随机变量的独立性

在第一章中，我们介绍了事件的独立性的概念 .

随机事件A和B相互独立是指A和B各自发生与否没有任何关系.

如果$P(AB) = P(A)P(B)$,就称A和B相互独立.

通俗地讲，随机事件X和Y 相互独立是指X和Y的各自取值情况没有任何关系.

随机变量的独立性

一般地，将随机变量X,Y的各种取值情况通过随机事件$A_x=\{X\le x\}$和$B_y=\{Y\le y\}$来体现，其中x,y均为任意实数.因此X和Y的各自取值情况没有任何关系表现为随机事件$A_x=\{X\le x\}$和$B_y=\{Y\le y\}$各自发生没有任何关系，即对于任意的实数x,y,随机事件A_x和B_y相互独立，从而有

$$P(A_xB_y)=P(A_x)P(B_y)$$

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

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

第四节　随机变量的独立性

一、随机变量独立性的定义

二、随机变量独立性的有关结论

三、小结

3.4.1 随机变量独立性的定义

1. 定义

设(X,Y)为二维随机变量，其分布函数为$F(x,y)$，(X,Y)关于X和Y的边缘分布函数分别为$F_X(x)$和$F_Y(y)$. 如果对于任意的实数x,y, 均有

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

就称随机变量X与Y相互独立.

3.4.1 随机变量独立性的定义

推广到n维随机变量的情况，设n维随机变量$(X_1, X_2, \dots, X_n)$的联合分布函数为$F(x_1, x_2, \dots, x_n)$, X_i的边缘分布函数为$F_{X_i}(x_i)$, $i = 1, 2, \dots, n$.若对任意n个实数$x_1, x_2, \dots, x_n$:

$P(X_1 \le x_1, X_2 \le x_2, \dots, X_n \le x_n) = P(X_1 \le x_1)P(X_2 \le x_2) \dots P(X_n \le x_n)$,

$$F(x_1, x_2, \dots, x_n) = F_{X_1}(x_1)\, F_{X_2}(x_2) \dots F_{X_n}(x_n),$$

则称随机变量$X_1, X_2, \dots, X_n$相互独立。

3.4.1 随机变量独立性的定义

注意：

在独立的情况下，联合分布函数$F(x_1,x_2,\dots,x_n)$与边缘分布函数$F_{X_i}(x_i)$是相互唯一确定的.

离散型随机变量$X_1,X_2,\dots,X_n$ 相互独立的条件等价于：

对$(X_1,X_2,\dots,X_n)$的人一个取值$(x_1,x_2,\dots,x_n)$，

$P(X_1=x_1,X_2=x_2,\dots,X_n=x_n)=P(X_1=x_1)P(X_2=x_2)\dots P(X_n=x_n)$,

连续型随机变量$X_1,X_2,\dots,X_n$ 相互独立的条件等价于：

$$f(x_1,x_2,\dots,x_n)=f_{X_1}(x_1)\,f_{X_2}(x_2)\cdots f_{X_n}(x_n)$$

在$f(x_1,x_2,\dots,x_n)$的连续点处成立，其中$f(x_1,x_2,\dots,x_n)$为$(X_1,X_2,\dots,X_n)$的联合密度函数， $f_{X_i}(x_i)$为X_i的边缘密度函数， $i=1,2,\dots,n$.

例 3.4.1 P79 设(X,Y)有如下联合分布律：

X	Y	
	0	1
0	1/4	b
1	a	1/4

且事件$\{X=0\}$与$\{X+Y=1\}$相互独立. (1)确定常数a, b. (2)判断X和Y是否独立？

解 (1) 由于事件$\{X=0\}$与$\{X+Y=1\}$相互独立，故有

$P(X=0,X+Y=1) = P(X=0)\,P(X+Y=1)$, 即$P(X=0,Y=1) = P(X=0)\,P(Y=1)$,

从而 $b = \left(\frac{1}{4}+b\right)\left(b+\frac{1}{4}\right)$, 因为 $\frac{1}{4}+b+a+\frac{1}{4}=1$，所以 $a = b = \frac{1}{4}$.

(2) 由于对(X,Y)的任一个取值$(i,j), i,j = 0,1$,下式均成立

$P(X=i,Y=j) = P(X=i)\,P(Y=j)$,故$X$和$Y$独立.

