probability
probability

第三章

多维随机变量

probability
probability

§3.1 二维随机变量及其分布

§3.2 随机变量的独立性

§3.3 条件分布

§3.4 随机变量的函数及其分布

§3.1 二维随机变量及其分布

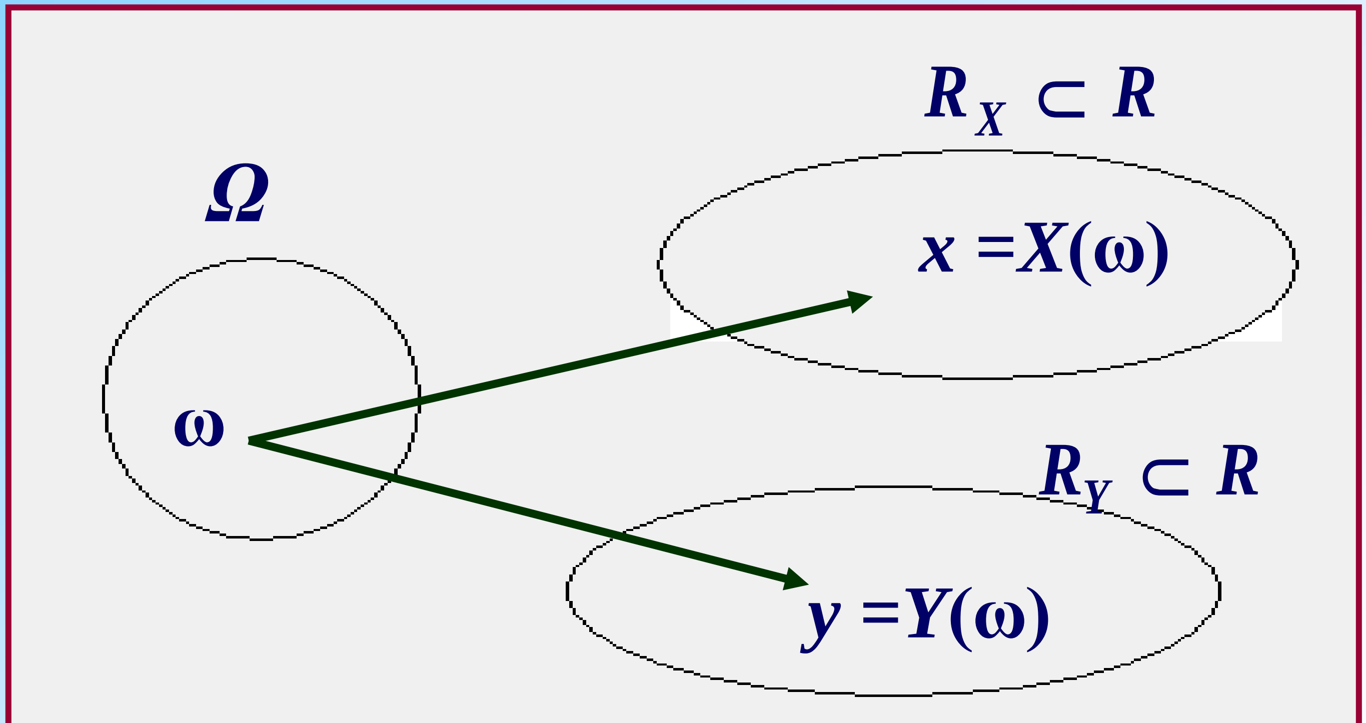
一.联合分布函数

多维随机变量的引入

定义 设随机试验E 的样本空间为Ω，对于每一样本点$\omega \in \Omega$，有两个实数 $X(\omega)$,$Y(\omega)$ 与之对应，称它们构成的有序数组(X, Y) 为**二维随机变量**.

注1： X, Y 都是定义在 Ω 上的随机变量, 有

$$F_X(x) = P\{X \le x\}, \qquad F_Y(x) = P\{Y \le y\}$$



例如，为描述一个人的身材特征，用身高 H 和体重 W 来描述.

假设 $\Omega=\{$电子科大全体男生$\}$，任选 1 名男生 $\omega\in\Omega$，相应的身高和体重是 $H(\omega)$ 与 $W(\omega)$.

即一个样本点 ω 对应着两个变量，(H, W) 是定义在 Ω 上的二维随机变量.

注2 记事件

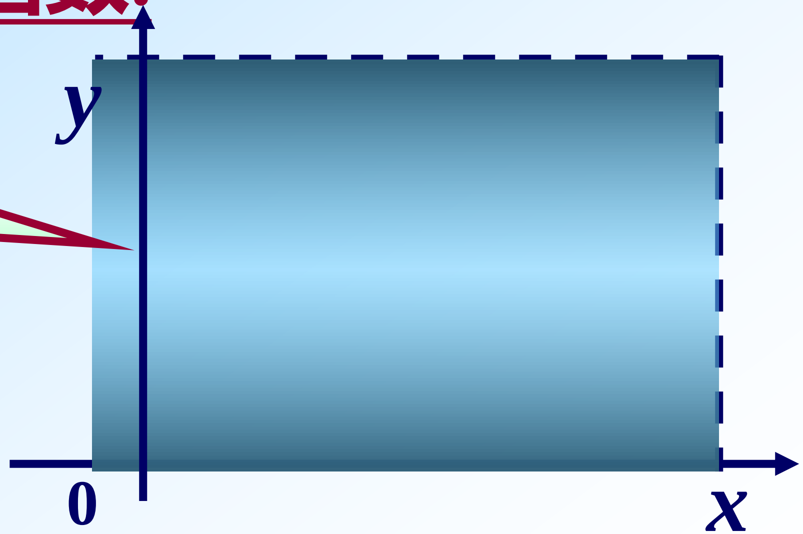
$$\{X \le x, Y \le y\} = \{X \le x\} \cap \{Y \le y\}$$

定义 对任意实数对 $(x, y) \in R^2$，称二元函数

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

为 (X, Y) 的**联合分布函数**.

