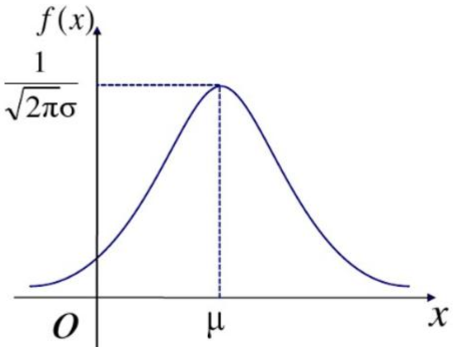
上节回顾



- 正态分布 $X \sim N(\mu, \sigma^2)$ $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

  $X \sim N(0, 1)$ $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

- 随机变量的函数 X是随机变量，$y = g(x)$是普通的实函数，则 $Y = g(X)$是 X 的函数。

定理1 X 为连续型随机变量，有概率密度函数

$$f_X(x)(-\infty < x < \infty)$$

若$g(x)$严格单调且处处可导，则$Y = g(X)$也是连续型随机变量。若令其中 $h(y)$ 是 $g(x)$ 的反函数，则Y的密度函数为

原函数 $y=g(x)$ 反函数$x=h(y)$

$$f_Y(y) = |h'(y)|\, f_X[h(y)] \quad y \in g(x)\text{的值域}$$

思考题：从南郊某地乘车到北区火车站有两条路可走，第一条路较短，但交通拥挤，所需时间T_1服从 $N(50, 100)$ 分布；第二条路线略长，但意外阻塞较少，所需时间T_2服从 $N(60, 16)$。

(1) 若有70分钟可用，问应走哪一条路？

(2) 若只有65分钟可用，又应走哪一条路？

分析：应该走在允许时间内有较大概率赶到火车站的路线。

解： (1) 若有70分钟可用

走第一条路线及时赶到的概率。

$$P\{T_1 \le 70\} = \Phi\left(\frac{70-50}{10}\right) = \Phi(2) = 0.9772$$

走第二条路线及时赶到的概率。

$$P\{T_2 \le 70\} = \Phi\left(\frac{70-60}{4}\right) = \Phi(2.5) = 0.9938$$

在这种场合，应走第二条路线。

(2) 若只有65分钟可用，又应走哪一条路？

走第一条路线及时赶到火车站的概率为

$$P\{T_1 \le 65\} = \Phi\left(\frac{65-50}{10}\right) = \Phi(1.5) = 0.9332$$

走第二条路线及时赶到的概率为

$$P\{T_2 \le 65\} = \Phi\left(\frac{65-60}{4}\right) = \Phi(1.25) = 0.8944$$

因此在这种场合，应走第一条路线更为保险。

Remark： 此处的μ和σ的含义分别是走完这段路程所需要的平均时间和均方差。

Chpt.3 Multi-Dimensional Random Variables and Their Distributions

第三章 多维随机变量及其分布

3.1.1 多维随机变量的定义

(1)很多随机现象中涉及多个变量，试验结果要用多个随机变量来表示。

考察一个人的健康时，需要检测身高、体重、血压、血糖...等多种因素，且这些因素本身还存在着关联.

(2)另外，当我们研究统计问题时也涉及到多个变量，比如，类似均值 $\dfrac{X_1+X_2+\cdots+X_n}{n}$，均方 $\dfrac{1}{n}(X_1^2+X_2^2+\cdots+X_n^2)$

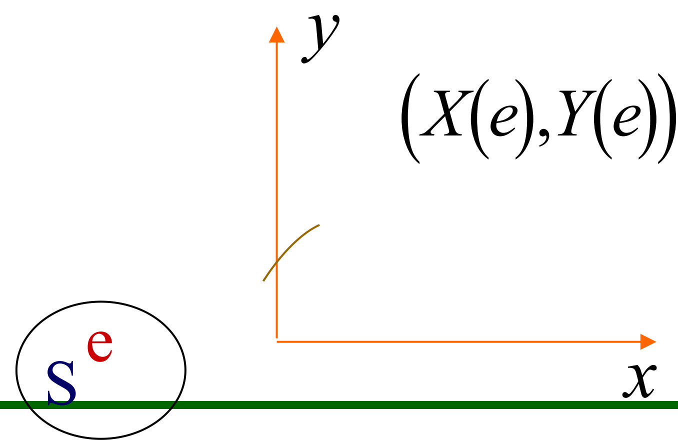
3.1 多维随机变量

[定义] 若随机变量$X_1,X_2,\ldots,X_n$定义在同一样本空间S上，就称这n个随机变量的整体$X=(X_1,X_2,\ldots,X_n)$ 为<u>n维随机向量</u>或<u>n维随机变量</u>(n-dimensional random variable).