3.4.1 随机变量独立性的定义

例3.4.2 P80 已知X和Y的分布律分别为 $X\sim\begin{pmatrix}-1 & 0 & 1\\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4}\end{pmatrix}$, $Y\sim\begin{pmatrix}0 & 1\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix}$,

而且$P\left(X^2=Y^2\right)=1$.(1)求X和Y的联合分布律；(2)判断X和Y是否独立?

解 (1)因为$P\left(X^2=Y^2\right)=1$，所以$P\left(X^2\neq Y^2\right)=0$，从而

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

再结合X和Y的分布律，可得到X和Y的联合分布律为

X	Y		$P(X=x_i)$
	0	1	
-1	0	1/4	1/4
0	1/2	0	1/2
1	0	1/4	1/4
$P(Y=y_i)$	1/2	1/2	

(2)由于

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

$P(X=-1)=\dfrac{1}{4}$,

$P(Y=0)=\dfrac{1}{2}$,

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

所以X和Y不独立.

3.4.1 随机变量独立性的定义

例3.4.3 **P81** 设二维随机变量(X,Y)的概率密度函数为

$$f(x,y)=\begin{cases} 3x, 0<x<1, 0<y<x \\ 0, \quad 其他 \end{cases}$$,判断X和Y是否独立?

解 由题意得， X的边缘密度函数为

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy = \begin{cases} \int_0^x 3xdy, 0<x<1, \\ 0, 其他. \end{cases} = \begin{cases} 3x^2, 0<x<1, \\ 0, 其他. \end{cases}$$

Y的边缘密度函数为

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx = \begin{cases} \int_y^1 3xdx, 0<y<1, \\ 0, 其他. \end{cases} = \begin{cases} \frac{3}{2}(1-y^2), 0<y<1, \\ 0, 其他. \end{cases}$$

由于当$0<x<1, 0<y<x$时， $f(x,y) \neq f_X(x) f_Y(y)$ **故X和Y不独立.**

Example Suppose that random variables (X, Y) obey uniform distribution in rectangular domain: $a \le x \le b\ , c \le y \le d$

Find: (1) joint pdf and marginal pdf of (X,Y);
(2) show that X and Y are independent of each other;
(3) Joint distribution function of (X, Y);
(4) $P(X \le b, Y \le \frac{c+d}{2})$

Solution: **(1)** joint pdf of (X, Y) is:

$$f(x,y)=\begin{cases}\dfrac{1}{(b-a)(d-c)} & a\le x\le b, c\le y\le d\\ 0 & \text{others}\end{cases}$$

Example

In the region of $a \le x \le b, c \le y \le d$, it has:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = \int_c^d \frac{dy}{(b-a)(d-c)} = \frac{1}{b-a}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx = \int_a^b \frac{dx}{(b-a)(d-c)} = \frac{1}{d-c}$$

Out of this region, $f_X(x) = 0, \quad f_Y(y) = 0$

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{others} \end{cases} \qquad f_Y(y) = \begin{cases} \dfrac{1}{d-c} & c \le y \le d \\ 0 & \text{others} \end{cases}$$

(2)

$$\because \quad f_X(x)\cdot f_Y(y)=\begin{cases}\dfrac{1}{b-a}\cdot\dfrac{1}{d-c} & a\le x\le b,\ c\le y\le d\\ 0 & \text{others}\end{cases}$$

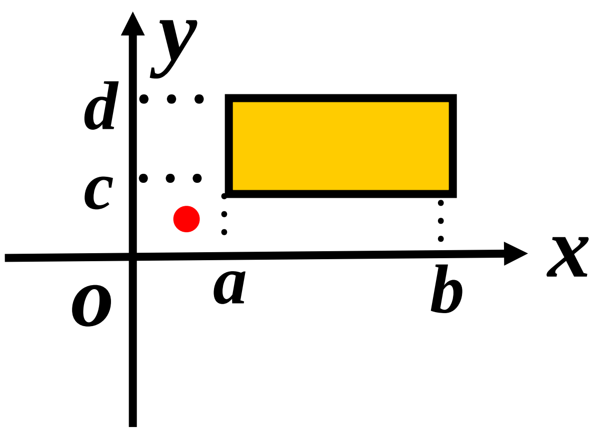
$$= f(x,y)$$

X and Y are independent of each other.