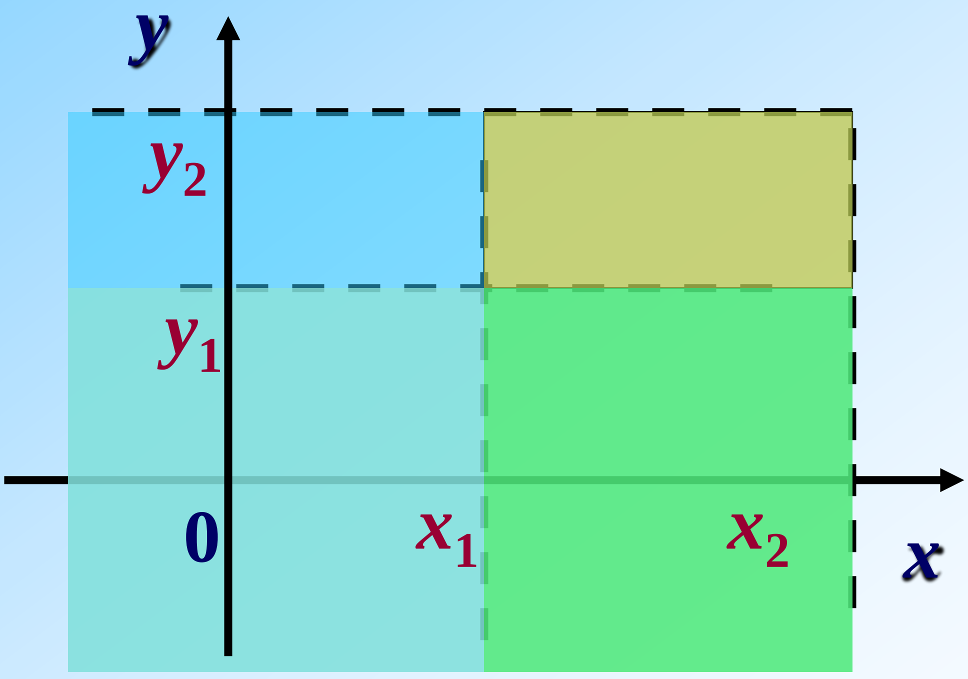
联合分布函数几何意义



联合分布

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

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



一维随机变量X、Y的分布函数$F_X(x)$和$F_Y(y)$称为(X, Y)的**边缘分布函数.**

由联合分布函数可确定边缘分布函数

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

$$= \lim_{y \to +\infty} F(x, y)$$

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

$$= \lim_{x \to +\infty} F(x, y)$$

思考 能否由边缘分布函数确定联合分布函数?

练习（X, Y）的联合分布函数为

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

$$F_X(x)=\begin{cases}1-e^{-2x} & x\ge 0\\ 0 & \text{其他}\end{cases}$$

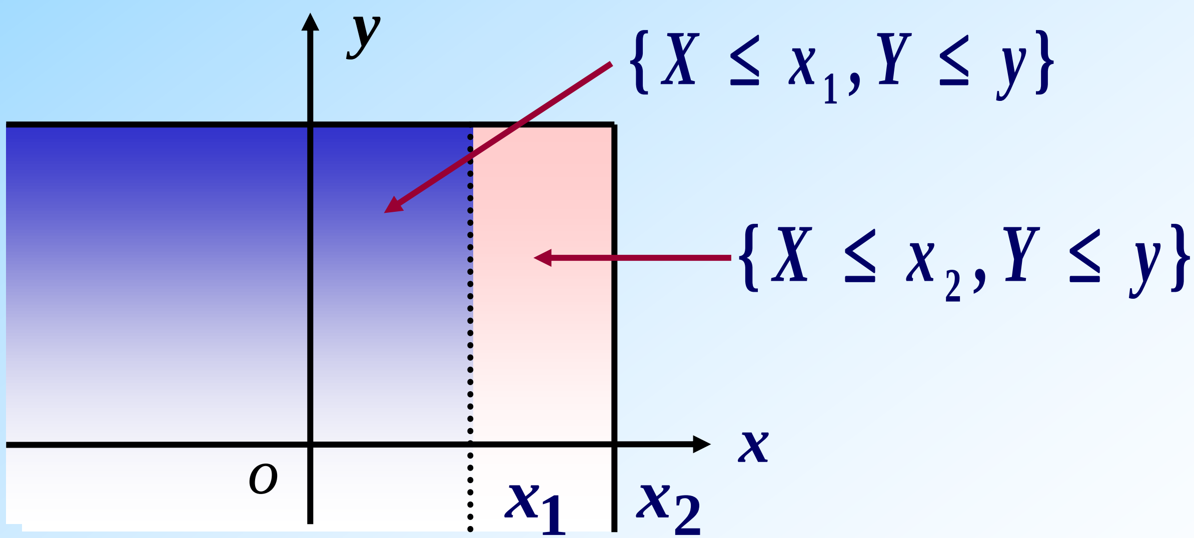
$$F_Y(y)=\begin{cases}1-e^{-3y} & y\ge 0\\ 0 & \text{其他}\end{cases}$$

联合分布函数的性质

1.单调不减性 $F(x,y)$**分别对**x , y**单调不减.**

当$x_1 < x_2, F(x_1, y) \le F(x_2, y), \quad \forall y \in R$;

