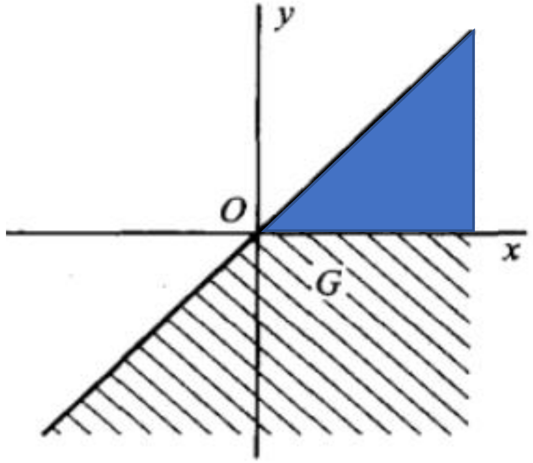
解 1. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)\,dxdy=\int_0^1\int_0^1\frac{6}{5}(x+y^2)\,dxdy=\int_0^1\left(\int_0^1\frac{6}{5}(x+y^2)\,dx\right)dy$

$$=\int_0^1\frac{6}{5}\left(\frac{x^2}{2}+y^2x\right)\Big|_{x=0}^{x=1}dy=\frac{6}{5}\int_0^1\left(\frac{1}{2}+y^2\right)dy=\frac{6}{5}\left(\frac{1}{2}y+\frac{y^3}{3}\right)_{y=0}^{y=1}=1$$

2. $\int_0^{1/4}\int_0^{1/4}f(x,y)\,dxdy=\int_0^{1/4}\left(\int_0^{\frac{1}{4}}\frac{6}{5}(x+y^2)\,dx\right)dy$

$$=\int_0^{\frac{1}{4}}\frac{6}{5}\left(\frac{x^2}{2}+y^2x\right)\Big|_{x=0}^{x=\frac{1}{4}}dy=\frac{6}{5}\int_0^{\frac{1}{4}}\left(\frac{1}{32}+\frac{y^2}{4}\right)dy=\frac{7}{640}$$

例: 假设 $f(x,y)=\begin{cases}ke^{-(2x+3y)}, & x>0,y>0\\ 0, & \text{otherwise}\end{cases}$ 是随机向量(X,Y)的联合概率密度函数，求 (1) k 的值；（2） (X,Y) 的联合分布函数；（3） 计算 $P(Y\le X)$. P66



解 (1) $1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\,dxdy=\int_{0}^{+\infty}\int_{0}^{+\infty}ke^{-(2x+3y)}\,dxdy$

$=k\int_0^{+\infty}e^{-2x}dx\int_0^{+\infty}e^{-3y}\,dy=k\left(-\frac{e^{-2x}}{2}\right)_{x=0}^{x=\infty}\left(-\frac{e^{-3y}}{3}\right)_{y=0}^{y=\infty}=\frac{k}{6}$ 得 $k=6$

(2) $F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f(u,v)\,dudv=\begin{cases}\int_0^y\int_0^x 6e^{-(2u+3v)}\,dudv, & x>0,y>0\\ 0, & 其它\end{cases}$

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

(3) $P(Y\le X)=\int_0^{\infty}\int_y^{\infty}6e^{-(2x+3y)}\,dxdy=\frac{3}{5}$

定义: 均匀分布

设 G 是平面上的有界区域，其面积为 A, 若二维连续型随机向量 (X,Y) 的联合概率密度为

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

则称 (X,Y) 在 G 上服从均匀分布。

类似地，设 G 为空间上的有界区域，其体积为 V, 若三维连续型随机向量 (X,Y,Z) 的联合概率密度为

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

则称 (X,Y,Z) 在 G 上服从均匀分布。

定义: 二维正态随机向量

设二维正态随机变量 (X,Y) 具有联合概率密度

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

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) 为具有参数 $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 的二维正态随机变量，记作

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

例 3.5 设$(X,Y)\sim N(0,0,\sigma^2,\sigma^2,0)$,求 $P\{X<Y\}$.

