第三章 多维随机变量及其分布

在实际问题中，试验结果有时需要同时用**两个或两个以上随机变量**来描述.

例如 1）用身高和体重来描述儿童的发育情况.

2）通过对含碳、含硫、含磷量的测定来研究钢的成分.

要研究这些随机变量之间的联系，就需要考虑多维随机变量及其取值规律。

——**多维分布**

主要内容

- 3.1_3.2 二维随机变量与边缘分布
- 3.4 相互独立的随机变量
- 3.5 两个随机变量的函数的分布

3.1_3.2 二维随机变量与边缘分布

二维随机变量的定义

定义： 设S为随机试验E的样本空间，

$$\forall e \in S \xrightarrow{\text{一定法则}} \exists \left(X(e), Y(e) \right) \in R^2$$

则称(X, Y)为二维随机变量。

讨论：

a. 二维随机变量作为一个整体的概率特性；

b. 其中每一个随机变量的概率特性、与整体概率特性之间的关系。

二维随机变量的分布函数

定义： 设(X, Y)为二维随机变量，对任何一对实数(x, y), 事件

$$(X \le x) \cap (Y \le y) \quad \textbf{(记为}\ (X \le x, Y \le y)\textbf{)}$$

的概率 $P\left(X \le x, Y \le y\right)$ 定义了一个二元实

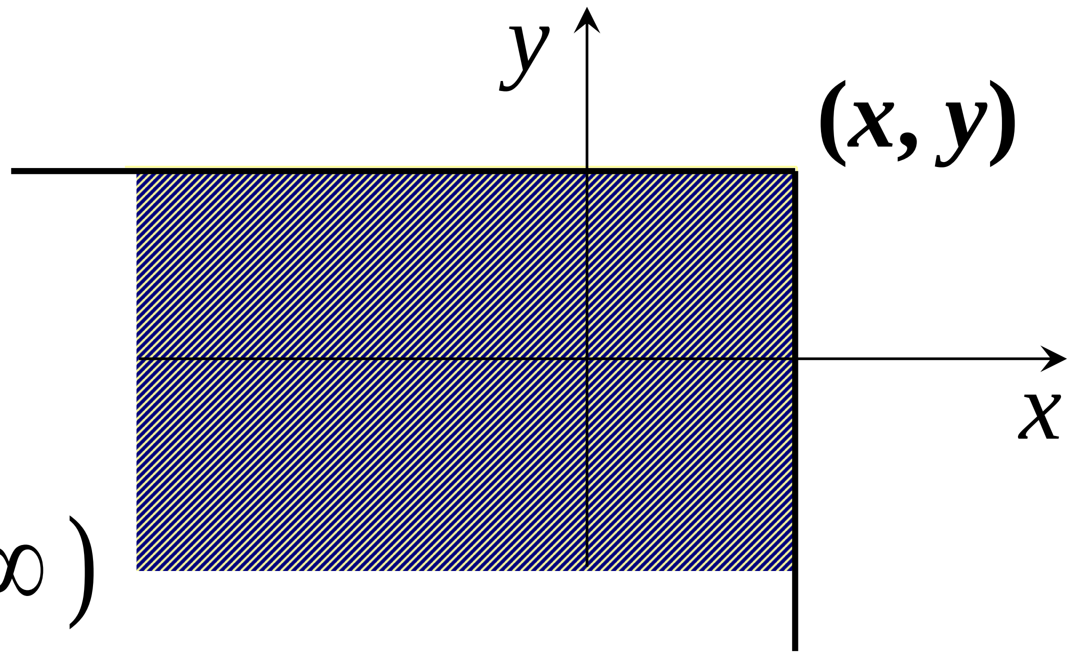
函数$F(x, y)$，称为二维随机变量(X, Y)的分布函数，或称为X与Y的联合分布函数，

即

$$F(x, y) = P\left(X \le x, Y \le y\right)$$

联合分布函数的几何意义

如果用平面上的点 (x, y) 表示二维随机变量 (X, Y)的一组可能的取值，则 $F(x, y)$ 表示 (X, Y) 的取值落入图中所示角形区域的概率.