对n维随机向量的研究从两个方面着手：

- 研究整体特性；　　联合分布
- 研究个体之间、个体与整体之间的关联关系。　　边缘分布

着重研究二维情形



3.1.2 多维随机变量的联合分布

多维随机变量$X_1,X_2,\ldots,X_n$，对于任意n个实数 $x_1, x_2,\ldots, x_n$ 或实向量$(x_1, x_2, \ldots, x_n)\in R^n$，称n元函数

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

为随机向量$X=(X_1,X_2,\ldots,X_n)$ 的(联合)分布函数.

注意：记

$$P\{X_1\le x_1,X_2\le x_2,\cdots,X_n\le x_n\}$$

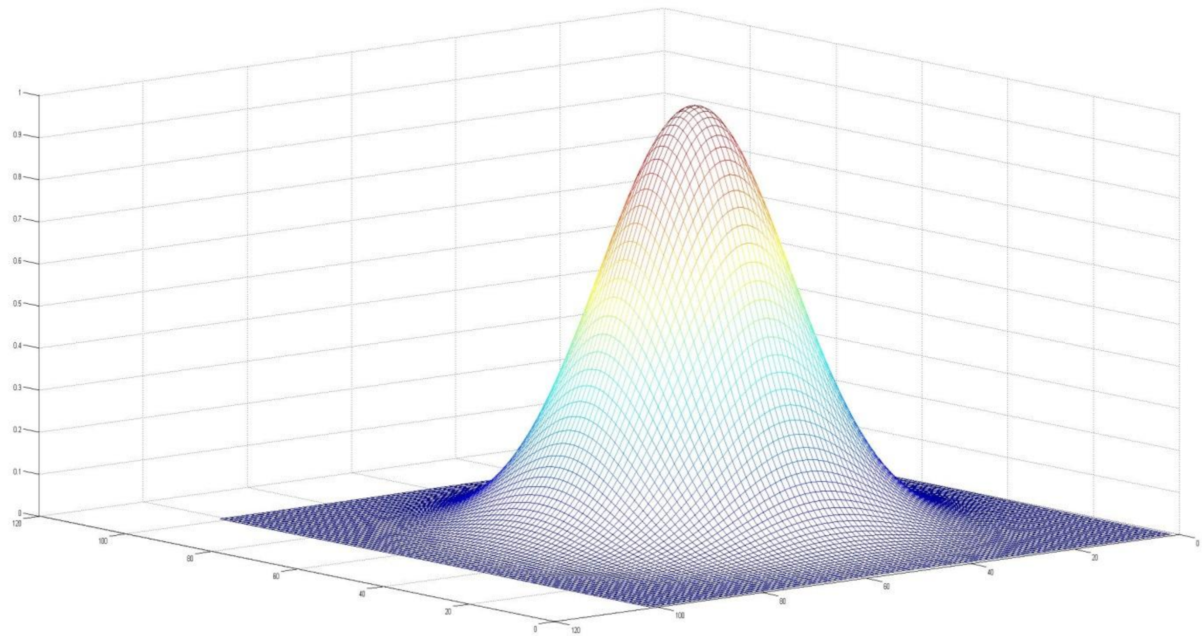
$$\overset{\Delta}{=} P\{(X_1\le x_1)\cap(X_2\le x_2)\cap\cdots\cap(X_n\le x_n)\}$$

