随机变量的独立性

设$(X,\ Y)$是二维随机变量，其联合分布函数为$F(x,\ y)$，又随机变量X的分布函数为$F_X(x)$，随机变量Y的分布函数为$F_Y(y)$．如果对于任意的$x,\ y$，有

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

则称$X,\ Y$是相互独立的随机变量．

说　明

(1). 由于

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

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

可知，随机变量 X 与 Y 相互独立，实际上是指：对于任意的 x, y，随机事件

$$\{X \le x\} \quad 与 \quad \{Y \le y\}$$

相互独立.

说 明

(2). 如果随机变量 X 与 Y 相互独立，则由

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

可知，

二维随机变量$(X,\ Y)$的联合分布函数$F(x,\ y)$可由其边缘分布函数$F_X(x)$与$F_Y(y)$唯一确定.

例 1

设二维随机变量$(X,\ Y)$的联合分布函数为

$$F(x,\ y)=\frac{1}{\pi^2}\left(\frac{\pi}{2}+\arctan\frac{x}{5}\right)\left(\frac{\pi}{2}+\arctan\frac{y}{10}\right)$$

$$(-\infty<x<+\infty,\ -\infty<y<+\infty)$$

试判断X与Y是否相互独立？

解：

X的边缘分布函数为

例 1（续）

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

$$=\lim_{y\to+\infty}\frac{1}{\pi^2}\left(\frac{\pi}{2}+\arctan\frac{x}{5}\right)\left(\frac{\pi}{2}+\arctan\frac{y}{10}\right)$$

$$=\frac{1}{\pi}\left(\frac{\pi}{2}+\arctan\frac{x}{5}\right)\qquad\left(x\in(-\infty,\quad+\infty)\right)$$

Y 的边缘分布函数为

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

$$=\lim_{x\to+\infty}\frac{1}{\pi^2}\left(\frac{\pi}{2}+\arctan\frac{x}{5}\right)\left(\frac{\pi}{2}+\arctan\frac{y}{10}\right)$$

例 1（续）

$$=\frac{1}{\pi}\left(\frac{\pi}{2}+\arctan\frac{y}{10}\right)\qquad\left(y\in(-\infty,\ +\infty)\right)$$

所以，对于任意的实数 x， y，有

$$F(x,\ y)=\frac{1}{\pi^2}\left(\frac{\pi}{2}+\arctan\frac{x}{5}\right)\left(\frac{\pi}{2}+\arctan\frac{y}{10}\right)$$

$$=\frac{1}{\pi}\left(\frac{\pi}{2}+\arctan\frac{x}{5}\right)\cdot\frac{1}{\pi}\left(\frac{\pi}{2}+\arctan\frac{y}{10}\right)=F_X(x)F_Y(y)$$

所以 X 与 Y 是相互独立的随机变量.

离散型随机变量的独立性

设$(X,\ Y)$是二维离散型随机变量，其联合分布律为

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

又随机变量X的分布律为

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

随机变量Y的分布律为

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

如果对于任意的$i,\ j$

$$p_{ij} = p_{i\cdot}\, p_{\cdot j}$$

则称$X,\ Y$是相互独立的随机变量.

例 2

设二维离散型随机变量 $(X, \quad Y)$ 的联合分布律为

Y \ X	1	2	3
1	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$
2	$\frac{1}{3}$	α	β

试确定常数 α，β 使得随机变量 X 与 Y 相互独立.

解：

由表，可得随机变量 X 与 Y 的边缘分布律为

例 2（续）

$X \backslash Y$	1	2	3	$p_{i\cdot}$
1	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$	$\frac{1}{3}$
2	$\frac{1}{3}$	α	β	$\frac{1}{3}+\alpha+\beta$
$p_{\cdot j}$	$\frac{1}{2}$	$\frac{1}{9}+\alpha$	$\frac{1}{18}+\beta$	

如果随机变量X与Y相互独立，则有

$$p_{ij} = p_{i\cdot} p_{\cdot j} \qquad \left(i=1,\ 2;\ \ j=1,\ 2,\ 3 \right)$$

由此得

例 2（续）

$$\frac{1}{9}=P\{X=1,\quad Y=2\}=P\{X=1\}P\{Y=2\}=\frac{1}{3}\cdot\left(\frac{1}{9}+\alpha\right)$$

由此得　$\alpha=\dfrac{2}{9}$;

又由

$$\frac{1}{18}=P\{X=1,\quad Y=3\}=P\{X=1\}P\{Y=3\}=\frac{1}{3}\cdot\left(\frac{1}{18}+\beta\right)$$

由此得　$\beta=\dfrac{1}{9}$.