当$y_1 < y_2, F(x, y_1) \le F(x, y_2), \quad \forall x \in R$



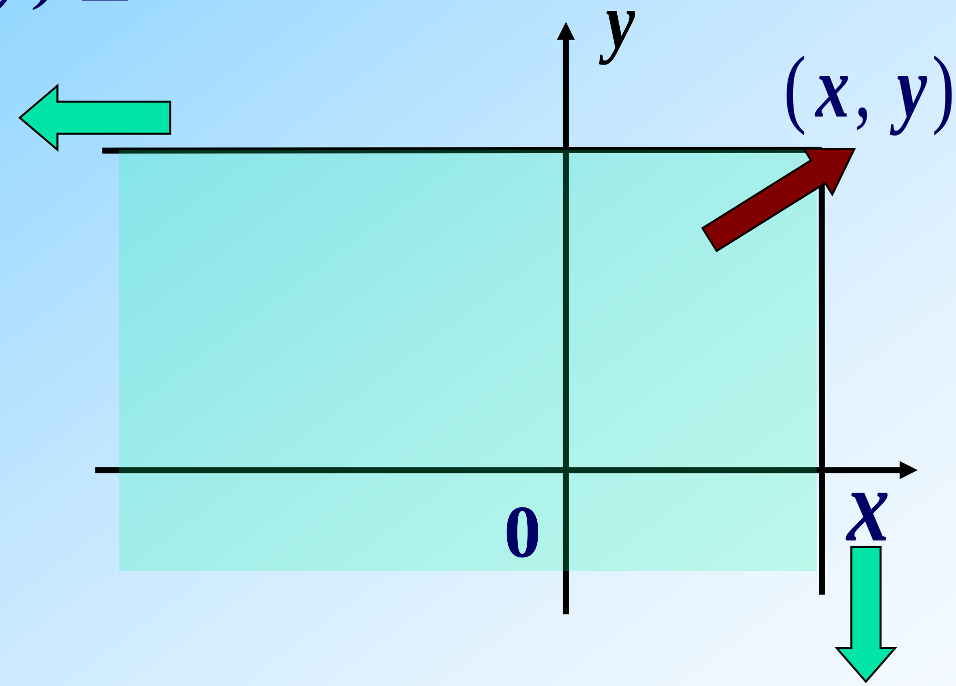
联合分布

2.有界性： $0 \le F(x, y) \le 1$

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

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

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



3.右连续性 $F(x, y)$ 分别关于x 或 y右连续.

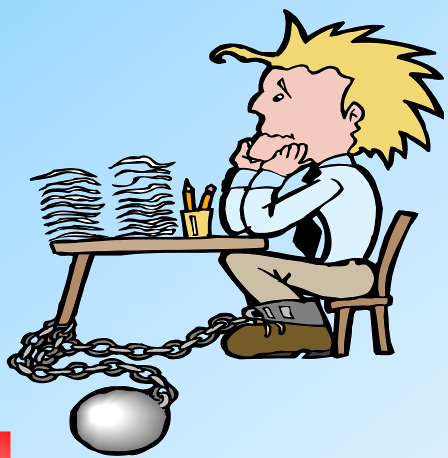
$$\lim_{x \to x_0^+} F(x, y) = F(x_0, y), \quad \lim_{y \to y_0^+} F(x, y) = F(x, y_0)$$

4.相容性：对任意 $x_1 < x_2, y_1 < y_2$，有

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

注 如果二元函数 $F(x, y)$ 满足上述4条性质，则必存在二维随机变量 (X, Y) 以 $F(x, y)$ 为分布函数.

参见教材 例3.1.1



思考 一维分布函数与二维分布函数的联系与区别?

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

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

$$(x_1, x_2, \ldots, x_n) \in R^n.$$

由$(X_1, X_2, \ldots, X_n)$的联合分布函数，可确定其中任意k 个分量的联合分布函数，称为**k维边缘分布函数.**

$$F_{X_1}(x_1) = F(x_1, +\infty, +\infty, \cdots, +\infty)$$

$$F_{X_1, X_2}(x_1, x_2) = F(x_1, x_2, +\infty, \cdots, +\infty)$$

二.联合分布律

定义 设二维随机变量(X,Y)至多取可列对数值：$(x_i, y_j),\ i,j=1,2,\dots.$

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

若 1) $p_{ij}\geq 0;\qquad i,j=1,2,\dots$

2) $\sum_i \sum_j p_{ij}=1.$

称(X,Y)为**二维离散型随机变量**，称式$(*)$为(X,Y)的联合分布律.

联合分布函数为

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

由联合分布律确定随机变量X , Y 的分布律

$$P\{X=x_i\}=\sum_{j=1}^{+\infty}p_{ij}=p_{i\cdot}\quad i=1,2,\dots$$

$$P\{Y=y_j\}=\sum_{i=1}^{+\infty}p_{ij}=p_{\cdot j}\quad j=1,2,\dots$$

用表格表示联合分布律和边缘分布律

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

例3.1.1　　例3.1.2

思考 能否用边缘分布律来确定联合分布律，原因是什么？

(参见教材 例3.1.3和例3.1.4)

多维随机变量的**联合分布**不仅与每个分量的**边缘分布**有关，而且还与每个**分量之间的联系**有关！

三.联合概率密度

定义 二维随机变量(X, Y)的联合分布函数为$F(x, y)$，如果存在非负的函数$f(x, y)$使得对任意实数对(x, y)，有

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

称(X, Y)是**连续型随机变量**，称$f(x, y)$为(X, Y)的**联合概率密度**.

密度性质

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

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

这两条可作为判断一个二元函数是否是联合概率密度的标准

3) 若$f(x,y)$在(x,y)处连续, 则

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

4) 若$G \subset R^2$,有

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

5) 关于X和Y的边缘概率密度为

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

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

带参变量的积分

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

$$f_X(x)=F'_X(x)=\int_{-\infty}^{+\infty} f(x,v)dv$$

例3.1.3 **例3.1.4** **例3.1.5**

求解边缘概率密度, 是求解含参变量积分问题.

难点:定积分的上下限.

解决方法:借助于图形

教材例3.1.7是通过讨论 $f(x,y)$的非零区域来求解.

