浙江大学
ZHEJIANG UNIVERSITY

第23讲　随机变量的独立性

相互独立的随机变量

第8讲中把A与B两个事件的独立性定义为$P(AB)=P(A)P(B)$，而随机变量的取值往往可以构成无数的事件，如$X=1, X<1$等，为此要定义两个随机变量的独立性必须包含两个随机变量的许多个事件间的独立。

设x, y为实数，设$A=\{X\le x\}, B=\{Y\le y\}$.

独立性定义：

设$F(x,y)$是二元随机变量(X,Y)的分布函数，$F_X(x)$是X的边际分布函数，$F_Y(y)$是Y的边际分布函数，若对所有x,y有：

$P(X \le x, Y \le y) = P(X \le x)P(Y \le y)$ 即

$$\boxed{F(x,y) = F_X(x)F_Y(y)}$$

称随机变量X,Y相互独立。

独立性等价判断：

离散型

用分布律判断。对一切i, j都成立 $\boxed{p_{ij} = p_{i\bullet} p_{\bullet j}}$

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

连续型

用密度函数判断。对在平面的点(x, y)

几乎处处成立 $\boxed{f(x, y) = f_X(x) f_Y(y).}$

即在平面上除去“面积”为零的集合以外，

上述等式处处成立。

- 例1：已知(X,Y)的联合分布律，试判断X与Y的独立性。

$X \backslash Y$	0	1	$P(X=i)$
1	$1/6$	$2/6$	$1/2$
2	$1/6$	$2/6$	$1/2$
$P(Y=j)$	$1/3$	$2/3$	

解：逐个检验$p_{ij} \overset{?}{=} p_{i\cdot} \times p_{\cdot j}$

$$P(X=1,Y=0)=1/6=P(X=1)P(Y=0)$$
$$P(X=2,Y=0)=1/6=P(X=2)P(Y=0)$$
$$P(X=1,Y=1)=2/6=P(X=1)P(Y=1)$$
$$P(X=2,Y=1)=2/6=P(X=2)P(Y=1)$$

因而X,Y是相互独立的.

需要检验所有等式成立才能得独立结论

- 例2：已知(X,Y)的联合分布律，试判断X与Y的独立性。

X \ Y	0	1	P(X=i)
1	$1/6$	$2/6$	$1/2$
2	$2/6$	$1/6$	$1/2$
P(Y=j)	$1/2$	$1/2$	

解：逐个检验$p_{ij} \overset{?}{=} p_{i\cdot} \times p_{\cdot j}$

$$P(X=1,Y=0)=1/6$$

$$P(X=1)P(Y=0)=1/2\times 1/2=1/4$$

故$P(X=1,Y=0)\neq P(X=1)P(Y=0)$

因而X与Y不相互独立.

只要有一对i,j使

$$p_{ij} \neq p_{i\cdot} \times p_{\cdot j}$$

就能判断不独立

例3:设X与Y是相互独立的随机变量，已知(X,Y)的联合分布律，求其余未知的概率值。

$X \backslash Y$	0	1	2	$P(X=x_i)$
1	0.01	0.2	0.04	0.25
2	0.03	0.6	0.12	0.75
$P(Y=y_j)$	0.04	0.8	0.16	

例4: (X,Y)具有概率密度$f(x,y)=\begin{cases}6e^{-(2x+3y)}, & x>0,y>0\\ 0, & 其他\end{cases}$,

问X和Y是否相互独立?

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

$$=\begin{cases}\int_0^{\infty}6e^{-(2x+3y)}dy=2e^{-2x}, & x>0\\ 0 & ,x\leq 0\end{cases}$$

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

$$=\begin{cases}\int_0^{\infty}6e^{-(2x+3y)}dx=3e^{-3y}, & y>0\\ 0 & ,y\leq 0\end{cases}$$

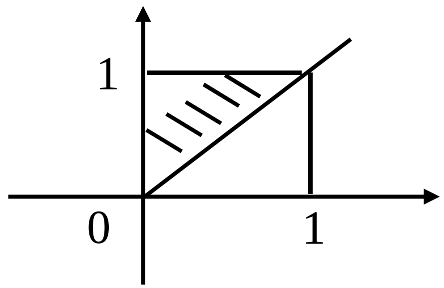
故有$f(x,y)=f_X(x)\cdot f_Y(y)$,因而$X,Y$是相互独立的。

结论：连续型变量独立，其联合密度函数一定能分解成x的函数与y的函数的乘积。即$f(x,y)=g(x)h(y)$.