(3) When $x<a$ or $y<c$

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

$$=\int_{-\infty}^{x}\int_{-\infty}^{y} o\,dudv \;=0$$



Section 2.7 The independence of random variables

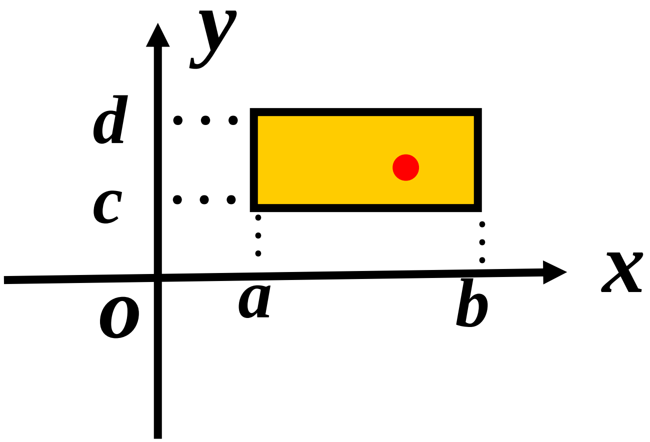
Example

When $a \le x \le b, \ c \le y \le d$

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

$$= \int_{a}^{x}\int_{c}^{y} \frac{1}{(b-a)(d-c)}\,dudv$$

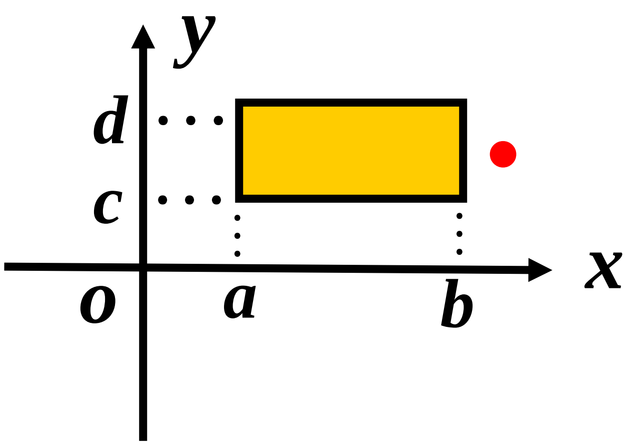
$$= \frac{(x-a)(y-c)}{(b-a)(d-c)}$$



When $x > b \ , c \le y \le d$

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

$$= \int_{a}^{b}\int_{c}^{y} \frac{1}{(b-a)(d-c)}\,dudv = \frac{(y-c)}{(d-c)}$$



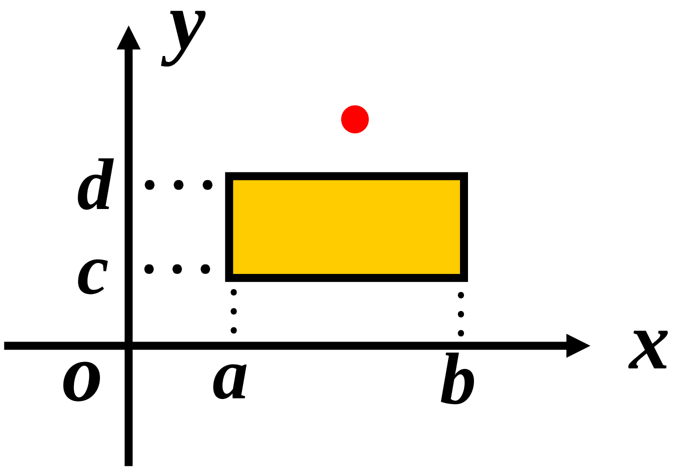
Section 2.7 The independence of random variables

Example

When $y > d$, $a \le x \le b$

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

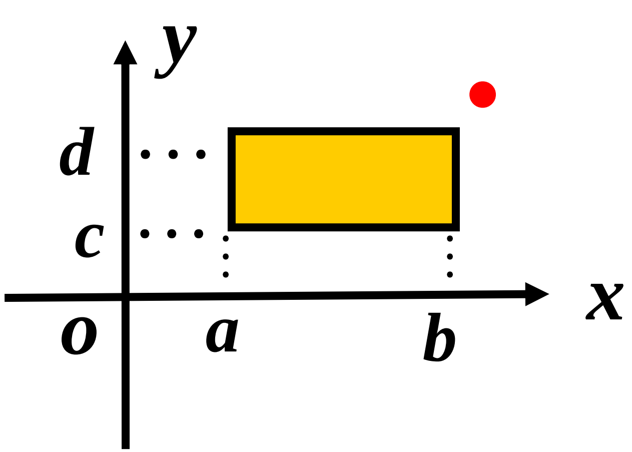
$$=\int_{c}^{d}\int_{a}^{x}\frac{1}{(b-a)(d-c)}\,dudv = \frac{(x-a)}{(b-a)}$$