Remark：注意到多维随机变量$X_1,X_2,\ldots,X_n$是针对同一试验而言的，是在同一个样本空间S上定义的随机变量。

3.1.2 多维随机变量的联合分布

引入图例

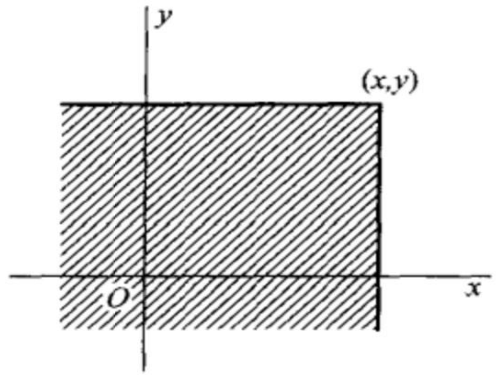
- 一个二维随机变量的高斯联合分布的pdf 概率密度函数如下图现实



联合分布的基本性质：以2维为例

- 对于给定任意y, 如果$x_1 < x_2$, 则$F(x_1, y) \le F(x_2, y)$，对y也成立
- $0 \le F(x, y) \le 1$且对于任意固定x,y有：
  $F(-\infty, y) = 0; F(x, -\infty) = 0; F(-\infty, -\infty) = 0; F(\infty, \infty) = 1$
  注：$F(\infty, y) \ne 1$，$F(x, \infty) \ne 1$
- $F(x, y)$关于x右连续，即$F(x+0, y) = F(x, y)$
  $F(x, y)$关于y右连续，即$F(x, y+0) = F(x, y)$
- 对于$x_1 < x_2, y_1 < y_2$有
  $F(x_2, y_2) + F(x_1, y_1) - F(x_2, y_1) - F(x_1, y_2) \ge 0$

对二维随机向量(X, Y)，分布函数 $F(x, y)=P\{X\leq x, Y\leq y\}$ 表示随机向量 (X, Y)落在 (x, y) 为顶点的位于该点左下方的无穷矩形内的概率.

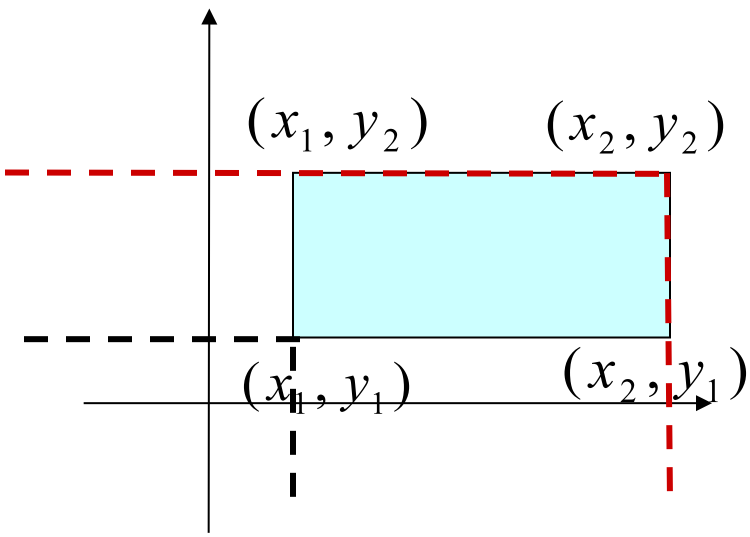


依照上述解释，(X, Y)落在 $[x_1 < x \leq x_2,\ y_1 < y \leq y_2]$ 的概率为

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

$$= F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2)$$

$$+ F(x_1, y_1)$$



3.1.2 多维随机变量的联合分布

n维联合分布的性质

- $0 \le F(x_1,x_2,\cdots,x_n) \le 1$
- $F(x_1,x_2,\cdots,x_{i-1},-\infty,x_{i+1},\cdots x_n) = 0 \qquad i = 1,2,\cdots,n$
- $F(\infty,\infty,\cdots,\infty) = 1$
- $F(x_1,x_2,\cdots,x_n)$是 $x_1,x_2,\cdots,x_n$的不减函数，即对任意的两个向量$(x_1,x_2,\cdots,x_n)$和 $(x_1',x_2',\cdots,x_n')$ ，只要 $x_i' \ge x_i$ (对任意$i = 1,2,\cdots,n$) ，总有

$$F(x_1',x_2',\cdots,x_n') \ge F(x_1,x_2,\cdots,x_n)$$

3.1.2 多维随机变量的联合分布

注意到联合分布的定义，以及

$$\{(X_1 \le x_1) \cap (X_2 \le x_2) \cap \cdots \cap (X_n \le x_n)\}$$

$$\subseteq \{(X_1 \le x_1') \cap (X_2 \le x_2') \cap \cdots \cap (X_n \le x_n')\}$$