但是，如果$(X,Y)\sim f(x,y)=\begin{cases}8xy, & 0<x<y<1\\ 0, & 其他\end{cases}$，则

$$f_X(x)=\begin{cases}\int_x^1 8xydy=4x(1-x^2), & 0<x<1\\ 0, & 其他\end{cases}$$

$$f_Y(y)=\begin{cases}\int_0^y 8xydx=4y^3, & 0<y<1\\ 0, & 其他\end{cases}$$



所以，当$0<y<x<1$时，$f(x,y)=0,\ f_X(x)\cdot f_Y(y)>0,$

因而X,Y不是相互独立的。

例5：证明，对于二元正态随机变量(X,Y)，

X与Y相互独立的充要条件是参数$\rho=0$.

证：因为(X,Y)的概率密度为：

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

其边际概率密度的乘积为：

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

"$\Leftarrow$" 如果$\rho=0$，则对于所有x, y，有$f(x,y)=f_X(x)f_Y(y)$
即X,Y相互独立。

"$\Rightarrow$" 反之，若X,Y相互独立，

由于$f(x,y), f_X(x), f_Y(y)$都是连续函数，

故对于所有的x, y，有$f(x,y)=f_X(x)f_Y(y)$

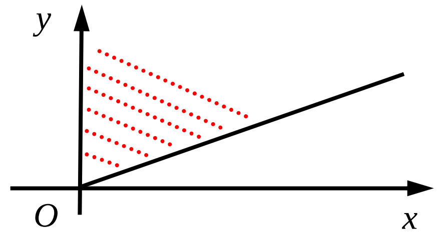
特别的有$f(\mu_1,\mu_2)=f_X(\mu_1)f_Y(\mu_2)$ ，

即$\dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}=\dfrac{1}{2\pi\sigma_1\sigma_2} \quad \Rightarrow \rho=0$

例6. 设甲，乙两种元件的寿命X, Y相互独立，服从同一分布，

其概率密度为：$f(x)=\begin{cases}\frac{1}{2}e^{-\frac{x}{2}}, x>0\\ 0, \quad 其它\end{cases}$

求甲元件寿命不大于乙元件寿命2倍的概率。



解：(X, Y)的联合密度为

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}\frac{1}{4}e^{-\frac{x+y}{2}} & x>0, y>0\\ 0 & 其他\end{cases}$$

$$P(X\le 2Y)=\iint\limits_{x\le 2y} f(x,y)dxdy=\int_0^{\infty}dx\int_{\frac{x}{2}}^{\infty}\frac{1}{4}e^{-\frac{x+y}{2}}dy \;=\int_0^{\infty}\frac{1}{2}e^{-\frac{x}{2}}e^{-\frac{x}{4}}dx$$

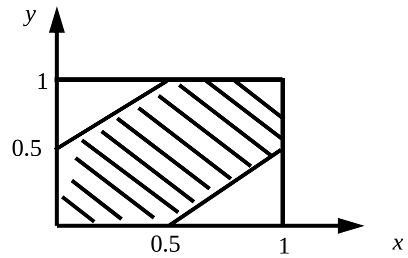
$$=\int_0^{\infty}\frac{1}{2}e^{-\frac{3x}{4}}dx=\frac{2}{3}$$

例7：在区间(0, 1)上任取两数，求这两数之差的绝对值小于0.5的概率。

解：设X, Y分别为(0, 1)上任取的两数，

则X与Y为独立且同分布的，均服从$U(0,1)$

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



$$P(|X-Y|<0.5)=\iint\limits_{|x-y|<0.5} f(x,y)dxdy$$

$$=\iint\limits_{\substack{|x-y|<0.5\\ 0<x,y<1}} 1dxdy=S_G=1-0.5^2=0.75$$

一般n元随机变量的一些概念和结果

- n元随机变量

设E是一个随机试验，它的样本空间是$S=\{e\}$;

设$X_1=X_1(e), X_2=X_2(e), \cdots, X_n=X_n(e)$是定义在$S$上的随机变量，

由它们构成的一个n元向量$(X_1, X_2, \cdots, X_n)$称为n元随机变量。

- n元随机变量的分布函数

对于任意n个实数$x_1, x_2, \cdots, x_n$，n元函数：

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

称为n元随机变量$(X_1, X_2, \cdots, X_n)$的分布函数。

- n元离散型随机变量的分布律

设$(X_1, X_2, \cdots, X_n)$的取值记为$(x_1, x_2, \cdots, x_n)$

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

$(x_1, x_2, \cdots, x_n)$ 取遍所有可能值，

称为n元离散型随机变量$(X_1, X_2, \cdots, X_n)$的分布律.