四.二维均匀分布

设$G \subset R^2$,面积为 $S(G)$,若二维随机变量(X, Y)的联合概率密度为

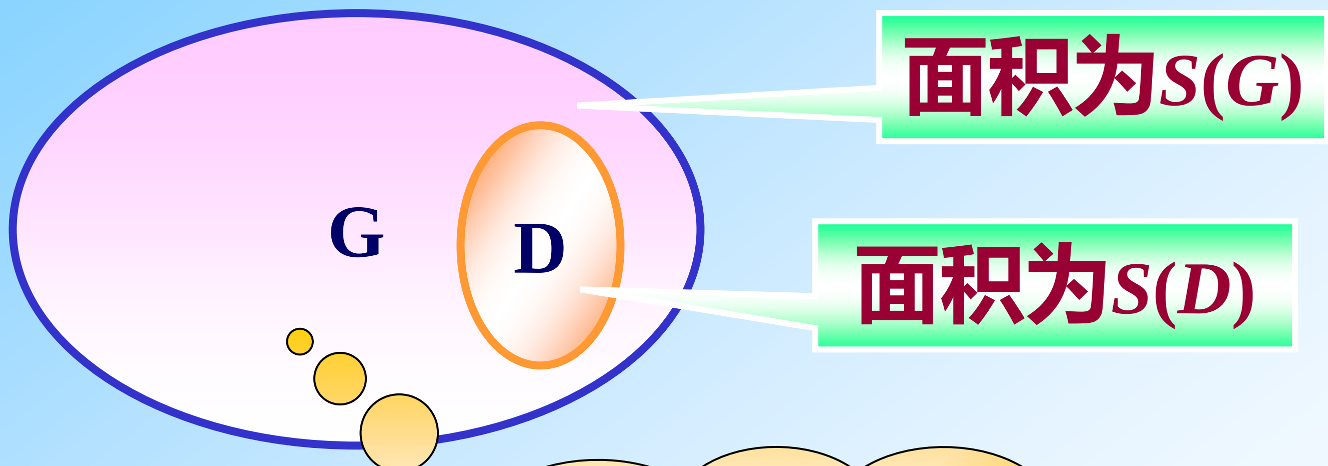
$$f(x,y)=\begin{cases} \dfrac{1}{S(G)} & (x,y)\in G \\ 0 & \text{其他} \end{cases}$$

则称(X, Y)在G上服从**均匀分布**.

1. (X, Y)在G上服从均匀分布,设$D \subset G$ 则有

联合分布

$$P\{(X,Y)\in D\}=\iint\limits_{D}\frac{1}{S(G)}d\sigma=\frac{S(D)}{S(G)}$$



概率值与区域 D 的形状、位置等均无关，只与 D 的面积有关。

2. 设$X \sim U(a,b)$，$(c,d) \subset (a,b)$ 则

$$P\{c < X \le d\} = \frac{d-c}{b-a} = \frac{(c,d)\text{的长度}}{(a,b)\text{的长度}}$$

借助于几何度量指标(长度, 面积, 体积等)计算概率, 可建立 “几何概型” .

例3.1.6　　例3.1.7

五.二维正态分布

定义 二维随机变量(X, Y)的联合概率密度为

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

其中$\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$均为常数,

且 $\sigma_1>0, \sigma_2>0, |\rho|<1,$

称(X, Y)服从**二维正态分布**，记为

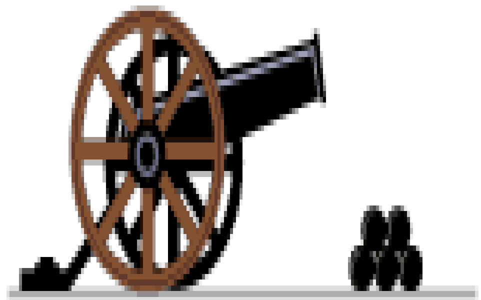
$$(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$$

命题3.1.1 若$(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$

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

见例3.1.10

例1 炮弹发射试验



炮弹在地面的命中点位置要由**两个随机变量**(X, Y)来确定.

例2 电子放大器的干扰电流由其 **振幅** 和 **相位** 两个随机变量来确定.

例3 飞机在空中的位置要由**三个随机变量** (X, Y, Z)来确定.

例3.1.1 在1,2,3,4 中随机取出一数 X ,再随机地从1~X 中取 一数Y， 求(X, Y)的联合分布律.

解 X 的分布律为

X	1	2	3	4
$P\{X=x\}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

$$p_{ij} = P\{X = i, Y = j\} = P(\{X = i\} \cap \{Y = j\})$$

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

联合分布

$$= \begin{cases} 0, & j > i; \\ \dfrac{1}{4}\dfrac{1}{i}, & j \le i. \end{cases} \qquad i\,,\ j = 1,\,2\,,\,3\,,\,4.$$

Y / X	1	2	3	4	
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
	25/48	13/48	7/48	3/48	1

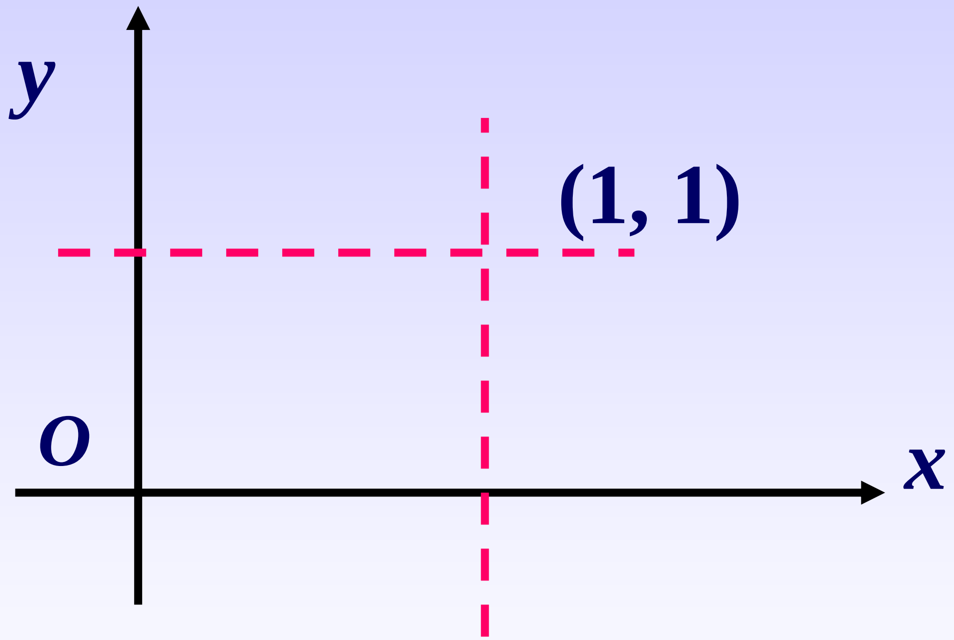
例3.1.2 (二维两点分布)

用剪刀随机的去剪一次悬挂有小球的绳子. 剪中的概率为p $(0< p<1)$.