$$P\{(X_1 \le x_1) \cap (X_2 \le x_2) \cap \cdots \cap (X_n \le x_n)\}$$

$$\le P\{(X_1 \le x_1') \cap (X_2 \le x_2') \cap \cdots \cap (X_n \le x_n')\}$$

即 $$F(x_1', x_2', \cdots, x_n') \ge F(x_1, x_2, \cdots, x_n)$$

3.1.3 离散多维随机变量的联合分布

随机向量只取有限组或可列组取值，就称为离散型随机向量.

列出所有各组可能值及取这些值的概率，就可得其概率分布.

Example 口袋中有2白球3黑球, 连取两次，每次任取一球. 设X为第一次得白球数，Y为第二次得白球数. 对

(1)有放回

(2)无放回

求：两种情况 (X,Y)的联合分布.

解 (1) 有放回：X与Y可能取的值都是0(黑球)与1（白球），各种情况搭配及相应概率如下（注意直观看待独立与否）：

{X=0, Y=0} 表示第一次取黑球且第二次也取黑球，因为有放回，两次取球相互独立的，其概率都是3/5，故

$$P\{X=0,Y=0\}=P\{X=0\}P\{Y=0\}=\frac{3}{5}\times\frac{3}{5}$$

同理 $P(X=0,Y=1)=\frac{3}{5}\times\frac{2}{5}$

$$P(X=1,Y=0)=\frac{2}{5}\times\frac{3}{5} \qquad P(X=1,Y=1)=\frac{2}{5}\times\frac{2}{5}$$

表1：有放回时的(X,Y)联合分布

X \ Y	0	1
0	$\frac{3}{5}\times\frac{3}{5}$	$\frac{3}{5}\times\frac{2}{5}$
1	$\frac{2}{5}\times\frac{3}{5}$	$\frac{2}{5}\times\frac{2}{5}$

(2) X与Y可能取的值与(1)相同，但因为无放回，两次结果是不独立的, 利用第一章乘积事件概率公式，得

$$P\{X=0,Y=0\}=P\{X=0\}P\{Y=0\mid X=0\}=\frac{3}{5}\times\frac{2}{4}$$

同理 $P\{X=0,Y=1\}=\frac{3}{5}\times\frac{2}{4}$

$P\{X=1,Y=0\}=\frac{2}{5}\times\frac{3}{4}$ $P\{X=1,Y=1\}=\frac{2}{5}\times\frac{1}{4}$

表2：无放回时的(X,Y)联合分布

X \ Y	0	1
0	$\frac{3}{5}\times\frac{2}{4}$	$\frac{3}{5}\times\frac{2}{4}$
1	$\frac{2}{5}\times\frac{3}{4}$	$\frac{2}{5}\times\frac{1}{4}$

3.1.3 离散多维随机变量的联合分布

一般，离散型的二维随机变量联合分布列为

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

或写成表格的形式如下表

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

二维矩阵

$$p_{ij} \geq 0$$

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

3.1.4 N维连续型随机变量的密度函数

[pdf-Definition] 对n维随机变量$X=(X_1,X_2,\cdots,X_n)$ 若存在n元可积的非负函数$f(x_1,x_2,\cdots,x_n)$,对于任意的$(x_1,x_2,\cdots,x_n)\in R^n$

$$F(x_1,x_2,\cdots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}f(y_1,y_2,\cdots,y_n)dy_1dy_2\cdots dy_n$$

则 $X=(X_1,X_2,\cdots,X_n)$ 是n维的连续随机变量，F 称为它的连续型分布，$f(x_1,x_2,\cdots,x_n)$ 称为 (联合)密度函数.