$(-\infty, +\infty)$

联合分布函数的性质

① 对每个变量单调不减

固定 x , 对任意的 $y_1 < y_2$,

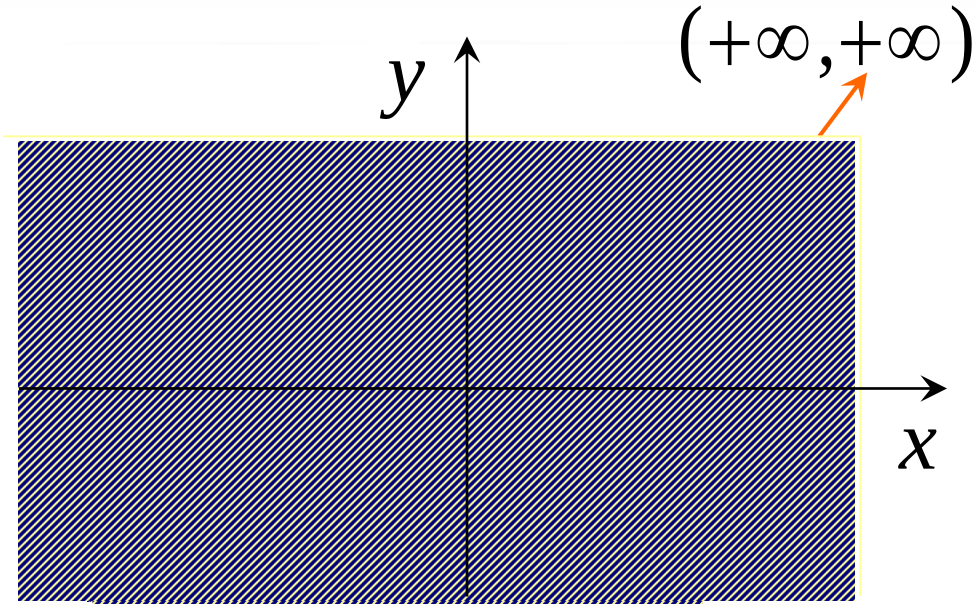
$$F(x, y_1) \le F(x, y_2)$$

固定 y , 对任意的 $x_1 < x_2$,

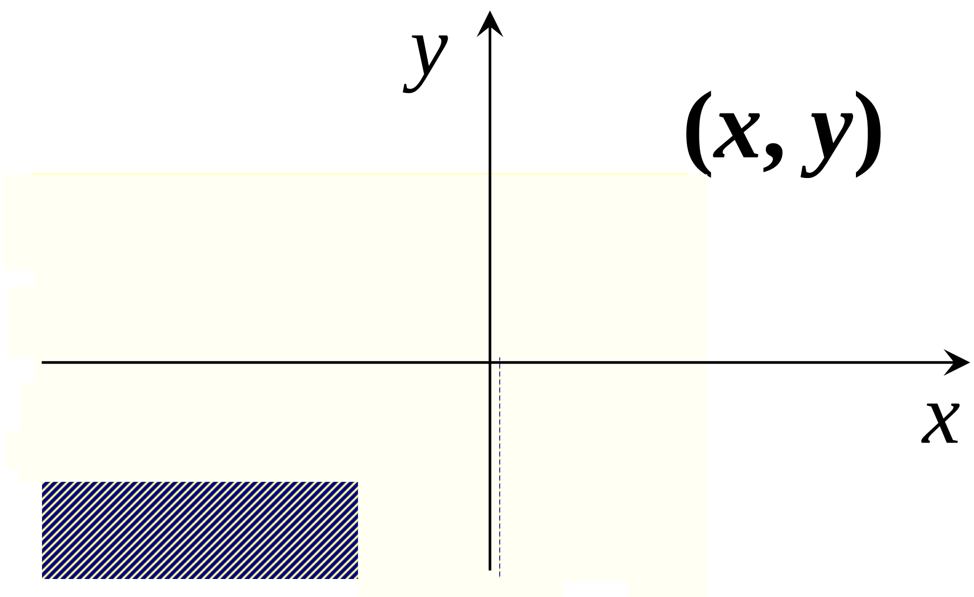
$$F(x_1, y) \le F(x_2, y)$$

联合分布函数的性质

② $0 \le F(x,y) \le 1$



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

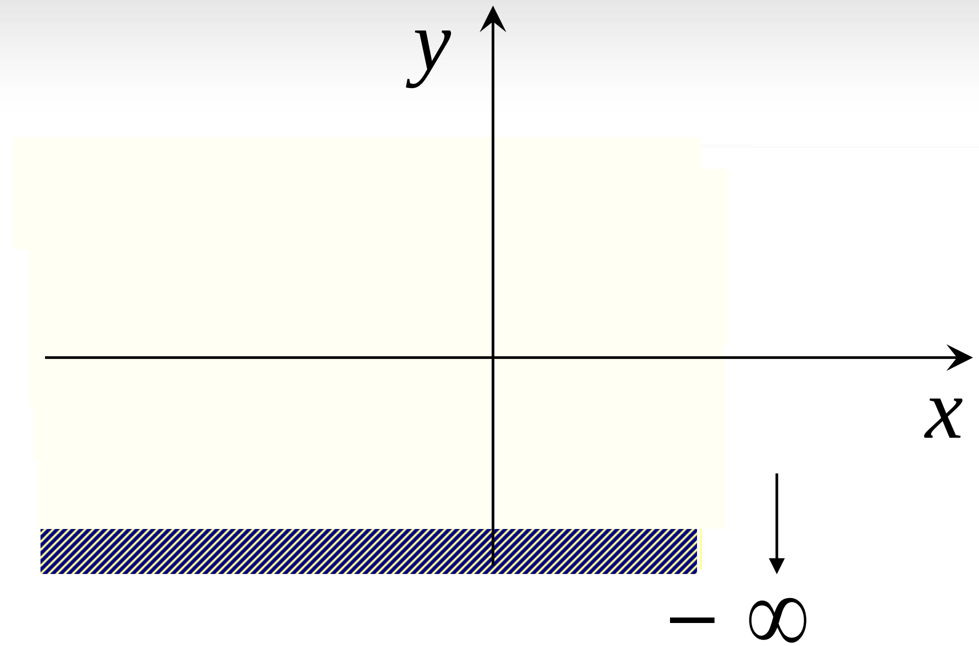


$$F(-\infty,-\infty)=0$$

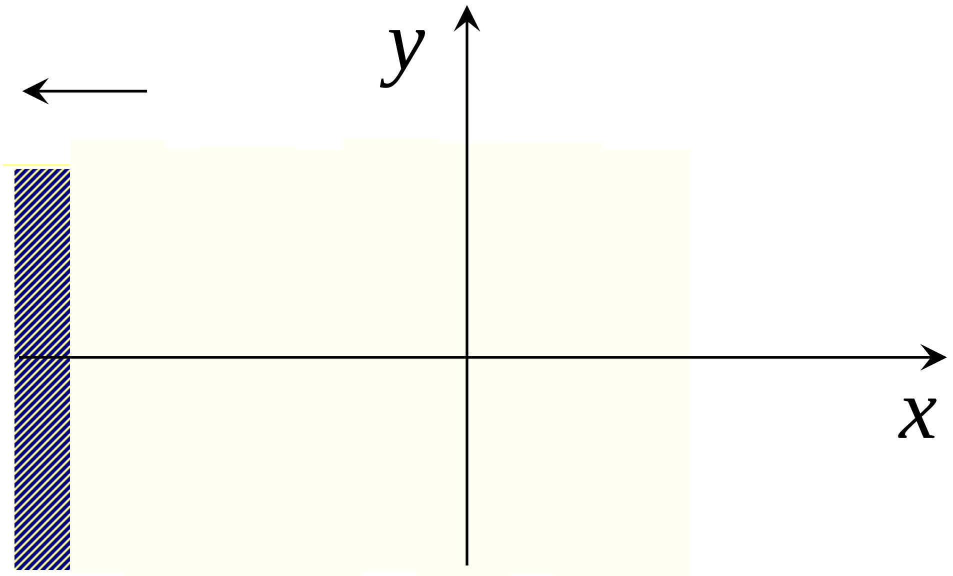
$(-\infty,-\infty)$

联合分布函数的性质

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



$-\infty$



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

联合分布函数的性质

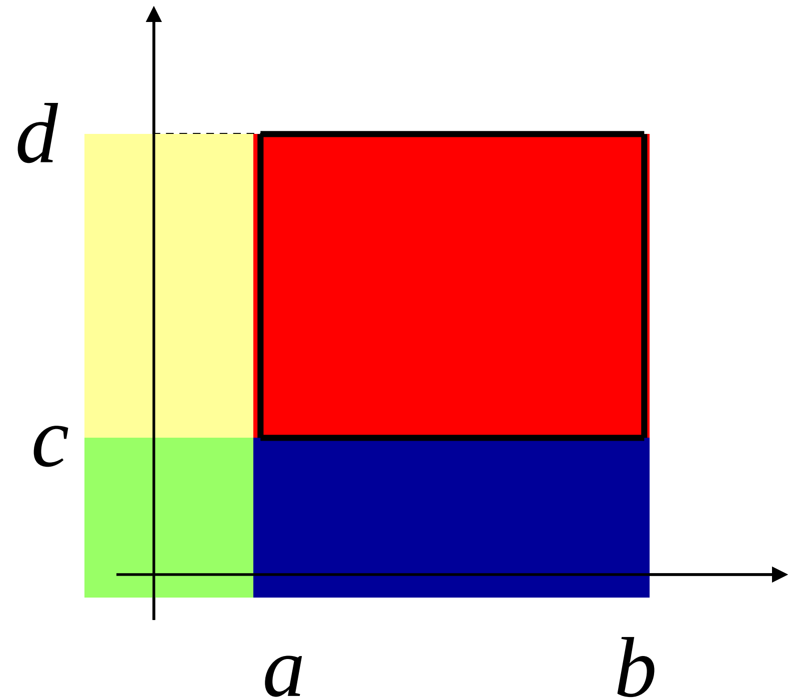
③ 对每个变量右连续

$$F(x, y) = F(x+0, y)$$

$$F(x, y) = F(x, y+0)$$

④ 对于任意 $a < b, c < d$

$$F(b,d) - F(b,c) - F(a,d) + F(a,c) \geq 0$$



事实上

$$F(b,d) - F(b,c)$$

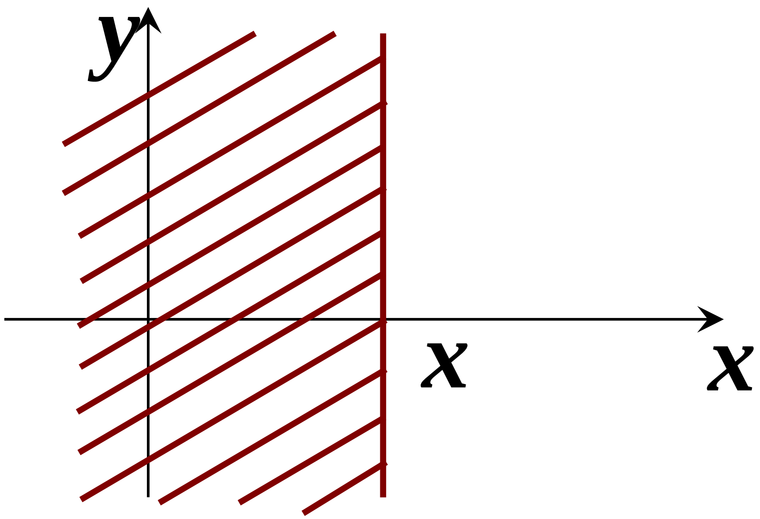
$$- F(a,d) + F(a,c)$$

$$= P(a < X \leq b, c < Y \leq d) \geq 0$$

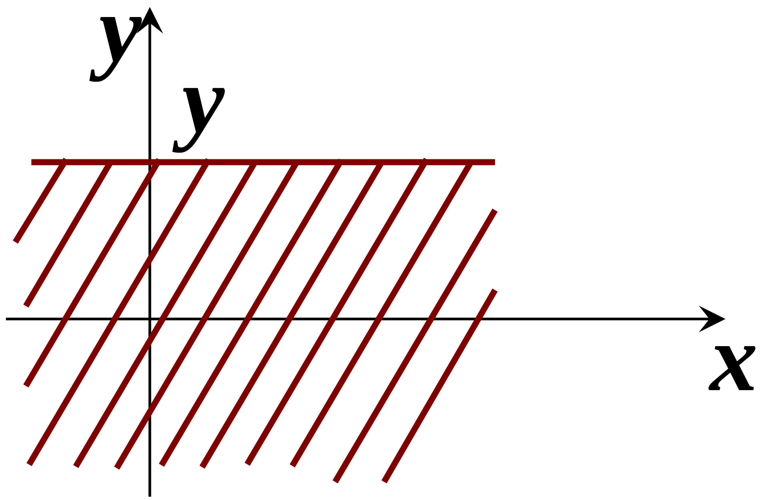
二维随机变量的边缘分布函数

由联合分布函数 $\Longrightarrow$ 边缘分布函数, 逆不真.

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



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



二维离散型随机变量及其概率特性

定义 若二维随机变量(X, Y)所有可能的取值为有限对或可列无穷对, 则称(X, Y)为二维离散型随机变量。

$a.$ 用联合分布律来描述二维离散型随机变量的整体概率特性；

$b.$ 用边缘分布律来描述整体与每个随机变量之间的关系。

联合分布律

设(X,Y)的所有可能的取值为

$$(x_i, y_j), \qquad i, j = 1, 2, \ldots\ldots$$

则称

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

为二维随机变量**(X,Y) 的分布律，**也称为**X和Y的联合分布律。**

注： $p_{ij} \geq 0, \qquad i, j = 1, 2, \ldots\ldots$

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