设X 表示剪中绳子的次数; Y 表示小球下落的次数.求(X, Y)的联合分布函数.

X \ Y	0	1
0	$1-p$	0
1	0	p

称(X, Y)服从**二维两点分布**

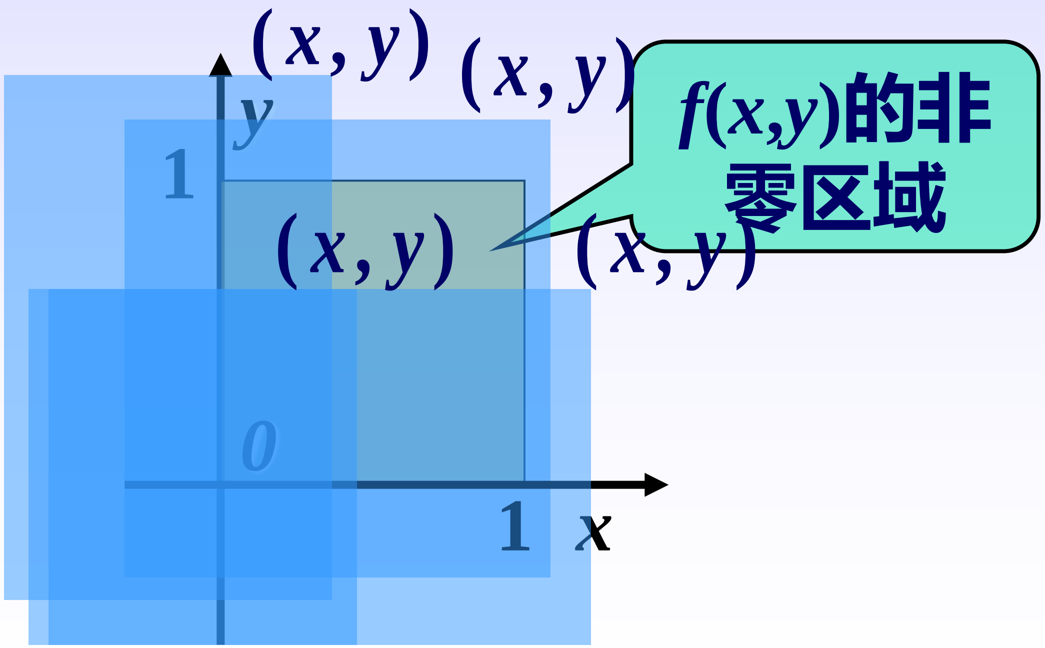


见教材例3.1.5.

例3.1.3 已知二维随机变量(X, Y)的联合概率密度为

$$f(x,y)=\begin{cases}4xy, & 0\le x\le 1,0\le y\le 1;\\ 0, & \text{其他}.\end{cases}$$

