第三章　多维随机变量及其分布

很多随机现象只用一个随机变量来描述是不够的，需要用几个随机变量同时来描述．如：

- 平面上一点的位置需要用两个坐标来表示；
- 天气通常由最高、最低气温，相对湿度，风力，降水量等因素决定；
- 钢材的质量有含碳量、含硫量和硬度等基本指标．

定义． 设 Ω 是某随机试验的样本空间，$X_1, X_2, \cdots, X_n$ 是该空间上的随机变量，称

$$X = (X_1, X_2, \cdots, X_n)$$

为 Ω 上的 n 维随机向量或 n 维随机变量，称 n 元函数

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

为 X 的联合分布函数，其中

$$x = (x_1, x_2, \cdots, x_n).$$

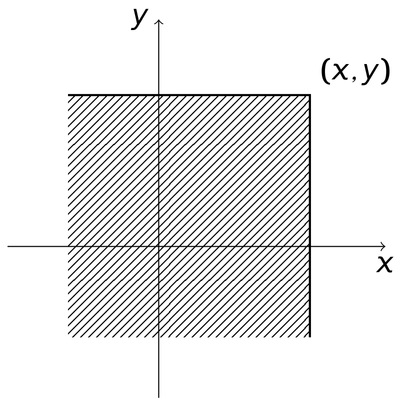
3.1　二维随机变量

定义 1. 设 (X, Y) 为定义在样本空间 Ω 上的两个随机变量，则称 (X, Y) 为二维随机向量或者二维随机变量，称二元函数

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

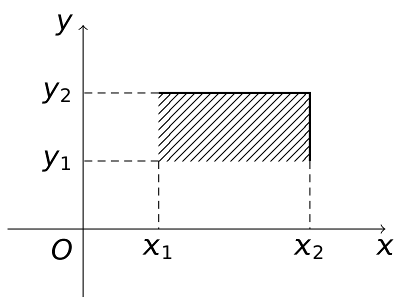
为 (X, Y) 的联合分布函数．

注记． 联合分布函数 $F(x, y)$ 表示事件 $\{X \le x\}$ 和事件 $\{Y \le y\}$ 同时发生的概率．



如果将二维随机变量 (X,Y) 看成是平面上随机点的坐标, 那么联合分布函数 $F(x,y)$ 在 (x,y) 处的函数值就是随机点 (X,Y) 落在以 (x,y) 为右上角的无穷矩形内的概率.

随机变量 (X,Y) 落在矩形区域 $(x_1,x_2]\times(y_1,y_2]$ 内的概率为



$$\begin{aligned}&P\{x_1<X\le x_2,y_1<Y\le y_2\}\\ &\qquad=F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1).\end{aligned}$$

任一二维联合分布函数 $F(x,y)$ 必具有如下四条基本性质:

1. 单调性　$F(x,y)$ 分别对 x 或 y 是单调不减的, 即

    - 当 $x_1<x_2$ 时, 有 $F(x_1,y)\le F(x_2,y)$.
    - 当 $y_1<y_2$ 时, 有 $F(x,y_1)\le F(x,y_2)$.

2. 有界性 对任意的 x 和 y, 有 $0 \le F(x,y) \le 1$, 且

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

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

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

3. 右连续性 对每个变量都是右连续的, 即

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

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

4. 非负性 对任意的 $x_1 < x_2$, $y_1 < y_2$ 有

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

定义 **2.** 如果二维随机变量 (X,Y) 的所有可能取值 (x_i,y_j) 只有有限对或者可列无限对，则称 (X,Y) 为离散型二维随机变量, 称

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

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} 满足下列条件*:*

1. $p_{ij} \ge 0, \quad i = 1,2,\cdots, j = 1,2,\cdots$

2. $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} p_{ij} = 1$；

二维离散型随机变量的联合分布也可以用表格表示：

$X \backslash Y$	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}	$\cdots$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$
$\ldots$	$\ldots$	$\ldots$		$\ldots$	

它们的联合分布函数可以由下面子式求出:

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

例 **1.** 一箱子装有 5 件产品，其中 2 件正品，3 件次品. 每次从中取 1 件产品检验质量，不放回地抽取, 连续两次. 定义随机变量 X 和 Y 如下:

$$X=\begin{cases}1, & \text{第一次取到次品}\\ 0, & \text{第一次取到正品}\end{cases}$$

$$Y=\begin{cases}1, & \text{第二次取到次品}\\ 0, & \text{第二次取到正品}\end{cases}$$

试求 (X,Y) 的分布律.