(X, Y) 的分布律

$X \backslash Y$	y_1	$\cdots$	y_j	$\cdots$
x_1	p_{11}	$\cdots$	p_{1j}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_i	p_{i1}	$\cdots$	p_{ij}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

二维离散型随机变量的边缘分布律

$$P(X=x_i)=\sum_{j=1}^{+\infty} p_{ij} \overset{\text{记作}}{=} p_{i\bullet}, \qquad i=1,2,\cdots$$

$$P(Y=y_j)=\sum_{i=1}^{+\infty} p_{ij} \overset{\text{记作}}{=} p_{\bullet j}, \qquad j=1,2,\cdots$$

由联合分布律可确定边缘分布律, 其逆不真.

联合分布律及边缘分布律

$X \backslash Y$	$y_1 \quad \cdots \quad y_j \quad \cdots$	$p_{i\bullet}$
x_1	$p_{11} \quad \cdots \quad p_{1j} \quad \cdots$	$p_{1\bullet}$
$\vdots$	$\vdots \quad\quad \vdots \quad\quad \vdots \quad\quad \vdots$	$\vdots$
x_i	$p_{i1} \quad \cdots \quad p_{ij} \quad \cdots$	$p_{i\bullet}$
$\vdots$	$\vdots \quad\quad \vdots \quad\quad \vdots \quad\quad \vdots$	$\vdots$
$p_{\bullet j}$	$p_{\bullet 1} \quad \cdots \quad p_{\bullet j} \quad \cdots$	1

例 设随机变量X在1, 2, 3, 4四个整数中等可能地取一个值，另一个随机变量Y在1-X中等可能地取一整数值，试求X与Y的联合分布律和边缘分布律。

解： X 取1时，Y 的可能取值为1，

X 取2时，Y的可能取值为1、2，

X 取3时，Y的可能取值为1、2、3，

X 取4时，Y的可能取值为1、2、3、4.

$$P(X=1,Y=1)=P(X=1)P(Y=1|X=1)=\frac{1}{4}\times 1=\frac{1}{4}$$

$$P(X=2,Y=1)=P(X=2)P(Y=1|X=2)=\frac{1}{4}\times\frac{1}{2}=\frac{1}{8}$$

同理有：

$$P(X=2,Y=2)=\frac{1}{8},$$

$$P(X=3,Y=1)=P(X=3,Y=2)$$

$$=P(X=3,Y=3)=\frac{1}{12},$$

$$P(X=4,Y=1)=P(X=4,Y=2)$$

$$=P(X=4,Y=3)=P(X=4,Y=4)=\frac{1}{16}.$$

故X与Y的联合分布律和边缘分布律如下：

X \ Y	1	2	3	4	$p_{i\bullet}$
1	1/4	0	0	0	1/4
2	1/8	1/8	0	0	1/4
3	1/12	1/12	1/12	0	1/4
4	1/16	1/16	1/16	1/16	1/4
$p_{\bullet j}$	25/48	13/48	7/48	1/16	1