试写出(X, Y)的联合分布函数。



解　1) 当 $x \le 0$，或 $y \le 0$ 时

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

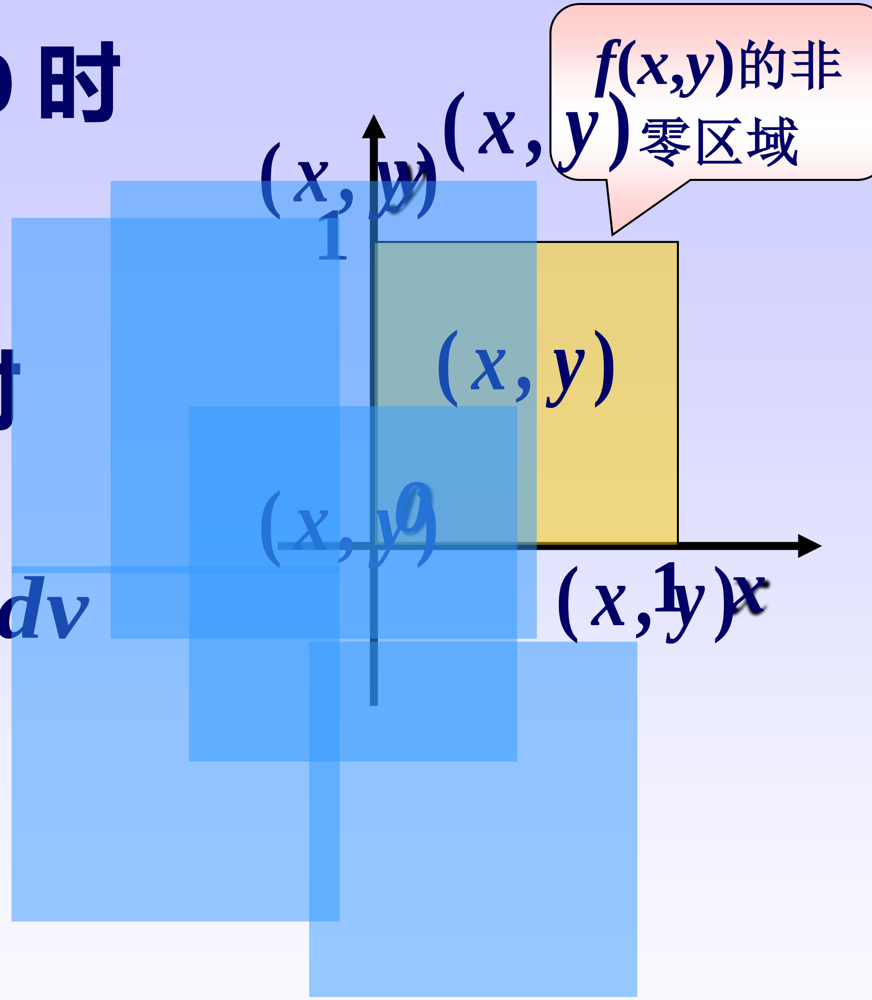
2) 当 $0 \le x < 1,\ 0 \le y < 1$ 时

$$F(x, y) = 4\int_0^x \int_0^y uv\, du dv$$

$$= x^2 y^2$$

3) 当 $0 \le x < 1,\ y \ge 1$ 时

$$F(x, y) = 4\int_0^x \int_0^1 uv\, du dv = x^2$$

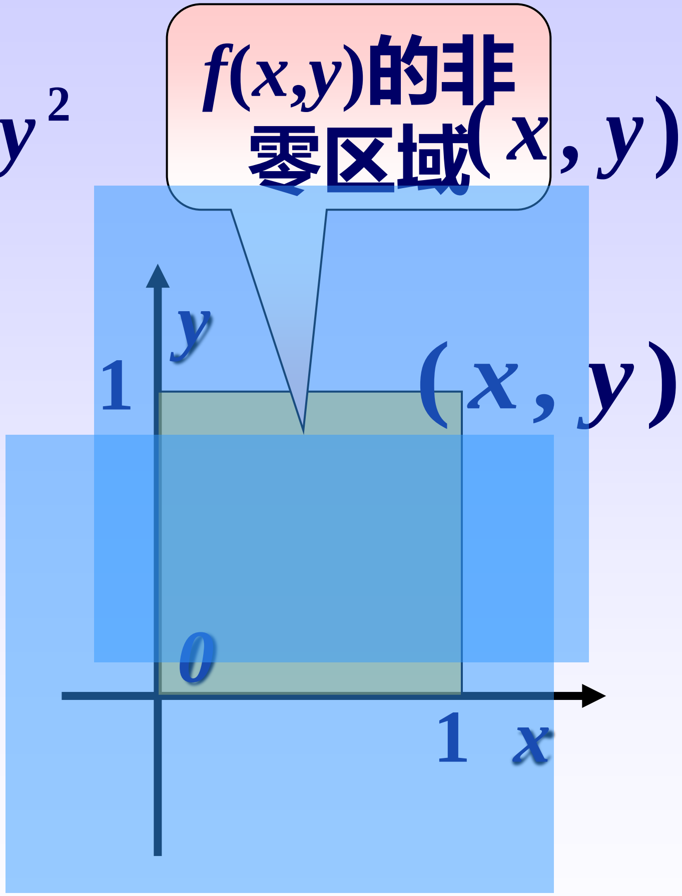


4）当 $0 \le y < 1$，$x \ge 1$ 时

$$F(x,y) = 4\int_0^1\int_0^y uv\,du\,dv = y^2$$

5）当 $1 \le y$，$x \ge 1$ 时

$$F(x,y) = 4\int_0^1\int_0^1 uv\,du\,dv = 1.$$



综上所叙得

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

例3.1.4 (教材例3.1.7)已知二维随机变量(X,Y)的联合概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{2x^2y}, & 1\le x<+\infty,\ \dfrac{1}{x}\le y\le x;\\ 0, & \text{其他}.\end{cases}$$

求关于Y的边缘概率密度$f_Y(y)$.

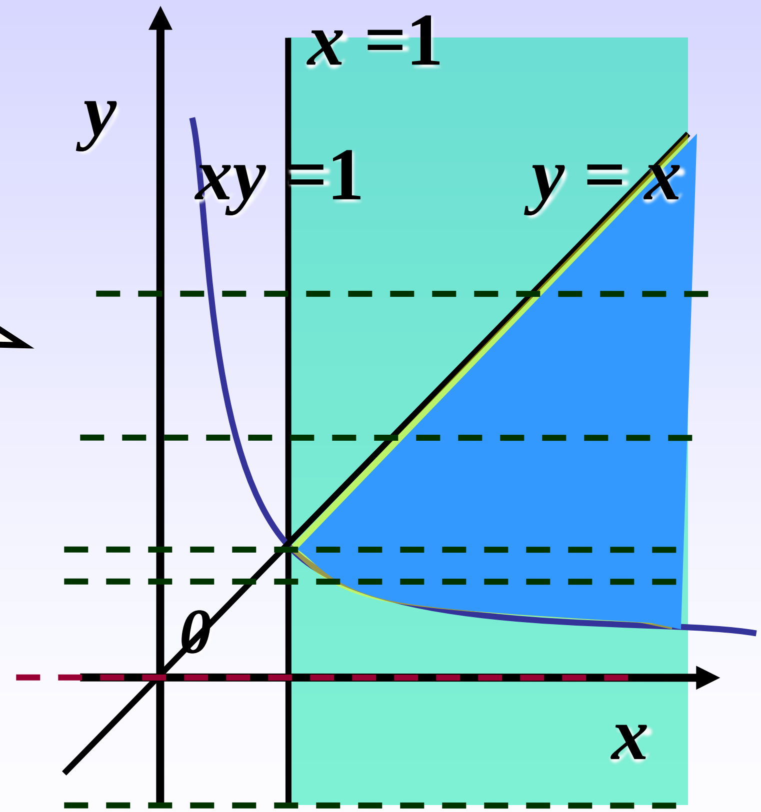
分析 $f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\,dx$

求Y的边缘概率密度, 需固定y对x求积分.

实质上是求含参变量的积分.

$$f(x,y)=\begin{cases}\dfrac{1}{2x^2y}, & 1\le x<+\infty,\ \dfrac{1}{x}\le y\le x;\\ 0, & \text{其他}.\end{cases}$$

对 y 的不同值，$f_Y(y)$ 的积分上下限不相同.