解**.** (X,Y) 可能取的值只有 4 对: $(0,0)$,$(0,1)$,$(1,0)$ 及 $(1,1)$, 按概率的乘法公式计算得

$$\begin{aligned}P\{X=0,Y=0\}&=P\{X=0\}P\{Y=0\mid X=0\}\\&=\frac{2}{5}\times\frac{1}{4}=0.1\\P\{X=0,Y=1\}&=\frac{2}{5}\times\frac{3}{4}=0.3\\P\{X=1,Y=0\}&=\frac{3}{5}\times\frac{2}{4}=0.3\\P\{X=1,Y=1\}&=\frac{3}{5}\times\frac{2}{4}=0.3\end{aligned}$$

故 (X,Y) 的分布律用表格表示为:

X \ Y	0	1
0	0.1	0.3
1	0.3	0.3

定义 **3.** 设二维随机变量 (X,Y) 的分布函数是 $F(X,Y)$, 如果存在一个非负函数 $f(x,y)$，使得对于任意的 x,y 有

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

则称 (X,Y) 为二维连续型随机变量．函数 $f(x,y)$ 称为 (X,Y) 的联合概率密度．

联合概率密度函数的基本性质：

1. $f(x,y)\geq 0, \forall x,y$；
2. $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1$；
3. 若函数 f 在点 (x,y) 处连续，则有
$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y).$$
4. 对任意的平面区域 D，
$$P\{(X,Y)\in D\} = \iint\limits_{(x,y)\in D} f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

特别地，
$$P\{a<X\leq b, c<Y\leq d\} = \int_a^b\int_c^d f(x,y)\,\mathrm{d}y\,\mathrm{d}x.$$

例 **2.** 设 G 是平面上的一个有界区域, 其面积为 A 二维随机变量 (x,y) 只在 G 中取值, 并且取 G 中的每一个点都是 ‘‘等可能的”，即 (x,y) 的概率密度为
$$f(x,y) = \begin{cases} C, & (x,y)\in G \\ 0, & 其它 \end{cases}$$
由概率的性质 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1$ 可得 $C=\frac{1}{A}$
$$故有 \qquad f(x,y) = \begin{cases} \frac{1}{A}, & (x,y)\in G \\ 0, & 其它 \end{cases}$$
如果一个二维随机变量 (X,Y) 以上式为概率密度, 则称 (X,Y) 服从区域 G 上的均匀分布.

例 **3.** 设 (X,Y) 的联合概率密度为
$$f(x,y) = \begin{cases} Cx^2y^2, & 0\leq x\leq 1, 0\leq y\leq 1 \\ 0, & otherwise \end{cases}$$
其中 C 是常数．

(1) 求常数 C；$\cdots\cdots\cdots$ 9

(2) 计算 $P\left\{X<\frac{1}{3},Y<\frac{1}{3}\right\}$. $\cdots\cdots\cdots$ $\frac{1}{729}$

例 **4.** 设二维随机变量 (X,Y) 具有概率密度

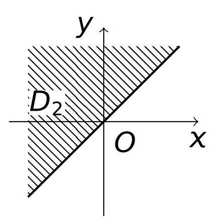
$$f(x,y)=\begin{cases} 2\mathrm{e}^{-(2x+y)}, & x>0,y>0 \\ 0, & \text{其他} \end{cases}$$

试求：*(1)* 分布函数 $F(x,y)$ *(2)* $P\{X<Y\}$

解**.** *(1)* 由条件易知

$$\begin{aligned} F(x,y) &= \int_{-\infty}^{y}\int_{-\infty}^{x} f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \begin{cases} \int_0^y\int_0^x 2\mathrm{e}^{-(2x+y)}\,\mathrm{d}x\,\mathrm{d}y, & x>0,y>0 \\ 0 & ,\text{其他}. \end{cases} \\ &= \begin{cases} \left(1-\mathrm{e}^{-2x}\right)\left(1-\mathrm{e}^{-y}\right), & x>0,y>0 \\ 0 & ,\text{其他}. \end{cases} \end{aligned}$$

把位于 XOY 平面的直线 $x=y$ 上方的区域记为 G



于是

$$\begin{aligned} P\{X<Y\} = P\{(x,y)\in G\} &= \iint_G f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_0^{+\infty}\int_x^{+\infty} 2\mathrm{e}^{-(2x+y)}\,\mathrm{d}y\,\mathrm{d}x = \frac{2}{3}. \end{aligned}$$

定义**.** 设 Ω 是某随机试验的样本空间，$X_1,X_2,\cdots,X_n$ 是该空间上的随机变量，称

$$(X_1,X_2,\cdots,X_n)$$

为 Ω 上的 n 维随机向量或 n 维随机变量，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 n 维随机变量 $(X_1,X_2,\cdots,X_n)$ 的分布函数或联合分布函数.