When $x > b$, $y > d$

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

$$=\int_{a}^{b}\int_{c}^{d}\frac{1}{(b-a)(d-c)}\,dudv = 1$$



So, joint distribution function of (***X***, ***Y***) is :

$$F(x,y)=\begin{cases} 0 & x<a \text{ or } y<c \\ \dfrac{(x-a)(y-c)}{(b-a)(d-c)} & a\le x\le b, c\le y\le d \\ \dfrac{y-c}{d-c} & x>b, c\le y\le d \\ \dfrac{x-a}{b-a} & y>d, a\le x\le b \\ 1 & x>b \ \& \ y>d \end{cases}$$

3.4.1 随机变量独立性的定义

例3.4.4 P81 $(X,Y) \sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$,证明$X$和$Y$独立等价于$\rho = 0$

证明 由例3.2.4 P 70知，X和Y的边缘分布均为正态分布，且

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

X和Y的联合密度函数为

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

X,Y的边缘密度函数为 $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$ $f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$

假设X和Y独立，故对一切(x,y),都有 $f(x,y) = f_X(x) f_Y(y)$

本题中取$(x,y) = (\mu_1,\mu_2)$,得$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} = \frac{1}{\sqrt{2\pi}\sigma_1} \cdot \frac{1}{\sqrt{2\pi}\sigma_2}$,所以$\rho = 0$

3.4.2 随机变量独立性的有关结论

定理3.4.1

(1) **n个随机变量**$X_1, X_2, \dots, X_n$相互独立的充要条件是：对一切使得$\{X_1 \in A_1\}, \{X_2 \in A_2\}, \cdots, \{X_n \in A_n\}$为事件的实数集$A_1, A_2, \dots, A_n$,

$$P(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n) = P(X_1 \in A_1)P(X_2 \in A_2) \dots P(X_n \in A_n)$$

(2) **设随机变量**$X_1, X_2, \dots, X_n$相互独立，且$f_1, f_2, \dots, f_n$是**n个恰当的实值函数，则**$f_1(X_1), f_2(X_2), \dots, f_n(X_n)$也相互独立.

3.4.2 随机变量独立性的有关结论

定义3. 4. 2（两个随机向量独立性的定义）

(1) 设（$X_1, X_2, \dots, X_n$）和（$Y_1, Y_2, \dots, Y_n$）为两个随机向量，其分布函数分别为$F_1(x_1, x_2, \dots, x_n)$和$F_2(y_1, y_2, \dots, y_n)$的联合分布函数为$F(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m)$. 若对任意的$x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m$,

$$F(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) = F_1(x_1, x_2, \dots, x_n)F_2(y_1, y_2, \dots, y_n).$$

则称随机向量（$X_1, X_2, \dots, X_n$）和（$Y_1, Y_2, \dots, Y_n$）相互独立.

3.4.2 随机变量独立性的有关结论

定理3. 4. 3

设随机向量（$X_1, X_2, \ldots, X_n$）和（$Y_1, Y_2, \ldots, Y_m$）相互独立.

（1）（$X_1, X_2, \ldots, X_n$）的子向量与（$Y_1, Y_2, \ldots, Y_m$）的子向量相互独，特别地，$X_i(i = 1, 2, \ldots, n)$和$Y_j(j = 1, 2, \ldots, m)$相互独立.

（2）若f和g是两个恰当的函数，则$f(X_1, X_2, \ldots, X_n)$和$g(Y_1, Y_2, \ldots, Y_m)$也相互独立.

例如：设随机变量X与Y相互独立，$f(x)$和$g(y)$是连续函数，则随机变量$f(x)$与$g(y)$也相互独立.

小结

定理 (1) 设 (X,Y) 为二维离散型随机变量，其分布律为

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

则 X 和 Y 相互独立的充要条件为

$$p_{ij}=p_{i\bullet}p_{\bullet j}\text{，}\quad i=1,2,\cdots, j=1,2,\cdots.$$

定理 (2) 设 (X,Y) 为二维连续型随机变量，其密度函数为 $f(x,y)$，则 X 和 Y 相互独立的充要条件为对平面上几乎所有的点 (x,y)，有

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