二维连续型随机变量及其概率特性

定义 设二维随机变量(X,Y)的分布函数为$F(x,y)$, 若存在非负可积函数$f(x,y)$,使得对于任意实数x, y有

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

则称(X,Y)为二维连续型随机变量， $f(x,y)$为(X,Y)的概率密度, 或称为X和Y的联合概率密度。

概率密度$f(x, y)$的性质

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

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

3. 若G是平面上的区域，则

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

4. 若$f(x, y)$在点(x, y)连续，则有

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

5. $P(X = a, Y = b) = 0$

$$P(X = a, -\infty < Y < +\infty) = 0$$

$$P(-\infty < X < +\infty, Y = a) = 0$$

二维连续型随机变量的边缘概率密度

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

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

$$F_Y(y)=\int_{-\infty}^{y}\boxed{\int_{-\infty}^{+\infty} f(u,v)du}dv$$

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

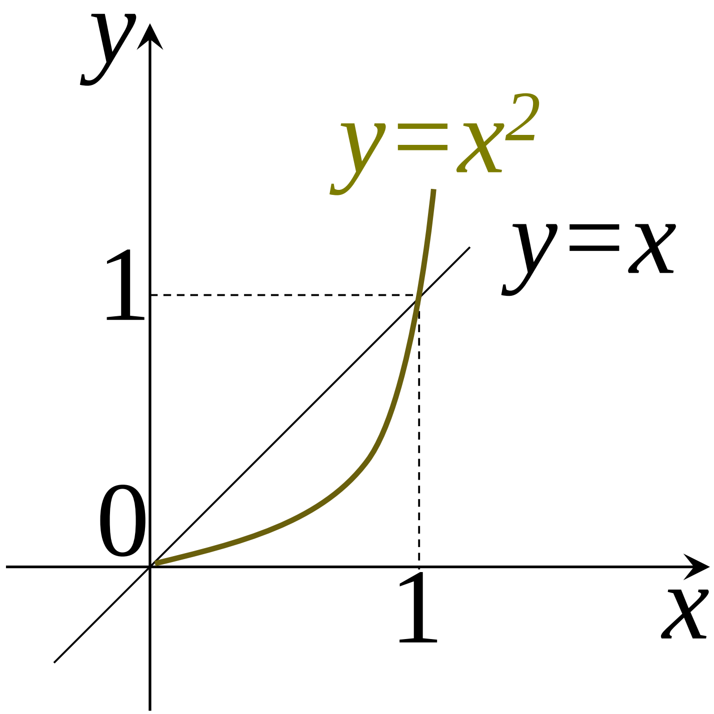
与离散型相同，已知联合分布可以求得边缘分布；反之则不能唯一确定.

例 设随机变量X 和 Y 具有联合概率密度

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

(1) 求边缘概率密度$f_X(x)$, $f_Y(y)$.

(2) 求概率$P(X<1/2,\ Y<1/2)$.

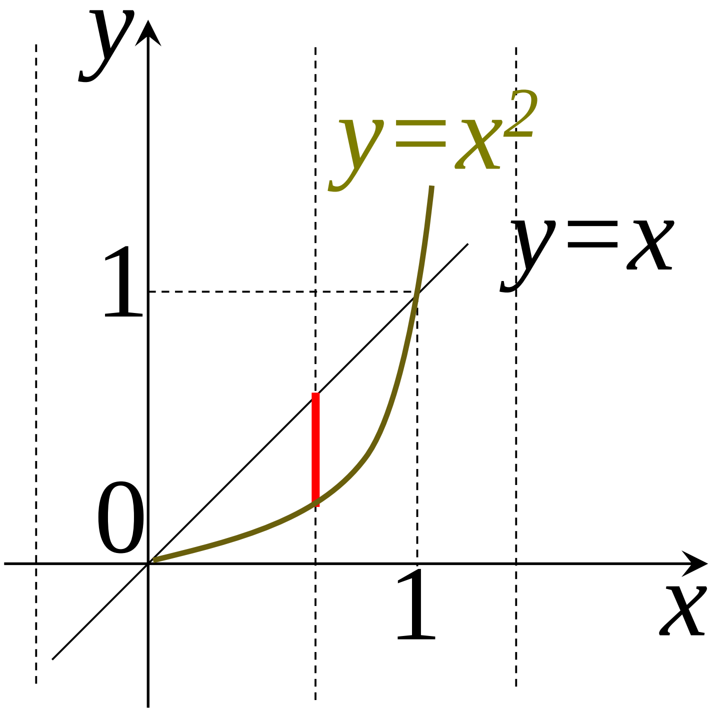


解:

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

$$=\begin{cases}\int_{x^2}^{x} 6dy, & 0\le x\le 1\\ 0, & 其他\end{cases}$$

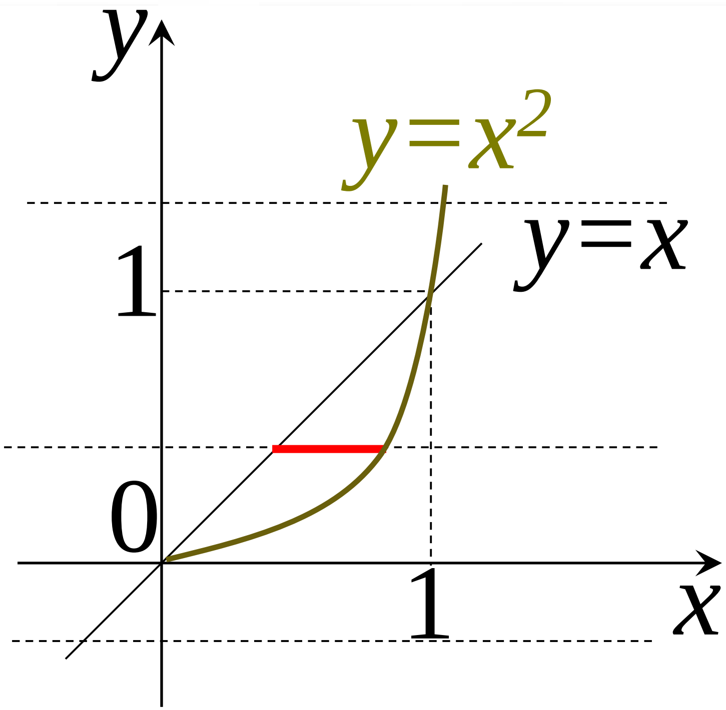
$$=\begin{cases}6(x-x^2), & 0\le x\le 1\\ 0, & 其他\end{cases}$$