而当$\alpha=\dfrac{2}{9}$，$\beta=\dfrac{1}{9}$ 时，联合分布律及边缘分布律为

例 2（续）

X \ Y	1	2	3	$p_{i\cdot}$
1	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$	$\frac{1}{3}$
2	$\frac{1}{3}$	$\frac{2}{9}$	$\frac{1}{9}$	$\frac{2}{3}$
$p_{\cdot j}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$	

可以验证，此时有

$$p_{ij} = p_{i\cdot}p_{\cdot j} \qquad (i=1, \ 2; \ \ j=1, \ 2, \ 3)$$

因此当$\alpha=\frac{2}{9}$，$\beta=\frac{1}{9}$时，X与Y相互独立.

例 3

将两个球等可能地放入编号为1, 2, 3的三个盒子中.

令：X：放入1号盒中的球数；

Y：放入2号盒中的球数.

试判断随机变量 X 与 Y 是否相互独立？

解：

X 的可能取值为 0，1，2；Y 的可能取值为 0，1，2.

由 §3.1 知 X 与 Y 的联合分布律及边缘分布律为

例 3（续）

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

$$P\{X=1,\quad Y=2\}=0 \neq P\{X=1\}P\{Y=2\}=\frac{4}{9}\cdot\frac{1}{9}$$

随机变量 X 与 Y 不独立.

连续型随机变量的独立性

设(X, Y)是二维连续型随机变量，其联合密度函数为$f(x, y)$，

又随机变量X的边缘密度函数为$f_X(x)$,随机变量Y的边缘密度函数为$f_Y(y)$，如果对于几乎所有的x, y有，

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

则称X, Y是相互独立的随机变量.

特别地，上式对$f(x, y)$的所有连续点(x, y)必须成立.

说　明

这里所谓的“对几乎所有的x，y”是指：

那些使得等式

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

不成立的全体点(x, y)所构成集合的“面积”为0.

例 4

设二维随机变量(X, Y)的密度函数为

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

试判断随机变量 X 与 Y 是否相互独立？

解：

当$0\le x\le 1$时，

$$f_X(x)=\int_{-\infty}^{+\infty} f(x, y)dy=\int_0^2\left(x^2+\frac{1}{3}xy\right)dy=2x^2+\frac{2}{3}x$$

例 4（续）

所以，随机变量 X 的密度函数为

$$f_X(x)=\begin{cases}2x^2+\dfrac{2}{3}x & 0\le x\le 1\\ 0 & \text{其它}\end{cases}$$

当 $0\le y\le 2$ 时，

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,\ y)dx=\int_0^1\left(x^2+\frac{1}{3}xy\right)dx=\frac{1}{3}+\frac{1}{6}y$$

所以，随机变量 Y 的密度函数为

例 4（续）

$$f_Y(y)=\begin{cases}\frac{1}{3}+\frac{1}{6}y & 0\le y\le 2\\ 0 & \text{其它}\end{cases}$$

$$f_X(x)=\begin{cases}2x^2+\frac{2}{3}x & 0\le x\le 1\\ 0 & \text{其它}\end{cases}$$

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

由于当$0<x<1$，$0<y<2$时，

$$f(x,\ y)\neq f_X(x)f_Y(y)$$

所以，随机变量X与Y不独立.