3.1.4 N维连续型随机变量的密度函数

[1] 显然，密度函数满足如下条件：

$$f(x_1, x_2, \cdots, x_n) \ge 0$$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(y_1, y_2, \cdots, y_n) dy_1 dy_2 \cdots dy_n = 1$$

[2] 对连续型随机向量，分布函数 $F(x_1, x_2, \cdots, x_n)$ 对每一变量都是连续的，且在密度函数 $f(x_1, x_2, \cdots, x_n)$ 的连续点，有偏导数

$$\frac{\partial^n F}{\partial x_1 \cdots \partial x_n} = f(x_1, x_2, \cdots, x_n)$$

[3] G是 R^n 中的任意一个区域，$(X_1, X_2, \cdots, X_n)$ 落入区域G内的概率为

$$P\{X \in G\} = \int_G \cdots \int f(y_1, y_2, \cdots, y_n) dy_1 dy_2 \cdots dy_n$$

Example 设二维随机向量(X,Y)的密度函数为

$$f(x,y)=\begin{cases} Ae^{-2(x+y)} & x>0, y>0 \\ 0 & elsewhere \end{cases}$$

1) 确定常数A；

2) 求分布函数

3) 计算 $P\{X<1, Y<2\}$；

4) 计算概率$P\{X+Y<1\}$

解 1) 由联合密度的性质，应有

$$1=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy=\int_{0}^{\infty}\int_{0}^{\infty} Ae^{-2(x+y)}dxdy=A/4$$

故 A=4

2) 分布函数 $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dudv$，我们来分块计算它.

当$x \le 0$或$y \le 0$时，$f(x,y)=0$，故 $F(x,y)=0$

当$x>0$且$y>0$时

$$F(x,y)=\int_0^x\int_0^y 4e^{-2(u+v)}dudv+\iint_{elsewhere}0dudv$$

$$=(1-e^{-2x})(1-e^{-2y})$$

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

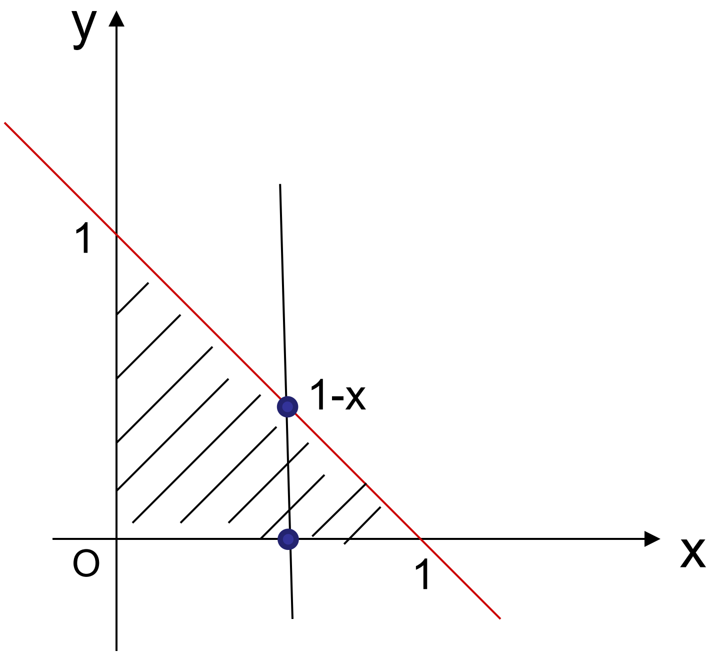
3) $P\{X<1,Y<2\}=F(1,2)=(1-e^{-2})(1-e^{-4})$

4) $P\{X+Y<1\}=\iint\limits_{\substack{x+y<1\\ x>0,\ y>0}} 4e^{-2(x+y)}dxdy$

二重积分化二次积分

$$=\int_0^1\left[\int_0^{1-x}4e^{-2(x+y)}dy\right]dx$$

$$=1-3e^{-2}$$



3.2 边际（边缘）分布

3.2.1 边缘分布的定义

n维随机变量 $X=(X_1,X_2,\cdots,X_n)$，具有联合分布 $F(x_1,x_2,\cdots,x_n)$，每一个随机变量 X_i 又有其自身的分布函数 $F_{X_i}(x_i)$，称之为**边缘分布函数**。