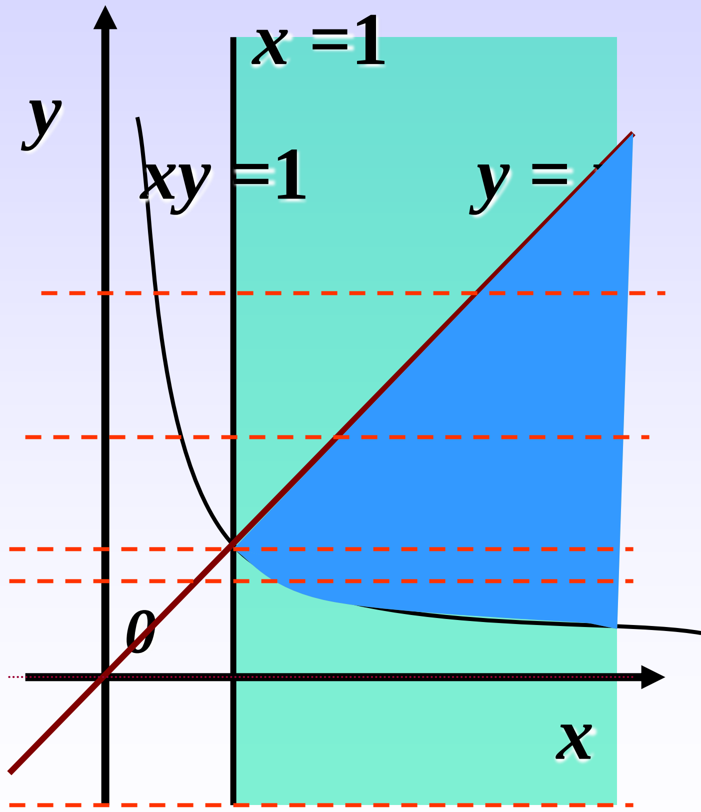


联合分布

$$f(x,y)=\begin{cases}\dfrac{1}{2x^2y}, & 1\le x<+\infty,\ \dfrac{1}{x}\le y\le x;\\ 0, & \text{其他}.\end{cases}$$

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

$$=\begin{cases}0, & y\le 0;\\ \displaystyle\int_{\frac{1}{y}}^{+\infty}\frac{1}{2x^2y}dx, & 0<y\le 1;\\ \displaystyle\int_{y}^{+\infty}\frac{1}{2x^2y}dx, & 1<y.\end{cases}$$



$$= \begin{cases} 0, & y \le 0; \\ \dfrac{1}{2}, & 0 < y \le 1; \\ \dfrac{1}{2y^2}, & 1 < y. \end{cases}$$

例3.1.5 已知二维随机变量(X, Y)的联合概率密度为

$$f(x,y)=\begin{cases} a(3x^2+xy), & 0\le x\le 1, 0\le y\le 2; \\ 0, & \text{其他}. \end{cases}$$

求 1）a，　　2）边缘概率密度.

3）$P\{X+Y>1\}$.

分析 1）利用性质 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dxdy=1$

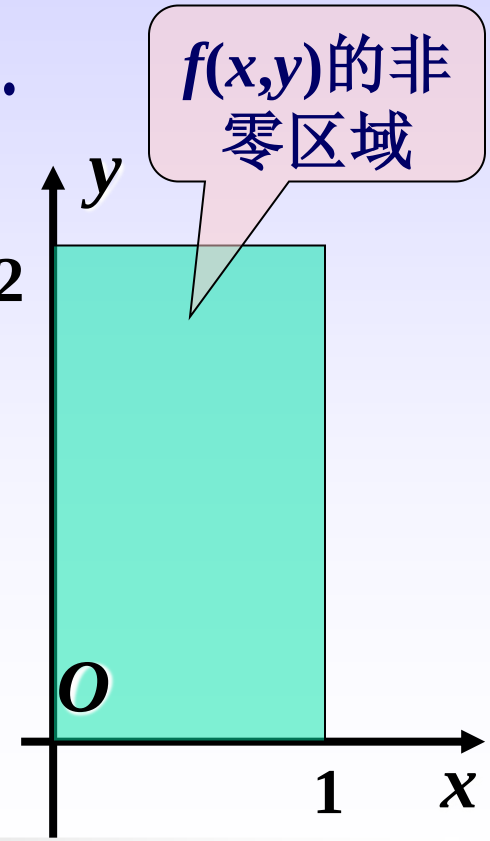
2）利用含参变量积分；

3）利用联合概率密度的性质：

若 $G \subset R^2$，有

$$p\{(x,y)\in G\} = \iint_G f(x,y)\,d\sigma$$

$$f(x,y) = \begin{cases} a(3x^2 + xy), & 0 \le x \le 1,\ 0 \le y \le 2; \\ 0, & \text{其他}. \end{cases}$$



解 1) 由 $\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} f(x,y)\,dx\right]dy = 1$

$$\Rightarrow \int_0^2\left[\int_0^1 a(3x^2 + xy)\,dx\right]dy = 1$$

$$\therefore \quad a = \frac{1}{3}.$$

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

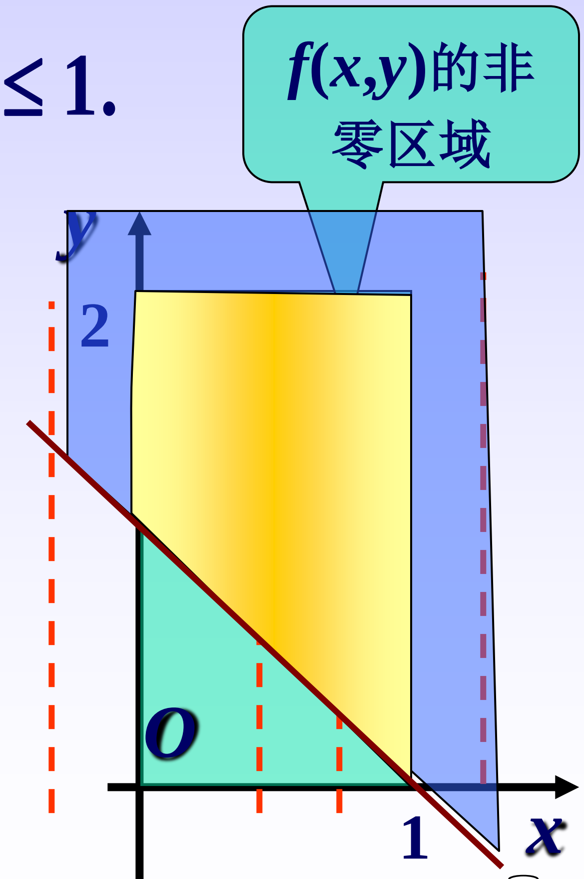
$$=\begin{cases} 0. & x<0\text{或}x>1; \\ 2x^2+\frac{2}{3}x. & 0\le x\le 1. \end{cases}$$