例 5

甲、乙两人约定在某地相会，假定每人的到达时间是相互独立的，且均服从中午12时到下午1时的均匀分布．试求先到者需等待10分钟以内的概率．

解：

设甲于12时 X 分到达，设乙于12时 Y 分到达．则随机变量 X 与 Y 相互独立，且都服从区间 $[0,\ 60]$ 上的均匀分布．

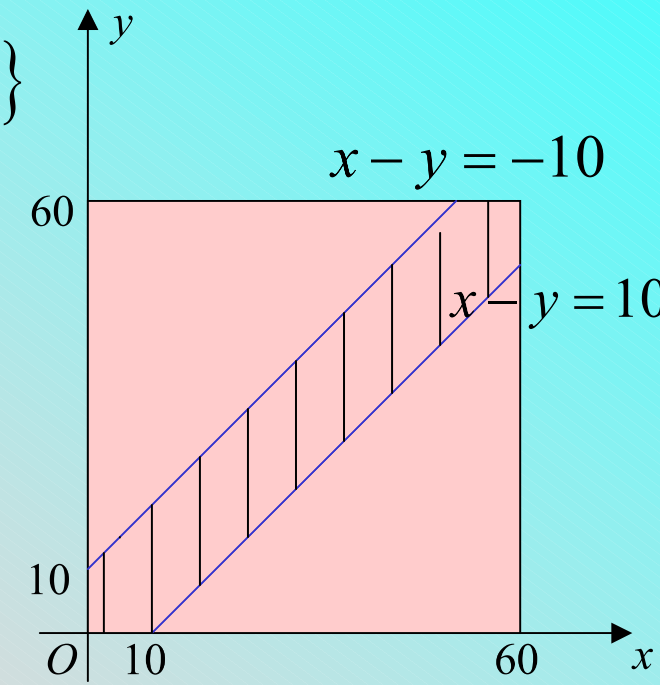
所以，$(X,\ Y)$的联合密度函数为

例 5（续）

$$f(x,\ y)=\begin{cases}\dfrac{1}{3600} & 0\le x\le 60,0\le y\le 60\\ 0 & \text{其它}\end{cases}$$

设：$A=\{$先到者等待时间不超过10分钟$\}$

则有，　$A=\{|X-Y|\le 10\}$



满足上述条件的点为图中直线

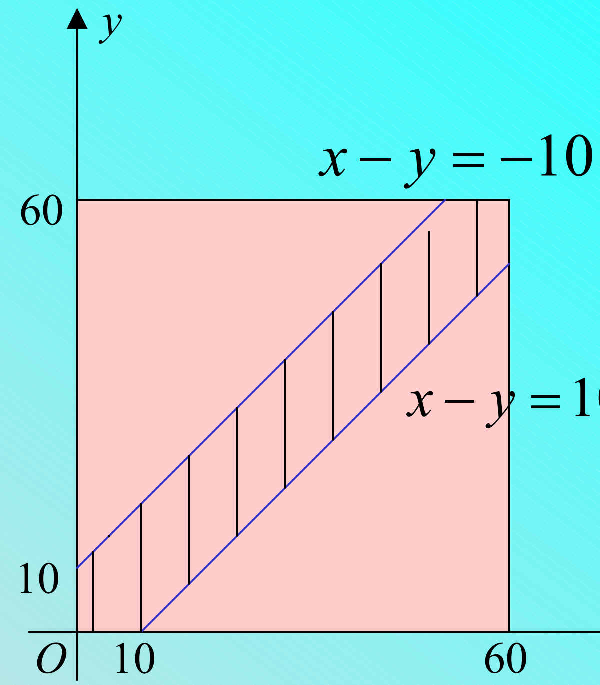
$$x-y=10$$

与直线

$$x-y=-10$$

之间的部分.

例 5（续）



所以，所求概率为

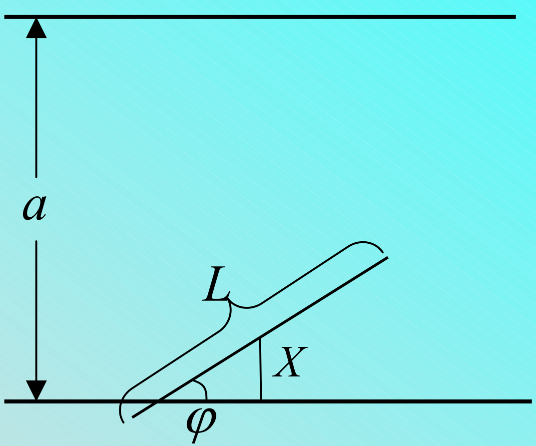
$$P(A)=P\{|X-Y|\le 10\}$$

$$=\iint\limits_{|x-y|\le 10} f(x,\ y)dxdy$$

$$=\frac{3600-50\times 50}{3600}=\frac{11}{36}$$

例6（Buffon投针问题）

平面上画有等距离为a的一些平行线，向此平面上任意投一根长度为$L(L<a)$的针，试求该针与任一平行直线相交的概率.



解：

设：X：针的中心到最近一条平行线的距离；

φ：针与X所在投影线的夹角.

例 6（续）

则随机变量 X 服从区间 $\left[0,\ \frac{a}{2}\right]$ 上的均匀分布；

随机变量 φ 服从区间 $[0,\ \pi]$ 上的均匀分布；

并且随机变量 X 与 φ 相互独立.