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

$$=\begin{cases}\int_y^{\sqrt{y}} 6dx, & 0\le y\le 1\\ 0, & 其他\end{cases}$$

$$=\begin{cases}6(\sqrt{y}-y), & 0\le y\le 1\\ 0, & 其他\end{cases}$$



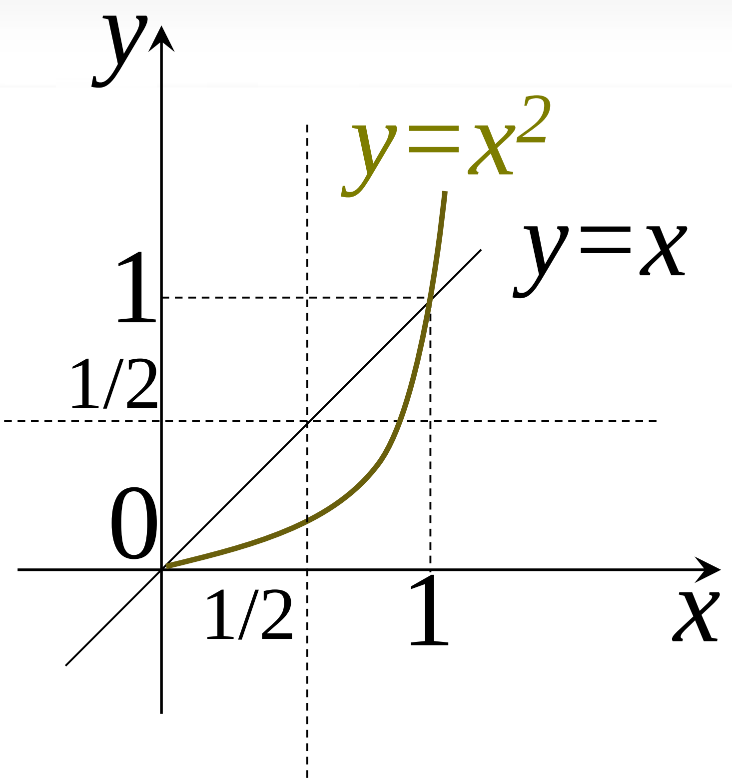
$$(2)\ P(X \le 1/2, Y \le 1/2)$$

$$= \int_{-\infty}^{1/2} \int_{-\infty}^{1/2} f(x,y)\,dx\,dy$$

$$= \int_{0}^{1/2} \int_{x^2}^{x} 6\,dy\,dx$$

$$= \int_{0}^{1/2} 6(x - x^2)\,dx$$

$$= 1/2$$



NANJING UNIVERSITY

3.4 相互独立的随机变量

（将事件独立性推广到

随机变量的独立性）

离散型 $\boldsymbol{X}$与$\boldsymbol{Y}$独立 $\Longleftrightarrow$ 对一切 $\boldsymbol{i}, \boldsymbol{j}$ 有

$$p_{ij} = p_{i\bullet} p_{\bullet j}$$

即 $P(X = x_i, Y = y_j) = P(X = x_i) P(Y = y_j)$

连续型 $\boldsymbol{X}$与$\boldsymbol{Y}$独立 $\Longleftrightarrow$ 对任何 x, y 有

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

二维随机变量（X, Y）中的X与Y相互独立,则边缘分布完全确定联合分布。

例 已知 (X, Y) 的分布律为

$X \backslash Y$	1	2
0	$1/6$	$1/6$
1	$2/6$	$2/6$

讨论X, Y是否独立？

解：

经计算可知$P\{X=0\}=1/3, P\{X=1\}=2/3$

$$P\{Y=1\}=1/2, P\{Y=2\}=1/2$$

则有 $P\{X=0,Y=1\}=1/6=P\{X=0\}P\{Y=1\}$

$$P\{X=0,Y=2\}=1/6=P\{X=0\}P\{Y=2\}$$

$$P\{X=1,Y=1\}=2/6=P\{X=1\}P\{Y=1\}$$

$$P\{X=1,Y=2\}=2/6=P\{X=1\}P\{Y=2\}$$

故 X,Y 相互独立。

例 已知 (X, Y) 的概率密度函数为

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

讨论 X, Y 是否独立？

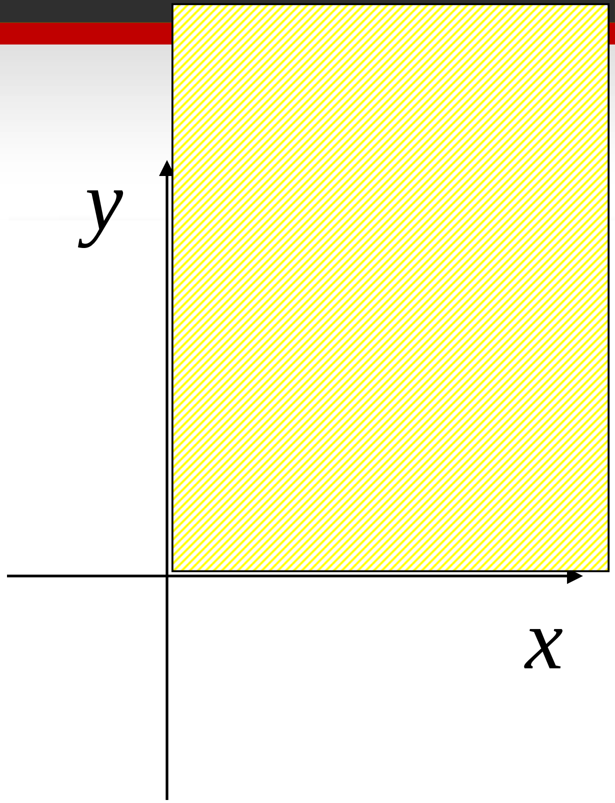
解:

X的边缘密度函数为

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

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

$$=\begin{cases}2e^{-2x}, & x>0,\\ 0, & \text{其他}.\end{cases}$$



同理，

$$f_Y(y) = \begin{cases} e^{-y}, & y > 0, \\ 0, & 其他 \end{cases}$$

显然，

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

故 X,Y 相互独立。

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

例 一负责人到达办公室的时间均匀分布在8-12时，他的秘书到达办公室的时间均匀分布在7-9时，设他们两人到达的时间相互独立，求他们到达办公室的时间相差不超过5分钟(1/12小时)的概率。