3) $P\{X+Y>1\}$

$$=\iint\limits_{x+y>1} f(x,y)dxdy$$

$$=\int_0^1[\int_{1-x}^2(x^2+\frac{1}{3}xy)dy]dx$$

$$=\frac{65}{72}.$$



例3.1.6（约会问题） 甲乙两艘轮船驶向一个不能同时停泊两艘轮船的码头停泊,它们在一昼夜内到达的时刻是等可能的.

如果甲的停泊时间为1小时，乙的停泊时间为二小时.求它们中的任意一艘都不须等待码头空出的概率.

分析 设甲在一昼夜到达的时刻为X,乙在一昼夜到达的时刻为Y. 所求的概率为

$$P\{X+1\le Y \text{或} Y+2\le X\}$$

设 $G=\{(x,y):0\le x,y\le 24\}$

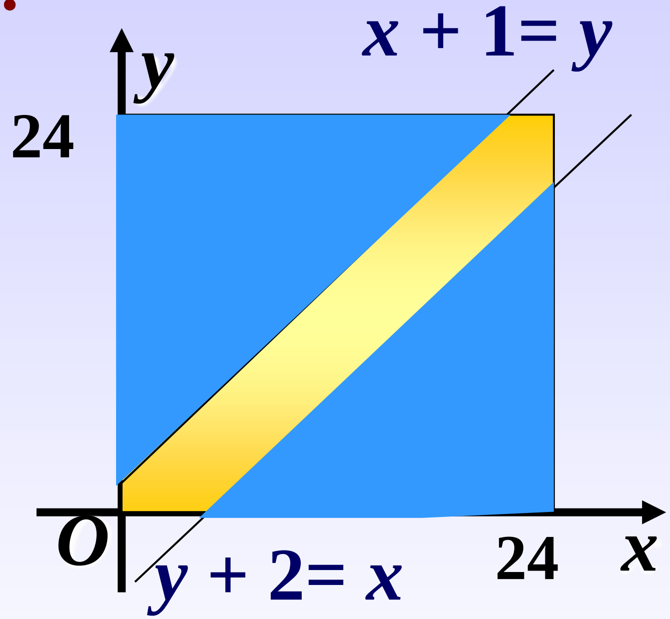
则 (*X*, *Y*)在G上服从**均匀分布**.

$P\{X+1\le Y \text{ 或 } Y+2\le X\}$

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

$$=\frac{S(D)}{S(G)}=\frac{0.5\times 23^2+0.5\times 22^2}{24^2}$$

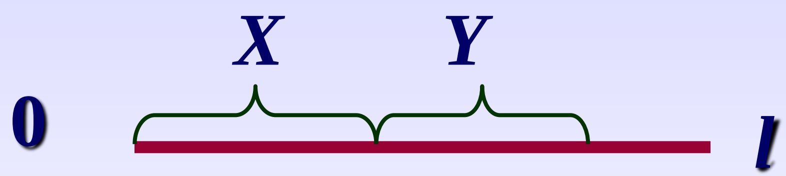
≈ 0.88



例3.1.7 把长为 l 的木棒，任意折成3段，求它们能构成一个三角形的概率.

分析

1) 可设第一段的长度为 X，第二段的长度为Y



有 $0 \le X \le l, 0 \le Y \le l, \quad X + Y \le l$

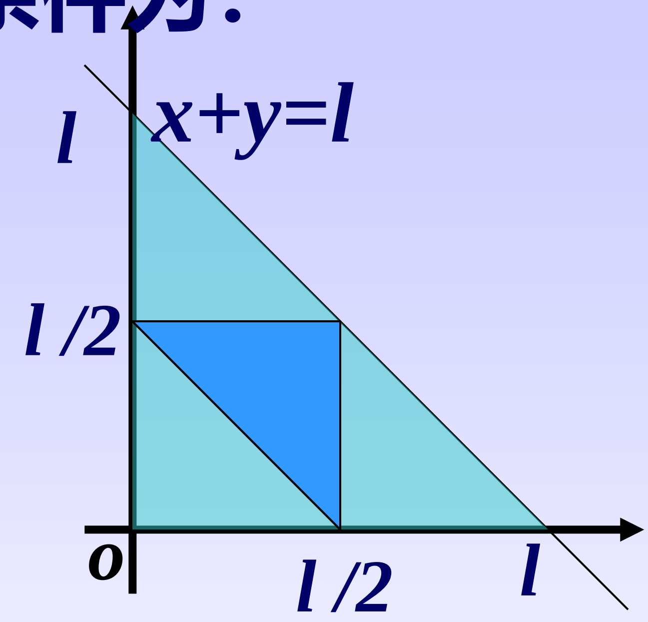
2) (X, Y) 在三角形

$$G = \{(x, y) \mid 0 \le x \le l, 0 \le y \le l, x + y \le l\}$$

上服从均匀分布.

3) 能构成三角形的充要条件为:

$$\begin{cases} 0 < x < \dfrac{l}{2} \\ 0 < y < \dfrac{l}{2} \\ x + y > \dfrac{l}{2} \end{cases}$$



解 设第一段的长度为 X，第二段的长度为Y，
(X,Y) 在 三 角 形

$$G = \{(x,y) \mid 0 \le x \le l, 0 \le y \le l, x + y \le l\}$$

上服从均匀分布.

所求概率为

$$p = P\{X < l/2, Y < l/2, X + Y < l/2\}$$

$$= \frac{\frac{1}{2}\left(\frac{l}{2}\right)^2}{\frac{1}{2}l^2} = \frac{1}{4}.$$