3.2 边缘分布

二维随机变量 (X,Y) 作为一个整体，有联合分布函数 $F(x,y)$，其分量 X 与 Y 都是随机变量，有各自的分布函数，分别记成 $F_X(x)$ 和 $F_Y(y)$，称为 X 和 Y 的边缘分布函数. 边缘分布由联合分布完全确定：

$$F_X(x)=F(x,+\infty),\qquad F_Y(y)=F(+\infty,y).$$

设二维离散型随机变量 (X,Y) 的联合分布律为

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

则随机变量 X 的边缘分布律为

$$p_{i\cdot}=P\{X=x_i\}=\sum_j p_{ij},\qquad i=1,2,\cdots$$

而随机变量 Y 的边缘分布律为

$$p_{\cdot j}=P\{Y=y_j\}=\sum_i p_{ij},\qquad j=1,2,\cdots$$

通常将这两个分布分别写在联合分布表右边和下边.

设二维连续型随机变量 (X,Y) 的联合概率密度为 $f(x,y)$，则 X，Y 的边缘概率密度分别定义为

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}y,\quad f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}x.$$

此时 X，Y 的边缘分布函数分别为

$$F_X(x)=\int_{-\infty}^{x}f_X(s)\,\mathrm{d}s,\quad F_Y(y)=\int_{-\infty}^{y}f_Y(t)\,\mathrm{d}t.$$

若 (X,Y) 服从矩形区域

$$D=\{(x,y):a\le x\le b,c\le y\le d\}$$

上的均匀分布，则边缘概率密度分别为

$$f_X(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{else} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{d-c} & y \in [c, d] \\ 0 & \text{else} \end{cases}$$

这表明：X 与 Y 都服从均匀分布．该结论对其他非矩形区域上的均匀分布一般不成立．

例 **1.** 把两封信随机地投入已经编好号的 *3* 个邮筒内, 设 X, Y 分别表示投入第 1,2 个邮筒内信的数目, 求 (X, Y) 的分布律及边缘分布律.

解**.** X, Y 各自的取值为 0, 1, 2. 由题设, (X, Y) 取 $(1, 2)$, $(2, 1)$, $(2, 2)$ 均不可能，因而相应的概率均为 0 再由古典概率计算得：

$$P\{X = 0, Y = 0\} = \frac{1}{3^2} = \frac{1}{9},\ P\{X = 0, Y = 1\} = \frac{2}{3^2} = \frac{2}{9},$$
$$P\{X = 0, Y = 2\} = \frac{1}{3^2} = \frac{1}{9},\ P\{X = 1, Y = 1\} = \frac{2}{3^2} = \frac{2}{9},$$

$P\{X = 1, Y = 0\}, P\{X = 2, Y = 0\}$ 可由对称性求得.

所有计算结果列表如下:

$X \backslash Y$	0	1	2	$p_{\cdot j}$
0	1/9	2/9	1/9	4/9
1	2/9	2/9	0	4/9
2	1/9	0	0	1/9
$p_{i\cdot}$	4/9	4/9	1/9	

X 和 Y 的边缘分布律可由 (X, Y) 的分布律确定.

例 **2.** 将 *2* 只红球和 *2* 只白球随机地投入已经编好号的 *3* 个盒子中去, 设 X 表示落入第 *1 1* 盒子内红球的数目, Y 表示落入第 *2* 个盒子内白球的数目, 求 (X, Y) 的分布律及边缘分布律.

$X \backslash Y$	0	1	2	$p_{\cdot j}$
0	16/81	16/81	4/81	4/9
1	16/81	16/81	4/81	4/9
2	4/81	4/81	1/81	1/9
$p_{i\cdot}$	4/9	4/9	1/9	

解. 不妨分别把 *2* 只红球和 *2* 只白球看作是有差别的 *(*例如编号），由古典概型计算得

$$P\{X=1,Y=1\}=\frac{\begin{pmatrix}2\\1\end{pmatrix}\cdot 2\cdot\begin{pmatrix}2\\1\end{pmatrix}\cdot 2}{3^4}=\frac{16}{81}$$

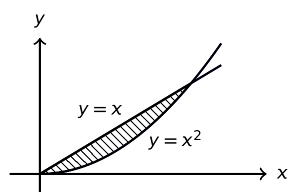
类似地计算出下表内的其他结果：

比较一下例 1 的表和例 2 的表，立即可以发现，两者有完全相同的边缘分布，而联合分布却是不相同的．由此可知，由边缘分布并不能唯一地确定联合分布．

例 **3.** 设二维随机变量 (X,Y) 在区域

$$G=\left\{(x,y)\mid 0\le x\le 1,x^2\le y\le x\right\}$$

上服从均匀分布, 求边缘概率密度 $f_X(x)$, $f_Y(y)$.