边缘分布与联合分布的关系

$$
\begin{aligned}
F_{X_i}(x_i) &= P\{X_i \le x_i\} \\
&= P\{\{X_1 \le \infty\}\cap\cdots\cap\{X_{i-1} \le \infty\}\cap\{X_i \le x_i\}\cap\{X_{i+1} \le \infty\}\cdots\{X_n \le \infty\}\} \\
&= P\{X_1 \le \infty,\cdots,X_{i-1} \le \infty, X_i \le x_i, X_{i+1} \le \infty,\cdots X_n \le \infty\} \\
&= F(\infty,\cdots,\infty,x_i,\infty,\cdots,\infty)
\end{aligned}
$$

3.2.2 离散随机变量的边缘分布

以离散型二维随机变量(X,Y)为例，$p_{ij}=P\left(X=x_i,Y=y_j\right),\ i,j=1,2,\cdots$

$$P\{X=x_i\}=\sum_{j=1}^{\infty}p_{ij}\overset{\Delta}{=}p_{i\bullet}$$

--称为(X,Y)关于X的边缘分布

$$P\{Y=y_j\}=\sum_{i=1}^{\infty}p_{ij}\overset{\Delta}{=}p_{\bullet j}$$

--称为(X,Y)关于Y的边缘分布

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

Example 求例1的边际分布.

$$P\{X=0\}=\frac{3}{5}\times\frac{3}{5}+\frac{3}{5}\times\frac{2}{5}$$ 类似得其它各概率，如下面两表.

Y \ X	0	1	$p_{\cdot j}$
0	$\frac{3}{5}\times\frac{3}{5}$	$\frac{2}{5}\times\frac{3}{5}$	$\frac{3}{5}$
1	$\frac{3}{5}\times\frac{2}{5}$	$\frac{2}{5}\times\frac{2}{5}$	$\frac{2}{5}$
$p_{i\cdot}$	$\frac{3}{5}$	$\frac{2}{5}$	

Y \ X	0	1	$p_{\cdot j}$
0	$\frac{3}{5}\times\frac{2}{4}$	$\frac{2}{5}\times\frac{3}{4}$	$\frac{3}{5}$
1	$\frac{3}{5}\times\frac{2}{4}$	$\frac{2}{5}\times\frac{1}{4}$	$\frac{2}{5}$
$p_{i\cdot}$	$\frac{3}{5}$	$\frac{2}{5}$	

Remark (1)与(2)的**联合分布不同，但边际分布相同**。这说明如果边际分布给定，联合分布却不能惟一确定，还要考虑分量间的相互关系.

以二维连续型随机变量(X,Y)为例，设其密度函数为f(x,y),
分布函数为F(x,y), 则X的边际分布函数

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

$$= \int_{-\infty}^{x} \Bigg[\underbrace{\int_{-\infty}^{\infty} f(x,y)dy}_{f_X(x)} \Bigg] dx$$

$$F_Y(y) = \int_{-\infty}^{y} \Bigg[\underbrace{\int_{-\infty}^{\infty} f(x,y)dx}_{f_Y(y)} \Bigg] dy$$

根据连续型随机变量的定义，X是连续型随机变量，它的密度函数就是$f_X(x)$。
同理，Y是连续型随机变量，其密度函数为$f_Y(y)$。
$f_X(x)$和$f_Y(y)$称为(X, Y) (或$f(x, y)$)的**边际密度**。

3.2.3 连续型随机变量的边缘分布

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)dy$$ --称为(X,Y)关于X的边际密度

$$f_Y(y)=\int_{-\infty}^{\infty} f(x,y)dx$$ --称为(X,Y)关于Y的边际密度

由此，也可以看出边际密度的两种求法：

[1] 由联合分布函数$F(x,y)$

计算$F(x,+\infty)$得$F_X(x)$，计算$F(+\infty,y)$得$F_Y(y)$，而后微分得到

$$f_X(x)=F_X'(x),\quad f_Y(y)=F_Y'(y)$$

[2] 由联合分布密度$f(x,y)$

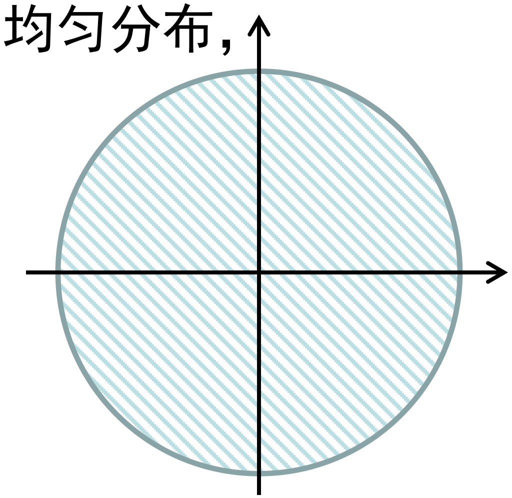
对y积分得到 $$f_X(x)=\int_{-\infty}^{\infty} f(x,y)dy$$