解 易知 $f(x,y)=\dfrac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{x^2+y^2}{2\sigma^2}}$, $-\infty<x,y<+\infty$,

所以 $$P\{X<Y\}=\iint\limits_{x<y}\frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{x^2+y^2}{2\sigma^2}}\mathrm{d}x\mathrm{d}y.$$

引进极坐标 $x=r\cos\theta$, $y=r\sin\theta$,

则 $$P\{X<Y\}=\int_{\frac{\pi}{4}}^{\frac{5}{4}\pi}\int_0^{+\infty}\frac{1}{2\pi\sigma^2}r\mathrm{e}^{-\frac{r^2}{2\sigma^2}}\mathrm{d}r\mathrm{d}\theta=\frac{1}{2}.$$

例: 设 (X,Y) 在圆域 G： $x^2+y^2\le 4$ 上服从均匀分布，求:

(1) (X,Y) 的联合概率密度

(2) $P(0<X<1,0<Y<1)$.

解 （1） 圆域 G 的半径为 2，面积为 $A=4\pi$ ，故 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}\frac{1}{4\pi}, & (x,y)\in G\\ 0, & \text{其他}\end{cases}$$

(2) $P(0<X<1,0<Y<1)=\int_0^1\int_0^1 f(x,y)\,dxdy=\int_0^1\int_0^1\frac{1}{4\pi}dxdy$

$$=\frac{1}{4\pi}$$

思考： $P(0<X<2,0<Y<2)=\ \dfrac{1}{4}$

定义: 连续型随机变量 X,Y 的边缘概率密度

求法：先求 X,Y 的边缘分布函数$F_X(x)$ 和 $F_Y(y)$, 再求导

设连续型随机变量 X 和 Y 的联合概率密度函数为 $f(x,y)$.

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

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

X 的边缘概率密度为

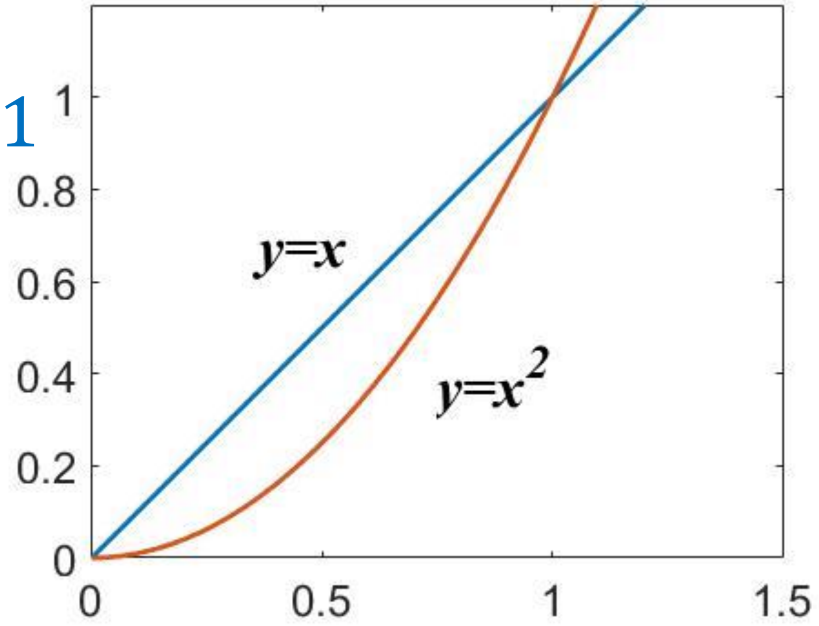
$$f_X(x) = \frac{dF_X(x)}{dx} = \int_{-\infty}^{+\infty} f(x,y)\,dy\,,$$

类似地，Y 的边缘概率密度为

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

例: 设随机变量 X 和 Y 的联合概率密度为 $f(x,y)=\begin{cases}6, & x^2\le y\le x\\ 0, & \text{其他}\end{cases}$

求边缘概率密度 $f_X(x), f_Y(y)$.