解. 不难得到 (X,Y) 的概率密度

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

则

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y=\begin{cases}6\left(x-x^2\right), & 0\le x\le 1,\\ 0, & \text{其他.}\end{cases}$$

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

虽然 (X,Y) 的联合分布在 G 上服从均匀分布, 但它们的边缘分布却不是均匀分布

3.3 条件分布

设二维离散型随机变量 (X,Y) 的联合分布为

$$P\{X=x_i,Y=y_j\}=p_{ij},$$

$$i=1,2,\cdots,j=1,2,\cdots$$

则 X 与 Y 的边缘分布分别为

$$p_{i\cdot}=P\{X=x_i\}=\sum_j p_{ij},\qquad i=1,2,\cdots$$

$$p_{\cdot j}=P\{Y=y_j\}=\sum_i p_{ij},\qquad j=1,2,\cdots$$

定义 **1.** 当 $p_{\cdot j}>0$ 时，称

$$P\{X=x_i|Y=y_j\}=\frac{P\{X=x_i,Y=y_i\}}{P\{Y=y_i\}}=\frac{p_{ij}}{p_{\cdot j}},\ i=1,2,\cdots$$

为 $Y=y_j$ 时 X 的条件分布律. 当 $p_{i\cdot}>0$ 时，称

$$P\{Y=y_j|X=x_i\}=\frac{P\{X=x_i,Y=y_i\}}{P\{X=x_i\}}=\frac{p_{ij}}{p_{i\cdot}},\ j=1,2,\cdots$$

为 $X=x_i$ 时 Y 的条件分布律.

不难验证以上两式均满足分布律的基本性质.

例 **1.** 把两封信随机地投入已经编好号的 *3* 个邮筒内, 设 X, Y 分别表示投入第 *1,2* 个邮筒内信的数目, 求在 $Y=0$ 条件下随机变量 X 的条件分布律.

解. 在 $Y=0$ 的条件下 X 的条件分布律为

$$P\{X=0\mid Y=0\}=\frac{P\{X=0,Y=0\}}{P\{Y=0\}}=\frac{1/9}{4/9}=\frac{1}{4},$$
$$P\{X=1\mid Y=0\}=\frac{P\{X=1,Y=0\}}{P\{Y=0\}}=\frac{2/9}{4/9}=\frac{1}{2},$$
$$P\{X=2\mid Y=0\}=\frac{P\{X=2,Y=0\}}{P\{Y=0\}}=\frac{1/9}{4/9}=\frac{1}{4}.$$

定义. 在 $X=x$ 条件下 Y 的条件分布函数定义为

$$\begin{aligned}F_{Y|X}(y|x)&=P\{Y\le y|X=x\}\\&=\lim_{h\to 0}P\{Y\le y|x\le X\le x+h\}\end{aligned}$$

定理. 若 $f(\cdot,\cdot)$ 在点 (x,y) 处连续，$f_X(\cdot)$ 在点 x 处连续，且 $f_X(x)>0$，则

$$F_{Y|X}(y|x)=\int_{-\infty}^{y}f_{Y|X}(t|x)\,\mathrm{d}t.$$

练习 **1.** 设 (X,Y) 的联合概率密度为

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

求 $Y=y$ 时 X 的条件概率密度，及 $P\left\{X\le\frac{1}{2}\middle|Y=\frac{1}{2}\right\}$.

3.4 随机变量的独立性

定义. 设随机变量 (X,Y) 的联合分布函数为 $F(x,y)$，边缘分布函数分别为 $F_X(x)$ 和 $F_Y(y)$，若对任意实数 x，y 有

$$F(x,y)=F_X(x)\cdot F_Y(y),$$

则称 X,Y 相互独立.

注记. X,Y 相互独立即是指对任意实数 x,y，事件 $\{X\le x\}$ 与 $\{Y\le y\}$ 相互独立, 也即

$$P\{X\le x,Y\le y\}=P\{X\le x\}P\{Y\le y\}.$$

$X \backslash Y$	1	2	3
1	1/6	1/9	1/18
2	1/3	α	β

设 (X,Y) 是离散型随机变量，则 X 和 Y 相互独立的充裕条件是：对于 (X,Y) 的所有可能取值 (x_i,y_j), 有

$$P\{X=x_i,Y=y_i\}=P\{X=x_i\}P\{Y=y_i\},$$

即

$$p_{ij}=p_{i\cdot}\cdot p_{\cdot j},\quad i,j=1,2,\cdots$$

…………………………………………………………………………

设 (X,Y) 是连续型随机变量，$f(x,y),f_X(x),f_y(y)$ 分别为 (X,Y) 的概率密度和边缘概率密度，则 X 和 Y 相互独立的充要条件是：

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

在 $f(x,y),f_x(x),f_Y(y)$ 的一切公共连续点上成立.