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

3.2.3 连续型随机变量的边缘分布

Example (X,Y)在圆形区域 $x^2+y^2\le 1$ 上服从如下的均匀分布，求边际密度.

$$f(x,y)=\begin{cases}1/\pi & x^2+y^2\le 1\\ 0 & elsewhere\end{cases}$$



解：因为 $|x|>1$时，$f(x,y)=0$, 此时$f_x(x)=0$;

$|x|\le 1$时，

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy = \int_{x^2+y^2\le 1}\frac{1}{\pi}dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}dy=\frac{2}{\pi}\sqrt{1-x^2}$$

$$f_X(x)=\begin{cases}\frac{2}{\pi}\sqrt{1-x^2} & |x|\le 1\\ 0 & |x|>1\end{cases}$$

$$f_Y(y)=\begin{cases}\frac{2}{\pi}\sqrt{1-y^2} & |y|\le 1\\ 0 & |y|>1\end{cases}$$

这里，虽然(X,Y)的联合分布是均匀分布，但边际分布却不是均匀分布.

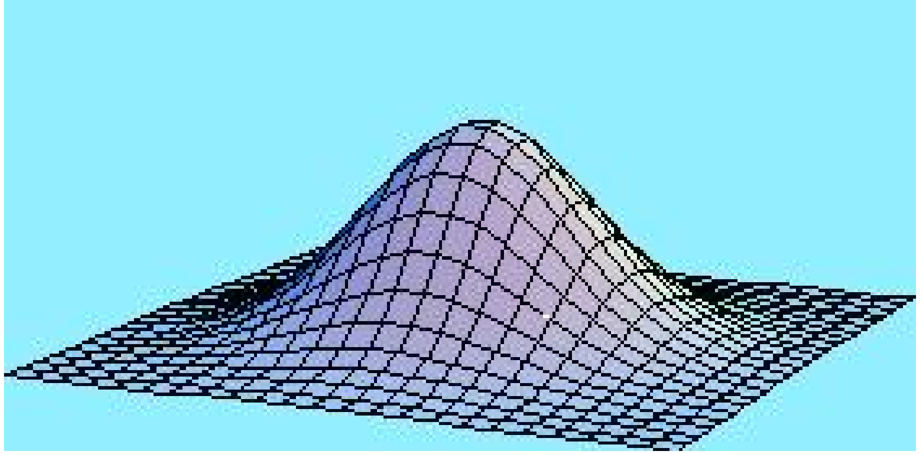
Example 设二维随机变量（X,Y）的分布概率密度是

$$f(x,y)=\frac{1}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2}$$

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

称(X,Y)为服从参数 $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 的二元正态分布，

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



求其边缘概率密度。

[解]：在上式的指数上对y配方，

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

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

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

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

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

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

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

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

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

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

$$\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right)^2\right]dy$$

$\Downarrow$ $$t=\frac{1}{\sqrt{1-\rho^2}}\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right), 则 dy=\sigma_2\sqrt{1-\rho^2}dt$$

$$=\frac{1}{2\pi\sigma_1}\exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt$$

$$=\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]$$

由标准正态分布可知 $\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt=\sqrt{2\pi}$

$$f_x(x)=\int_{-\infty}^{\infty}f(x,y)dy=\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}$$

$$f_y(y)=\int_{-\infty}^{\infty} f(x,y)dx=\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\}$$

$$X \sim N(\mu_1,\sigma_1^2)$$

$$Y \sim N(\mu_2,\sigma_2^2)$$

上述结论是说，二元正态分布的边际分布仍是正态分布，并且与ρ无关.

但反过来不正确，即若(X,Y)的边际分布都是正态分布，其联合分布却未必是二元正态分布.

思考： **设 (X,Y)的联合密度为**

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

求边际分布.