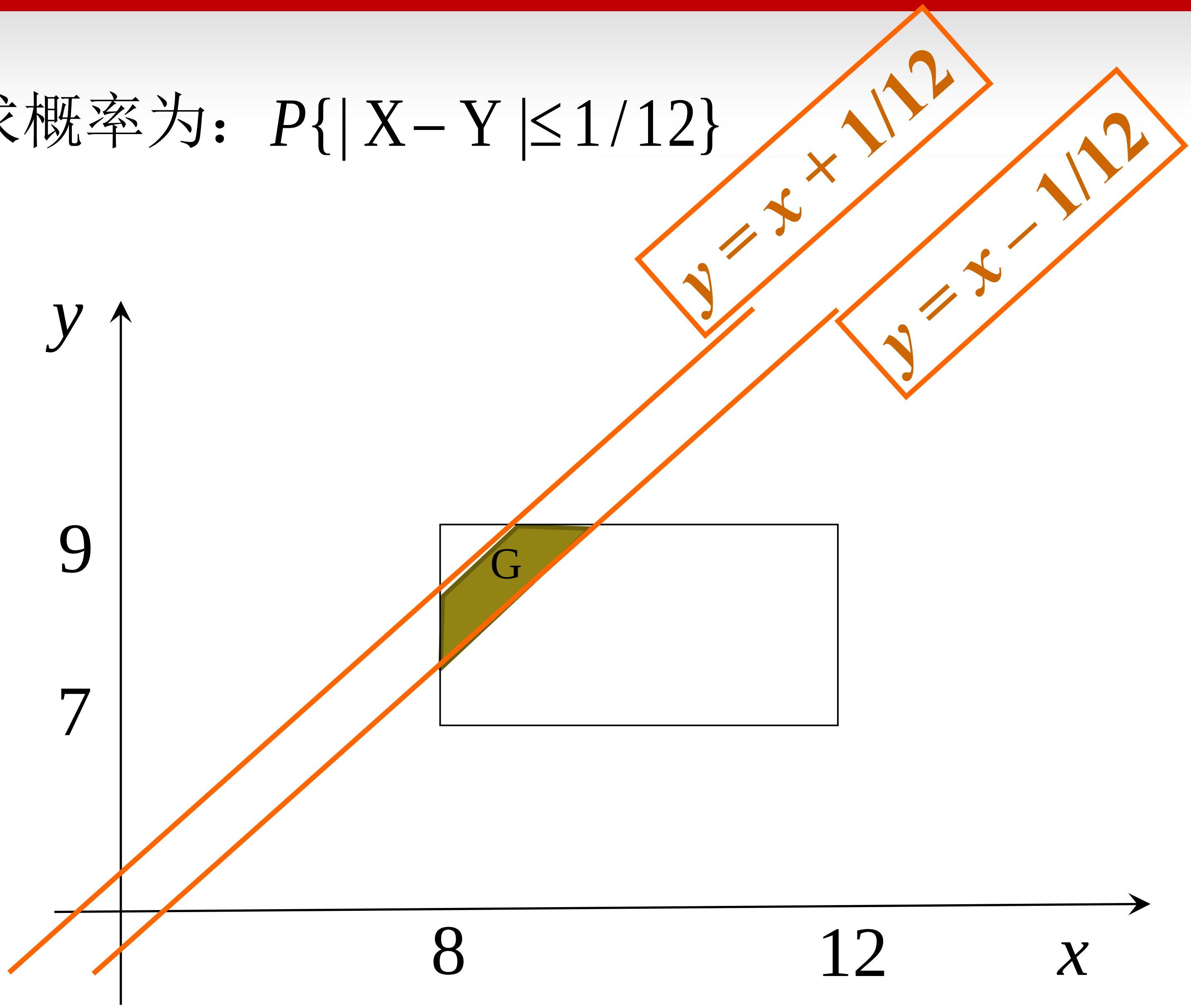
解： 设X和Y分别是负责人和秘书到达办公室的时间，则X和Y的概率密度分别为：

$$f_X(x)=\begin{cases}1/4, & 8<x<12, \\ 0, & 其他,\end{cases} \qquad f_Y(y)=\begin{cases}1/2, & 7<y<9, \\ 0, & 其他,\end{cases}$$

因为X与Y相互独立，则(X, Y)的概率密度为：

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}1/8, & 8<x<12, 7<y<9, \\ 0, & 其他,\end{cases}$$

所求概率为：$P\{|X-Y|\le 1/12\}$



$$P\{|X-Y|\le 1/12\}$$

$$=\iint_G f(x,y)dxdy$$

$$=\frac{1}{8}S_G$$

$$=\frac{1}{8}(\frac{1}{2}\times\frac{13}{12}\times\frac{13}{12}-\frac{1}{2}\times\frac{11}{12}\times\frac{11}{12})$$

$$=\frac{1}{48}$$

3.5 两个随机变量的函数的分布

$Z = X + Y$

设(X, Y)的联合概率密度为$f(x, y)$, 则

$$F_Z(z) = P(Z \le z)$$

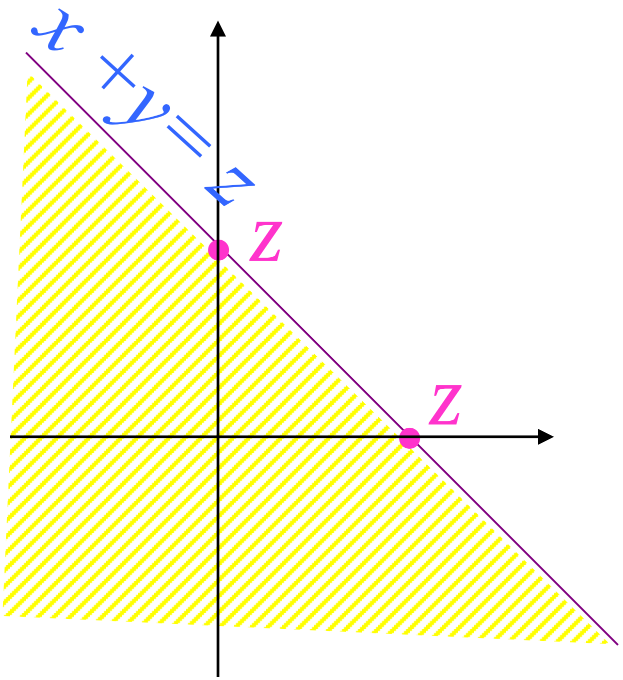
$$= P(X + Y \le z)$$

$$= \iint_{x+y\le z} f(x, y)dxdy$$

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

或

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



$$-\infty < z < +\infty$$

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

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

上式对z求导数：

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x) dx \quad -\infty < z < +\infty \quad (1)$$

同理：

$$f_Z(z) = \int_{-\infty}^{+\infty} f(z-y, y) dy \quad -\infty < z < +\infty \quad (2)$$

特别地，若X ,Y相互独立，则

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)dx \overset{\text{记作}}{=} f_X(z) * f_Y(z)$$

$$-\infty < z < +\infty \quad (3)$$

或 $$f_Z(z)=\int_{-\infty}^{+\infty} f_X(z-y) f_Y(y)dy \overset{\text{记作}}{=} f_X(z) * f_Y(z)$$

$$-\infty < z < +\infty \quad (4)$$

称之为函数$f_X(z)$与$f_Y(z)$的卷积

例 在一简单电路中，两电阻R_1和R_2串联连接，设R_1，R_2相互独立，它们的概率密度均为

$$f(x)=\begin{cases}\dfrac{10-x}{50}, & 0\le x\le 10,\\ 0, & \text{其他}\end{cases}$$

求总电阻$R=R_1+R_2$的概率密度.