例 **1.** 设二维离散型随机变量 (X,Y) 的分布律如下表所示：当 α,β 取何值时，X 和 Y 相互独立？

解**.** X,Y 的边缘分布律分别为

X	*1*	*2*
p_k	$\frac{1}{3}$	$\frac{1}{3}+\alpha+\beta$

Y_k	*1*	*2*	*3*
p_k	$\frac{1}{2}$	$\frac{1}{9}+\alpha$	$\frac{1}{18}+\beta$

若 X 和 Y 相互独立, 则有

$$\frac{1}{9}=P\{X=1,Y=2\}=P\{X=1\}P\{Y=2\}=\frac{1}{3}\times\left(\frac{1}{9}+\alpha\right),$$
$$\frac{1}{18}=P\{X=1,Y=3\}=P\{X=1\}P\{Y=3\}=\frac{1}{3}\times\left(\frac{1}{18}+\beta\right).$$

解得 $\alpha=\frac{2}{9},\beta=\frac{1}{9}$, 此时对所有的 x_i,y_j, $p_{ij}=p_{i\cdot}p_{\cdot j}$ 均成立，即 X,Y 相互独立.

例 **2.** 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是相互独立的, 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.

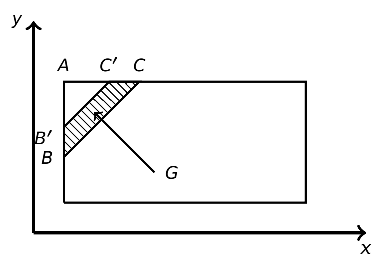
解**.** 设 X 和 Y 分别是负责人和他的秘书到达办公室的时间, 由题设 X 和 Y 的概率密度分别为

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

因为 X,Y 相互独立, 故 (X,Y) 的概率密度为

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

按题意, 要求的概率为 $P\left\{|X-Y|\leqslant\frac{1}{12}\right\}$, 画出区域: $\{(x,y)\big||x-y|\leqslant\frac{1}{12}\}$, 以及长方形区域 $\{(x,y)\mid 8<x<12; 7<y<9\}$.



它们的公共部分是四边形 $B'C'CB$, 记为 G 显然, 仅当 (X,Y) 在 G 内取值, 他们两人到达的时间相差才不超过 $\dfrac{1}{12}$ 小时.

因此, 所求的概率为

$$P\left\{|X-Y|\leqslant\frac{1}{12}\right\}=\iint_G f(x,y)\,\mathrm{d}x\,\mathrm{d}y=\frac{1}{8}\times(G\text{的面积})$$

而

$$\begin{aligned}G\text{的面积}&=\triangle ABC\text{的面积}-\triangle AB'C'\text{的面积}\\&=\frac{1}{2}\left(\frac{13}{12}\right)^2-\frac{1}{2}\left(\frac{11}{12}\right)^2=\frac{1}{6},\end{aligned}$$

于是,

$$P\left\{|X-Y|\leqslant\frac{1}{12}\right\}=\frac{1}{48}.$$

即负责人和他的秘书到达办公室的时间相差不超过 5 分钟的概率为 $\frac{1}{48}$.

例 **3.** 若 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}8xy, & 0\leqslant x\leqslant y\leqslant 1;\\ 0, & \text{其他}.\end{cases}$$

问 X 与 Y 是否相互独立?

解**.** 为判断 X 与 Y 是否独立, 只需看边缘概率密度函数的乘积是否等于联合概率密度函数. 为此先求边缘概率密度函数.

当 $x<0$ 或 $x>1$ 时, $f_X(x)=0$. 而当 $0\leqslant x\leqslant 1$ 时, 有

$$f_X(x)=\int_x^1 8xy\,\mathrm{d}y=8x\left(\frac{1}{2}-\frac{x^2}{2}\right)=4x\left(1-x^2\right).$$

解**.** 因此

$$f_X(x)=\begin{cases}4x\left(1-x^2\right), & 0\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

同样, 当 $y<0$ 或 $y>1$ 时, $f_Y(y)=0$. 而当 $0\leqslant y\leqslant 1$ 时, 有

$$f_Y(y)=\int_0^x 8xy\,\mathrm{d}x=4y^3.$$

因此

$$f_Y(y)=\begin{cases}4y^3, & 0\leqslant y\leqslant 1;\\ 0, & \text{其他}.\end{cases}$$

由此得 $f(x,y)\neq f_X(x)f_Y(y)$, 所以 X 与 Y 不独立.

练习 **1.** 设 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}9x^2y^2, & 0\leq x\leq 1, 0\leq y\leq 1\\ 0, & \text{其它}\end{cases}$$

判断 X 与 Y 是否相互独立.

定理 **1.** 设 X 和 Y 是相互独立的随机变量, $h(x)$ 和 $g(y)$ 是 $(-\infty,\infty)$ 上的连续函数, 则 $h(X)$ 和 $g(Y)$ 也是相互独立的随机变量.