解: $f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy=\begin{cases}\int_{x^2}^{x} 6\,dy, & 0\le x\le 1\\ 0, & \text{其他}\end{cases}$

$$=\begin{cases}6(x-x^2), & 0\le x\le 1\\ 0, & \text{其他}\end{cases},$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\,dx=\begin{cases}\int_{y}^{\sqrt{y}} 6\,dx=6(\sqrt{y}-y), & 0\le y\le 1\\ 0, & \text{其他}\end{cases},$$

注1: 求 $f_X(x)$ 时，要计算$\int_{-\infty}^{+\infty} f(x,y)\,dy$，这是一个对 y 的积分，积分以后是 x 函数。积分时，需要注意y的取值范围；而首先需要确定x的范围；

定义: 独立的连续型随机变量

设连续型随机变量 X 和 Y 的联合概率密度函数为 $f(x,y)$. 其边缘概率密度分别为$f_X(x)$ 和 $f_Y(y)$. 如果对任意的 x,y 都有

$$f(x,y) = f_X(x) \cdot f_Y(y)$$

则称 X 和 Y 是相互独立的

Review: 独立的离散型随机变量

两个离散型随机变量 X 和 Y 是独立的，如果对每一对 (x_i, y_j) 都有

$$P\left(X = x_i, Y = y_j\right) = P(X = x_i) \cdot P(Y = y_j)$$

注：常用的两种判定独立性的定义

随机变量的独立性，见教材P73

定义 3.7　设 X 和 Y 为两个随机变量，若对于任意的 x 和 y，有

$$P\{X \leqslant x, Y \leqslant y\} = P\{X \leqslant x\}P\{Y \leqslant y\},$$

则称 X 和 Y 是相互独立(mutually independent)的.

若二维随机变量(X,Y)的分布函数为 $F(x,y)$，其边缘分布函数分别为 $F_X(x)$ 和 $F_Y(y)$，则上述独立性条件等价于对所有 x 和 y，有

$$F(x,y) = F_X(x)F_Y(y). \tag{3-13}$$

例: 设随机变量 (X,Y) 在圆域 $x^2+y^2\le 1$ 内服从均匀分布。问 X 和 Y 是否独立? P74

解 (X,Y) 的联合概率密度为 $f(x,y)=\begin{cases}1/\pi, & x^2+y^2\le 1\\ 0, & \text{otherwise}\end{cases}$

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy=\begin{cases}\frac{2}{\pi}\sqrt{1-x^2}, & -1\le x\le 1\\ 0, & \text{otherwise}\end{cases},$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\,dx=\begin{cases}\frac{2}{\pi}\sqrt{1-y^2}, & -1\le y\le 1\\ 0, & \text{otherwise}\end{cases},$$

可见, 在圆内 $f(x,y)\neq f_X(x)\cdot f_Y(y)$. **故** X 和 Y 不独立

例 3.14 设 X 和 Y 分别表示两个元件的寿命(单位:h),又设 X 与 Y 相互独立,且它们的概率密度分别为

$$f_X(x)=\begin{cases} e^{-x}, & x>0, \\ 0, & \text{其他}; \end{cases} \quad f_Y(y)=\begin{cases} e^{-y}, & y>0, \\ 0, & \text{其他}. \end{cases}$$

求 X 和 Y 的联合概率密度 $f(x,y)$.

解 由 X 和 Y 相互独立,可知

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

注: 注意在将 $f_X(x)$ 和 $f_Y(y)$ 相乘时，定义域需从一维变为二维

二维连续型随机变量的条件分布

定义 3.5 设(X,Y)是二维离散型随机变量，对于固定的j，若$P\{Y=y_j\}>0$，则称

$$P\{X=x_i \mid Y=y_j\}=\frac{P\{X=x_i, Y=y_j\}}{P\{Y=y_j\}}, \quad i=1,2,\cdots$$

为在 $Y=y_j$ 条件下随机变量 X 的条件分布律(conditional distribution)

对于连续型随机变量(X,Y)，因为$P\{X=x, Y=y\}=0$，所以不能直接由定义 3.5 来定义条件分布，但是对任意的$\varepsilon>0$，如果$P\{y-\varepsilon<Y\leqslant y+\varepsilon\}>0$，则可以考虑

$$P\{X\leqslant x \mid y-\varepsilon<Y\leqslant y+\varepsilon\}=\frac{P\{X\leqslant x, y-\varepsilon<Y\leqslant y+\varepsilon\}}{P\{y-\varepsilon<Y\leqslant y+\varepsilon\}}.$$

如果上述条件概率当$\varepsilon\to 0^+$时的极限存在，自然可以将此极限值定义为在$Y=y$条件下X的条件分布.