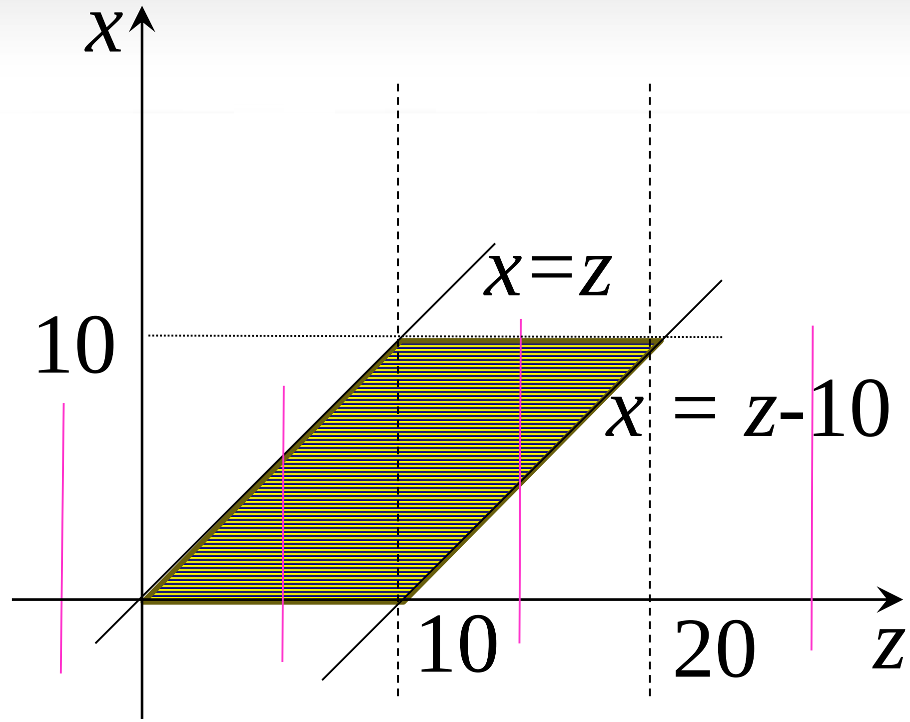
解：（直接利用公式）

$$f_R(z)=\int_{-\infty}^{+\infty} f(x)f(z-x)dx$$

先确定被积函数非零时x和z的取值范围：$\begin{cases}0\le x\le 10\\ 0\le z-x\le 10\end{cases}$ 看下页图

$$f_R(z)$$

$$=\int_{-\infty}^{+\infty} f(x)f(z-x)dx$$



$$=\begin{cases}\int_0^z \frac{10-x}{50}\bullet\frac{10-(z-x)}{50}dx, & 0\le z<10,\\ \int_{z-10}^{10} \frac{10-x}{50}\bullet\frac{10-(z-x)}{50}dx, & 10<z\le 20,\\ 0, & \text{其他}\end{cases}$$

$$
= \begin{cases} \dfrac{1}{15000}(600z - 60z^2 + z^3), & 0 \le z < 10, \\ \dfrac{1}{15000}(20 - z)^3, & 10 < z \le 20, \\ 0, & \text{其他} \end{cases}
$$

例 已知(X,Y)的联合概率密度为

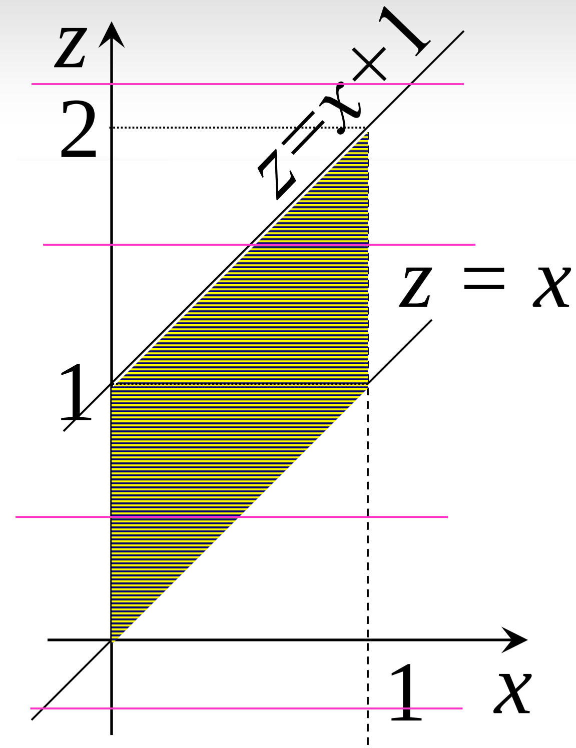
$$f(x,y)=\begin{cases}1, & 0<x<1,0<y<1\\ 0, & \text{其他}\end{cases}$$

$Z=X+Y$,求$f_Z(z)$

解法一（直接利用公式）

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

先确定被积函数非零时x和z的取值范围：$\begin{cases}0<x<1\\ 0<z-x<1\end{cases}$ 看下页图



$$f_Z(z)=$$

$$\int_{-\infty}^{+\infty} f(x,z-x)dx=\begin{cases}\int_0^z 1dx, & 0\le z<1,\\ \int_{z-1}^{1} 1dx, & 1\le z<2,\\ 0, & 其他\end{cases}$$

$$f_Z(z)=\begin{cases} z, & 0\le z<1 \\ 2-z, & 1\le z<2 \\ 0, & \text{其他} \end{cases}$$