定义 1. 设 $F(x_1,x_2,\cdots,x_n)$, $F_{X_i}(x_i)$ $(i=1,2,\cdots,n)$ 分别是 n 维随机变量 $(X_1,,X_2,\cdots,X_n)$ 的分布函数和边缘分布函数, 若对任意的实数 $x_1,x_2,\cdots,x_n$, 有

$$F(x_1,x_2,\cdots,x_n)=F_{X_1}(x_1)F_{x_2}(x_2)\cdots F_{X_n}(x_n)$$

则称 $X_1,X_2,\cdots,X_n$ 是相互独立的.

. . . . 连续型随机变量 $X_1,X_2,\cdots,X_n$ 相互独立的充分必要条件是

$$f(x_1,x_2,\cdots,x_n)=f_{x_1}(x_1)f_{x_2}(x_2)\cdots f_{x_n}(x_n)$$

离散型随机变量 $X_1,X_2,\cdots,X_n$ 相互独立的充分必要条件是

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

定义. 若对所有的 $x_1,\ldots,x_m;y_1,\ldots,y_n$ 有

$$F(x_1,\ldots,x_m,y_1,\ldots,y_n)$$
$$=F_1(x_1,\ldots,x_m)F_2(Y_1,\ldots,Y_n)$$

其中 F_1,F_2,F 分别为 $(X_1,\ldots,X_m),(Y_1,\ldots,Y_n)$ 和 $(X_1,\ldots,X_m,Y_1,\ldots,Y_n)$ 的分布函数，则称随机变量 $(X_1,\ldots,X_m)$ 和 $(Y_1,\ldots,Y_n)$ 是相互独立的.

定理 2. 设 $(X_1,\ldots,X_m)$ 和 $(Y_1,\ldots,Y_n)$ 相互独立,则 $X_i(i=1,2,\ldots,m)$ 和 $Y_j(j=1,2,\ldots,n)$ 相互独立，又若 h,g 是连续函数，则 $h(X_1,\ldots,X_m)$ 与 $g(Y_1,\ldots,Y_n)$ 相互独立.

3.5 两个随机变量的函数的分布

设离散型随机变量 (X,Y) 的联合分布为

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

令 $Z=g(X,Y)$，则 Z 也是一个离散型随机变量，其分布可按如下步骤求得

1. 根据函数关系列出 Z 的所有可能值；
2. 对 Z 的每个可能值 z，$P\{Z=z\}$ 等于所有满足 $g(x_i,y_j)=z$ 的 p_{ij} 之和.

例 **1.** 设二维随机变量 (X,Y) 的分布律为

$X \backslash Y$	-1	0	1
0	0.1	0.2	0.1
1	0.3	0.1	0.2

求 $Z=X+Y$ 的分布律.

解**.** Z 的取值范围为 $-1,0,1,2$, 并且

$$P\{Z=-1\}=P\{X=0,Y=-1\}=0.1,$$

$$P\{Z=0\}=P\{X=0,Y=0\}+P\{X=1,Y=-1\}=0.5,$$

$$P\{Z=1\}=P\{X=0,Y=1\}+P\{X=1,Y=0\}=0.2,$$

$$P\{Z=2\}=P\{X=1,Y=1\}=0.2$$

故 Z 的分布律为

Z	-1	0	1	2
p_k	0.1	0.5	0.2	0.2

对连续型随机变量 (X,Y)，求 $Z=g(X,Y)$ 的密度函数的基本方法是

1. 根据函数关系先求 Z 的分布函数
$$F_Z(z)=P\{Z\le z\}=P\{g(X,Y)\le z\}$$
2. 然后对 $F_Z(z)$ 求导可得 Z 的概率密度.

例 **2.** 设连续型随机变量 (X,Y) 的密度函数为 $f(x,y)$，求 $Z=X+Y$ 的概率密度.

解**.** $Z=X+Y$ 的分布函数为:

$$F_Z(z)=P\{Z\leqslant z\}=\iint_{x+y\leqslant z} f(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

化成累次积分，得

$$F_z(z)=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z-y} f(x,y)\,\mathrm{d}x\right]\mathrm{d}y.$$

对积分 $\int_{-\infty}^{z-y} f(x,y)\,\mathrm{d}x$ 作变量代换, 令 $x=u-y$, 得

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

于是

$$\begin{aligned} F_Z(z) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{z} f(u-y,y)\,\mathrm{d}u\,\mathrm{d}y \\ &= \int_{-\infty}^{z}\left[\int_{-\infty}^{+\infty} f(u-y,y)\,\mathrm{d}y\right]\mathrm{d}u. \end{aligned}$$

上式两边对 z 求导数, 即得 Z 的概率密度

$$f_Z(z) = \int_{-\infty}^{+\infty} f(z-y,y)\,\mathrm{d}y.$$

由 X,Y 的对称性, $f_Z(z)$ 又可写成

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

如果 X 与 Y 相互独立，概率密度分别为 $f_X(x)$ 和 $f_Y(y)$，则 $Z=X+Y$ 的概率密度为

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\,\mathrm{d}x \\ &= \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y)\,\mathrm{d}y. \end{aligned}$$

上述公式称为卷积公式.