- n元连续型随机变量的概率密度

若存在非负函数$f(x_1, x_2, \cdots, x_n)$，使得对于任意实数$x_1, x_2, \cdots, x_n$

$$F(x_1, x_2, \cdots x_n) = \int_{-\infty}^{x_n}\int_{-\infty}^{x_{n-1}} \cdots \int_{-\infty}^{x_1} f(x_1, x_2, \cdots x_n) dx_1 dx_2 \cdots dx_n$$

$f(x_1, x_2, \cdots, x_n)$称为n元连续型随机变量$(X_1, X_2, \cdots, X_n)$的概率密度.

n元随机变量的边际分布

若$(X_1, X_2, \cdots, X_n)$的分布函数$F(x_1, x_2, \cdots x_n)$

（分布律$p(x_1, x_2, \cdots x_n)$,概率密度$f(x_1, x_2, \cdots x_n)$）已知，

则$(X_1, X_2, \cdots, X_n)$的$k(1 \le k \le n)$元边际分布就随之确定. 比如：

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

$$p_{X_1}(x_1) = P(X_1 = x_1) = \sum_{x_2, x_3, \cdots x_n} P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n)$$

$$f_{X_1}(x_1) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, x_2, \cdots, x_n) dx_2 dx_3 \cdots dx_n$$

n元随机变量的相互独立

$(X_1,\ldots,X_n)$的分布函数为$F(x_1,x_2,\cdots,x_n)$，若对于所有的 $x_1,x_2,\cdots,x_n$,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。

离散

$(X_1,\ldots,X_n)$的分布律为$p(x_1,x_2,\cdots,x_n)$，若对于所有的 $x_1,x_2,\cdots,x_n$,有： $p(x_1,x_2,\cdots,x_n)=p_{X_1}(x_1)p_{X_2}(x_2)\cdots p_{X_n}(x_n)$

则称$X_1,X_2,\cdots,X_n$是相互独立的。

连续

$(X_1,\ldots,X_n)$的概率密度为$f(x_1,x_2,\cdots,x_n)$，若对于$x_1,\cdots,x_n$, 几乎处处有： $f(x_1,x_2,\cdots,x_n)=f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n)$

则称$X_1,X_2,\cdots,X_n$是相互独立的。

$(X_1, X_2, \cdots, X_m)$与$(Y_1, Y_2, \cdots, Y_n)$的独立性

设$(X_1, X_2, \cdots, X_m)$的分布函数为$F_1(x_1, x_2, \cdots x_m)$,

$(Y_1, Y_2, \cdots, Y_n)$的分布函数为$F_2(y_1, y_2, \cdots y_n)$,

$(X_1, X_2, \cdots, X_m, Y_1, Y_2, \cdots, Y_n)$的分布函数为$F(x_1, x_2, \cdots x_m, y_1, y_2, \cdots y_n)$

若$F(x_1, x_2, \cdots, x_m, y_1, y_2, \cdots, y_n) = F_1(x_1, x_2, \cdots x_m) F_2(y_1, y_2, \cdots y_n)$

对一切$x_1, \cdots x_m, y_1, \cdots y_n$成立，称$(X_1, \cdots, X_m)$与$(Y_1, \cdots, Y_n)$相互独立.

定理：设$(X_1, X_2, \cdots, X_m)$与$(Y_1, Y_2, \cdots, Y_n)$相互独立，

(1)则$X_i\,(i = 1, 2, \cdots, m)$与$Y_j\,(j = 1, 2, \cdots, n)$相互独立。

(2)若$h(x_1, x_2, \cdots, x_m)$和$g(y_1, y_2, \cdots, y_n)$是连续函数，

则$h(X_1, X_2, \cdots, X_m)$和$g(Y_1, Y_2, \cdots, Y_n)$相互独立。