所以二维随机变量 $(X,\ \varphi)$ 的联合密度函数为

$$f(x,\ y)=\begin{cases}\dfrac{2}{\pi a} & 0\le x\le\dfrac{a}{2},\ 0\le y\le\pi\\ 0 & \text{其它}\end{cases}$$

例 6（续）

设：$A=\{$针与任一直线相交$\}$

则 $A=\left\{\dfrac{X}{\sin\varphi}<\dfrac{L}{2}\right\}=\left\{X<\dfrac{L}{2}\sin\varphi\right\}$

所以，

$$P(A)=P\left\{X<\frac{L}{2}\sin\varphi\right\}=\iint\limits_{x<\frac{L}{2}\sin y} f(x,\quad y)dxdy$$

$$=\int_0^{\pi}dy\int_0^{\frac{L}{2}\sin y}\frac{2}{\pi a}dx=\frac{2}{\pi a}\int_0^{\pi}\frac{L}{2}\sin y\,dy=\frac{2L}{\pi a}$$

说　　明

由本题的答案 $$P(A)=\frac{2L}{\pi a}$$

我们有圆周率 π 的近似计算公式：

$$\pi=\frac{2L}{a}\cdot\frac{1}{P(A)}$$

若我们投针 N 次，其中有 n 次与平行线相交，则以

$\frac{n}{N}$ 作为 $P(A)$ 的近似值代入上式，得

$$\pi\approx\frac{2L}{a}\cdot\frac{N}{n}$$

说　　明

历史上，确有些学者做过此项实验，下表就是一些有关资料（其中把a折算为1）：

实验者	年　份	针　长	投掷次数	相交次数	π 的近似值
Wolf	1850	0.8	5000	2532	3.1596
Smith	1855	0.6	3204	1218.5	3.1554
De Morgan	1860	1.0	600	382.5	3.137
Fox	1884	0.75	1030	489	3.1595
Lazzerini	1901	0.83	3408	1808	3.1415929
Reina	1925	0.5419	2520	859	3.1759

说 明

上述的计算方法就是一种概率方法，它概括起来就是：

首先建立一个概率模型，它与我们感兴趣的某些量（如上面的常数π）有关.

然后设计适当的随机试验，并通过这个试验的结果来确定这些量.

现在，随着计算机的发展，已按上述思路建立起一类新的计算方法——$Monte-Carlo$方法.

例 7（正态随机变量的独立性）

设二维随机变量$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, r)$

则(X, Y)的联合密度函数为

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}$$

$$\cdot \exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2r(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

又随机变量 X 的边缘密度函数为

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \qquad (-\infty < x < +\infty)$$

例 7（续）

随机变量 Y 的边缘密度函数为

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}} \qquad (-\infty<y<+\infty)$$

所以，当 $r=0$ 时，$(X, \quad Y)$ 的联合密度函数为

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

$$=f_X(x)\cdot f_Y(y)$$

这表明，随机变量 X 与 Y 相互独立；

例 7（续）

反之，如果随机变量 X 与 Y 相互独立，则对任意的实数 x， y，有

$$f(x,\quad y)=f_X(x)\cdot f_Y(y)$$

特别地，我们有

$$f(\mu_1,\quad \mu_2)=f_X(\mu_1)\cdot f_Y(\mu_2)$$

即，

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}=\frac{1}{\sqrt{2\pi}\,\sigma_1}\cdot\frac{1}{\sqrt{2\pi}\,\sigma_2}$$

例 7（续）

由此得，$r=0$.

综上所述，我们有以下重要结论：

二元正态随机变量 $N(\mu_1,\ \mu_2,\ \sigma_1^2,\ \sigma_2^2,\ r)$ 相互独立的充分必要条件为：

$$r=0.$$

§4随机变量的独立性

n维随机变量的独立性

设$(X_1, X_2, \cdots, X_n)$是n维随机变量，其联合分布函数为$F(x_1, x_2, \cdots, x_n)$，又随机变量X_i的分布函数为$F_{X_i}(x_i)$，$(i=1, 2, \cdots, n)$．如果对于任意的n维实数组$(x_1, x_2, \cdots, x_n)$，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$X_1, X_2, \cdots, X_n$是相互独立的随机变量.

注意： 若 X,Y 独立，f(x),g(y) 是连续函数，则 f(X),g(Y) 也独立。