例 **3.** 设 X 和 Y 是两个相互独立的随机变量, 它们都服从正态分布 $N(0,1)$, 其概率密度为

$$\begin{aligned} f_X(x) &= \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}, -\infty < x < +\infty, \\ f_Y(y) &= \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}, -\infty < y < +\infty \end{aligned}$$

求 $Z=X+Y$ 的概率密度.

解. 由卷积公式可得

$$\begin{aligned}f_Z(z) &= \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\,\mathrm{d}x \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{x^2}{2}} \cdot \mathrm{e}^{-\frac{(z-x)^2}{2}}\,\mathrm{d}x \\ &= \frac{1}{2\pi}\mathrm{e}^{-\frac{z^2}{4}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\left(x-\frac{z}{2}\right)^2}\,\mathrm{d}x.\end{aligned}$$

令 $t = x - \frac{z}{2}$, 得

$$f_Z(z) = \frac{1}{2\pi}\mathrm{e}^{-\frac{z^2}{4}} \int_{-\infty}^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t = \frac{1}{2\sqrt{\pi}}\mathrm{e}^{-\frac{z^2}{4}}$$

即 Z 服从正态分布 $N(0,2)$.

- 若 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$，且 X 与 Y 相互独立，则
$$Z = X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$
- 若 $X_i \sim N(\mu_i, \sigma_i^2)(i = 1, 2, \ldots, n)$, 且它们相互独立，则 $Z = X_1 + X_2 + \cdots + X_n$ 服从正态分布且
$$Z \sim N(\mu_1 + \cdots + \mu_n, \sigma_1^2 + \cdots + \sigma_n^2).$$

例 **4.** 设随机变量 X, Y 相互独立, 其概率密度分别为

$$f_X(x) = \begin{cases} 1, & 0 \leqslant x \leqslant 1, \\ 0, & \text{其他}, \end{cases} \qquad f_Y(y) = \begin{cases} \mathrm{e}^{-y}, & y > 0, \\ 0, & \text{其他}. \end{cases}$$

求随机变量 $Z = X + Y$ 的概率密度.

解. 由条件知, (X, Y) 的概率密度为

$$f(x, y) = f_X(x) f_Y(y) = \begin{cases} \mathrm{e}^{-y}, & 0 \leqslant x \leqslant 1, y > 0, \\ 0, & \text{其他}, \end{cases}$$

则 Z 的分布函数为

$$F_Z(z) = P\{Z \leqslant z\} = P\{X + Y \leqslant z\} = \iint_{x+y \leqslant z} f(x, y)\,\mathrm{d}x\,\mathrm{d}y.$$

1. 当 $z < 0$ 时, $F_z(z) = 0$；

2. 当 $0 \leqslant z < 1$ 时,$F_z(z) = \int_0^z \left[\int_0^{z-x} \mathrm{e}^{-y}\,\mathrm{d}y\right]\mathrm{d}x = \mathrm{e}^{-z} + z - 1$;

3. 当 $z \geqslant 1$ 时, $F_z(z) = \int_0^1 \left[\int_0^{z-x} \mathrm{e}^{-y}\,\mathrm{d}y\right]\mathrm{d}x = 1 + (1-\mathrm{e})\mathrm{e}^{-z}$.

综上所述得 Z 的分布函数

$$F_z(z) = \begin{cases} 0, & z < 0, \\ \mathrm{e}^{-z} + z - 1, & 0 \leqslant z < 1, \\ 1 + (1-\mathrm{e})\mathrm{e}^{-z}, & z \geqslant 1, \end{cases}$$

故 $Z = X + Y$ 的概率密度为

$$f_Z(z) = F_Z'(z) = \begin{cases} 0, & z < 0 \\ 1 - \mathrm{e}^{-z}, & 0 \leqslant z < 1 \\ (\mathrm{e}-1)\mathrm{e}^{-z}, & z \geqslant 1 \end{cases}$$

练习 **1.** 设某种商品在一周内的需要量是一个随机变量，其概率密度函数为

$$f(x) = \begin{cases} x\mathrm{e}^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

如果各周的需要量相互独立，求两周需要量的概率密度函数.

例 **5.** 设二维随机变量 (X,Y) 在矩形域 $G = \{(x,y) \mid 0 \leqslant x \leqslant 2, 0 \leqslant y \leqslant 1\}$ 上服从均匀分布, 试求边长为 X 和 Y 的矩形面积 S 的概率密度 $f(s)$.