定义3.6: 设对任意的 $\varepsilon > 0$, 且 $P(y-\varepsilon < Y \le y+\varepsilon) > 0$, 如果极限

$$\lim_{\varepsilon\to 0^+} P(X \le x \mid y-\varepsilon < Y \le y+\varepsilon) = \lim_{\varepsilon\to 0^+} \frac{P(X \le x,\ \ y-\varepsilon < Y \le y+\varepsilon)}{P(y-\varepsilon < Y \le y+\varepsilon)}$$

存在, 则称此极限为在 $Y=y$ 的条件下 X 的条件分布函数，记作,

$$F_{X|Y}(x|y) = P(X \le x \mid Y = y)$$

1. 若 $f_Y(y) > 0$ 且连续，则不难验证 $F_{X|Y}(x|y) = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)}\,du$

2. 故有 $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$, $-\infty < x < \infty$, $f_Y(y) > 0$

3. 相似地， $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$, $-\infty < x < \infty$, $f_X(x) > 0$

例 3.11 设$(X,Y)\sim N(0,0,1,1,\rho)$，求 $f_{X|Y}(x\mid y)$ 与 $f_{Y|X}(y\mid x)$.

解 易知 $f(x,y)=\dfrac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}(-\infty<x,y<+\infty)$，所以

$$f_{X|Y}(x\mid y)=\frac{f(x,y)}{f_Y(y)}=\frac{1}{\sqrt{2\pi(1-\rho^2)}}e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}};$$

$$f_{Y|X}(y\mid x)=\frac{f(x,y)}{f_X(x)}=\frac{1}{\sqrt{2\pi(1-\rho^2)}}e^{-\frac{(y-\rho x)^2}{2(1-\rho^2)}}.$$

例: 随机变量 (X,Y) 的联合概率密度如下，求

$$f(x,y)=\begin{cases}\frac{6}{5}(x+y^2),\ 0\le x\le 1,\ 0\le y\le 1\\ 0,\quad \text{otherwise}\end{cases}$$

1. $X=0.8$ 时Y的条件分布函数
2. $Y\le 0.5$的概率
3. $X=0.8$ 的条件下 $Y\le 0.5$ 的概率.

解: X 和 Y 的边缘概率密度分别为

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\begin{cases}\int_0^1\frac{6}{5}(x+y^2)dy, 0\le x\le 1\\ 0,\qquad 其他\end{cases}=\begin{cases}\frac{6}{5}x+\frac{2}{5}, 0\le x\le 1\\ 0,\qquad 其他\end{cases}$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx=\begin{cases}\int_0^1\frac{6}{5}(x+y^2)dx, 0\le y\le 1\\ 0,\qquad 其他\end{cases}=\begin{cases}\frac{6}{5}y^2+\frac{3}{5}, 0\le y\le 1\\ 0,\qquad 其他\end{cases}$$

当 $X=x$ 时， Y 的条件概率密度为

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\frac{\frac{6}{5}(x+y^2)}{\frac{6}{5}x+\frac{2}{5}},0\le x\le 1,0\le y\le 1$$

1. $X=0.8$ 时Y的条件概率密度为

$$f_{Y|X}(y|0.8)=\frac{f(0.8,y)}{f_X(0.8)}=\frac{\frac{6}{5}(0.8+y^2)}{\frac{6}{5}\times 0.8+\frac{2}{5}}=\frac{12+15y^2}{17},\qquad 0\le y\le 1$$

2. $P(Y\le 0.5)=\int_{-\infty}^{0.5}f_Y(y)dy\ =\int_0^{0.5}\left(\frac{6}{5}y^2+\frac{3}{5}\right)dy\ =\frac{7}{20}$

3. $P(Y\le 0.5|X=0.8)=\int_{-\infty}^{0.5}f_{Y|X}(y|0.8)dy=\int_0^{0.5}\left(\frac{12+15y^2}{17}\right)dy=\frac{53}{136}.$

练习: 设 (X, Y) 的联合分布如下

$Y \backslash X$	1	2	4
1	0.15	0.30	0.35
3	0.05	0.12	a

(1) 求 a 的值

(2) 求 X 和 Y 的边缘分布律.

(3) X 和 Y 是否独立?

参考答案 (1) $a = 0.03$

(2)

X	1	2	4
p_k	0.20	0.42	0.38

Y	1	3
p_k	0.20	0.42

(3) X 和 Y 不独立，因为 $P(X = 1, Y = 1) = 0.15 \neq P(X = 1) \times P(Y = 1) = 0.20 \times 0.80 = 0.16$