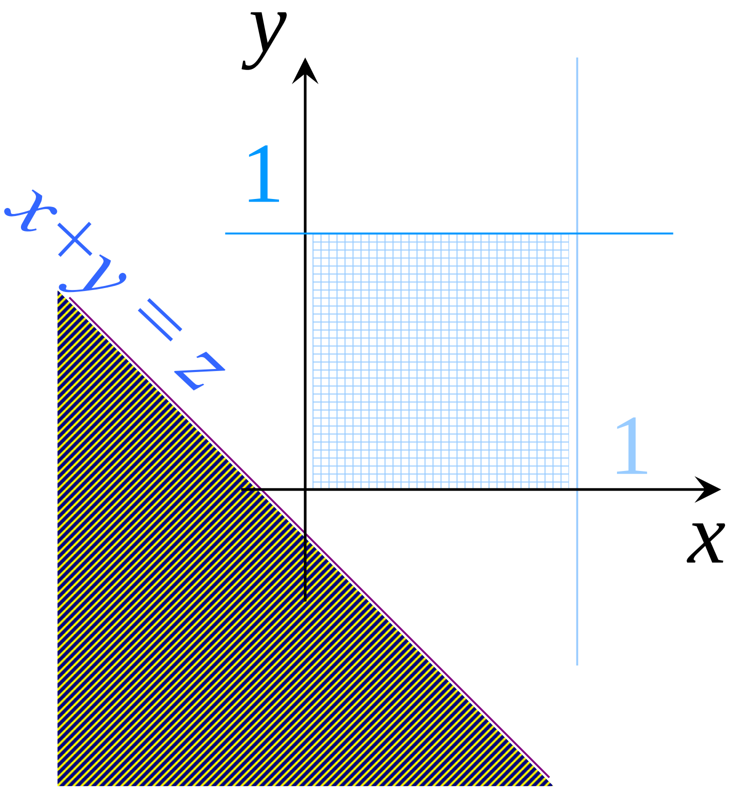
解法二 从分布函数出发

$$F_Z(z)=P(X+Y\le z)$$

$$=\iint\limits_{x+y\le z} f(x,y)dxdy$$

当$z<0$时，

$$F_Z(z)=0 \Longrightarrow f_Z(z)=0$$



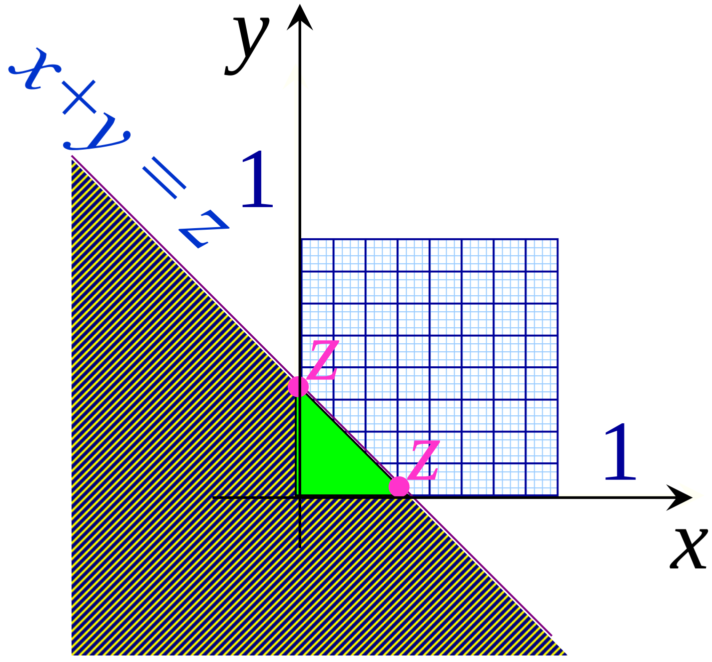
当$0 \le z < 1$时，

$$F_Z(z) = \int_0^z dx \int_0^{z-x} 1dy$$

$$= \int_0^z (z-x)dx$$

$$= z^2/2$$

$$\Rightarrow \quad f_Z(z) = z$$



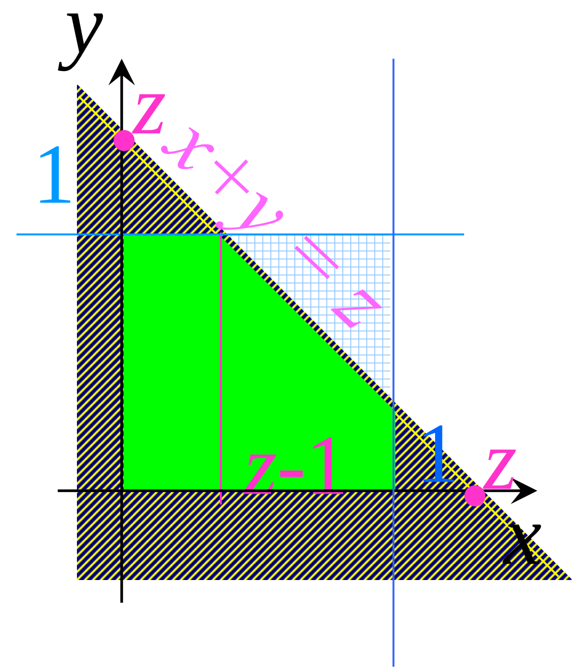
当$1 \le z < 2$ 时，

$$F_Z(z) = (z-1) + \int_{z-1}^{1} dx \int_{0}^{z-x} 1 dy$$

$$= z - 1 + \int_{z-1}^{1} (z-x) dx$$

$$= 2z - z^2/2 - 1$$

$$f_Z(z) = 2 - z$$



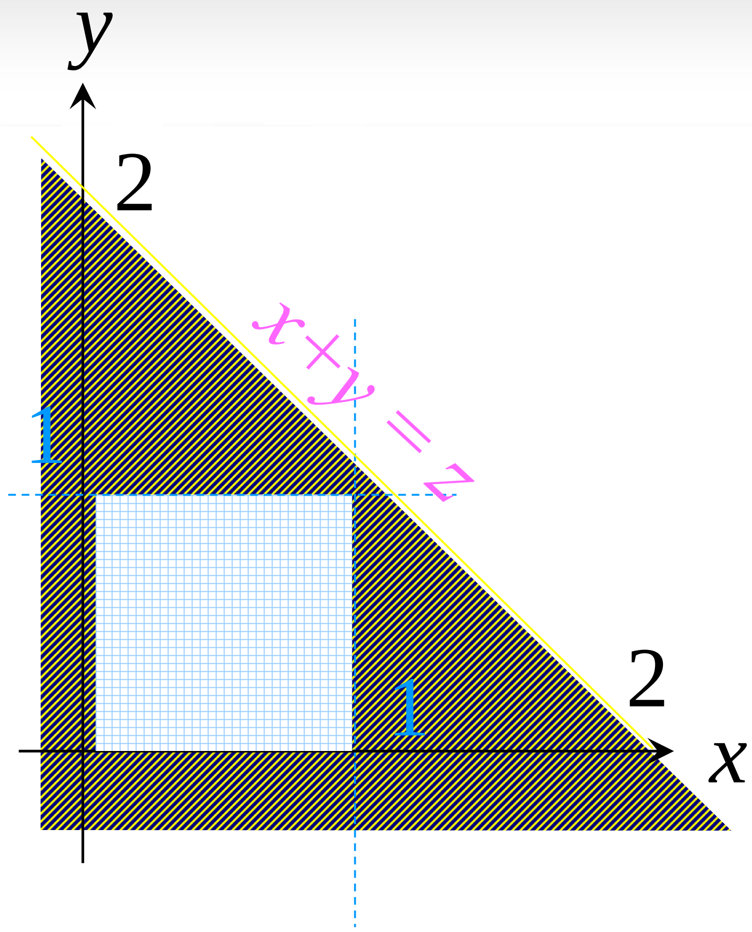
当 $z \geq 2$时，

$$F_Z(z) = 1$$

$$f_Z(z) = 0$$

综上，

$$f_Z(z) = \begin{cases} z, & 0 \leq z < 1 \\ 2 - z, & 1 \leq z < 2 \\ 0, & \text{其他} \end{cases}$$



正态随机变量的结论

若X,Y相互独立，$X \sim N(\mu_1,\sigma_1^2), Y \sim N(\mu_2,\sigma_2^2)$

则 $X+Y \sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$

推广 若 $X_1,X_2,\cdots,X_n$ 相互独立

$$X_i \sim N(\mu_i,\sigma_i^2), i=1,2,\cdots,n$$

则 $\sum_{i=1}^{n} X_i \sim N(\sum_{i=1}^{n}\mu_i,\sum_{i=1}^{n}\sigma_i^2)$

THE END