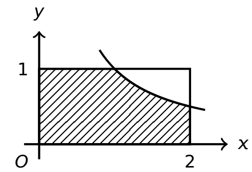
解**.** 由已知, (X,Y) 的概率密度为

$$f(x,y) = \begin{cases} \dfrac{1}{2}, & (x,y) \in G, \\ 0, & \text{其他}. \end{cases}$$

令 $F(s)$ 为 S 的分布函数, 则

$$F(s) = P\{S \leqslant s\} = \iint_{xy \leq s} f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

显然,



1. 当 $s \leqslant 0$ 时, $F(s) = 0$;
2. 当 $s \geqslant 2$ 时, $F(s) = 1$;
3. 当 $0 < s < 2$ 时, 有

$$\begin{aligned} F(s) &= \iint_{xy \leqslant s} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1 - \frac{1}{2}\int_s^2 \left[\int_{\frac{s}{x}}^1 \mathrm{d}y\right]\mathrm{d}x \\ &= \frac{s}{2}(1 + \ln 2 - \ln s) \end{aligned}$$

于是

$$F(s) = \begin{cases} 0, & s \leqslant 0, \\ \dfrac{s}{2}(1 + \ln 2 - \ln s), & 0 < s < 2, \\ 1, & s \geqslant 2. \end{cases}$$

故 S 的概率密度为

$$f(s) = F'(s) = \begin{cases} \dfrac{1}{2}(\ln 2 - \ln s), & 0 < s < 2, \\ 0, & \text{其他}. \end{cases}$$

设随机变量 X 与 Y 相互独立，其分布函数分别为 $F_X(x)$ 和 $F_Y(y)$. 设 $M = \max\{X, Y\}$ 的分布函数为 $F_{\max}(z)$，则有

$$F_{\max}(z) = P\{M \leq x\} = P\{X \leq x, Y \leq z\} = F_X(z)F_Y(z).$$

设 $N = \min\{X, Y\}$ 的分布函数为 $F_{\min}(z)$，则有

$$\begin{aligned} F_{\min}(z) &= P\{N \leq z\} = 1 - P\{N > z\} \\ &= 1 - P\{X > z, Y > z\} = 1 - [1 - F_X(z)][1 - F_Y(z)] \\ &= F_X(z) + F_Y(z) - F_X(z)F_Y(z). \end{aligned}$$

设 $X_1, X_2, \cdots, X_n$ 是 n 个相互独立的随机变量, 它们的分布函数分别为 $F_{X_i}(x_i)\,(i = 1, 2, \cdots, n)$

则 $M=\max(X_1,X_2,\cdots,X_n)$ 及 $N=\min(X_1,X_2,\cdots,X_n)$ 的分布函数分别为

$$F_{\max}(z)=F_{X_1}(z)F_{X_2}(z)\cdots F_{X_n}(z)$$

$$F_{\min}(z)=1-\left[1-F_{X_1}(z)\right]\left[1-F_{x_2}(z)\right]\cdots\left[1-F_{X_n}(z)\right]$$

$X_1,X_2,\cdots,X_n$ 相互独立且具有相同分布函数 $F(x)$ 时, 有

$$F_{\max}(z)=[F(z)]^n$$

$$F_{\min}(z)=1-[1-F(z)]^n$$

一般地，设连续型随机变量 (X,Y) 的联合分布函数为 $F(x,y)$，则 $U=\max\{X,Y\}$ 的分布为

$$F_U(u)=F(u,u).$$

设其边缘分布函数分别为 $F_X(x)$ 和 $F_Y(y)$，则 $V=\min\{X,Y\}$ 的分布为

$$F_V(v)=F_X(v)+F_Y(v)-F(v,v).$$

$$\begin{aligned}F_V(v)&=P\{\min\{X,Y\}\leq v\}=P\{X\leq v \text{ 或 } Y\leq v\}\\&=P\{X\leq v\}+P\{Y\leq v\}-P\{X\leq v,Y\leq v\}\\&=F_X(v)+F_Y(v)-F(v,v)\end{aligned}$$

例 **6.** 设某种型号的电子元件的寿命 (以小时计) 近似服从分布 $N(160,\ 20^2\ )$, 随机地选取 4 只, 求其中没有一只寿命小于 180 小时的概率.

解**.** 将随机选出的 4 只电子元件的寿命分别记为 T_1,T_2,T_3,T_4. 按题意, $T_i\sim N\left(160,20^2\right),i=1,2,3,4$, 其分布函数为 $F(t)$. 令 $T=\min\{T_1,T_2,T_3,T_4\}$, 则

$$F_T(t)=P\{T\leqslant t\}=1-[1-F(t)]^4.$$

于是 $P\{T\geqslant 180\}=1-P\{T<180\}=[1-F(180)]^4$. 依据第二章第四节的引理知

$$F(180)=\Phi\left(\frac{180-160}{20}\right)=\Phi(1)$$

故所求的概率为

$$P\{T\geqslant 180\}=[1-\Phi(1)]^4=(1-0.8413)^4=(0.1587)^4.$$