# §3.2 二维 r.v.的条件分布

## 二维离散 r.v.的条件分布律

设二维离散型 r.v. $(X, Y)$的分布

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

若

$$p_{i\bullet}=P(X=x_i)=\sum_{j=1}^{\infty}p_{ij}>0$$

则称

$$\frac{P(X=x_i, Y=y_j)}{P(X=x_i)}=\frac{p_{ij}}{p_{i\bullet}}\overset{\text{记作}}{=}P(Y=y_j \mid X=x_i)$$

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

为**在 $X=x_i$ 的条件下, $Y$ 的条件分布律**

若 $p_{\bullet j} = P(Y = y_j) = \sum_{i=1}^{\infty} p_{ij} > 0,$

则称 $\dfrac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \dfrac{p_{ij}}{p_{\bullet j}} \overset{\text{记作}}{=} P(X = x_i | Y = y_j)$

$i = 1, 2, \cdots$

为在 $Y = y_j$ 的条件下 $X$ 的条件分布律

类似乘法公式

$$P(X = x_i, Y = y_j) \underset{\text{或}}{=} P(X = x_i) P(Y = y_j | X = x_i)$$

$$= P(Y = y_j) P(X = x_i | Y = y_j)$$

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

类似于全概率公式

$$P(X=x_i)=\sum_{j=1}^{\infty}p_{ij}=\sum_{j=1}^{\infty}P(X=x_i,Y=y_j)$$

$$=\sum_{j=1}^{\infty}P(X=x_i \mid Y=y_j)P(Y=y_j)$$

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

$$P(Y=y_j)=\sum_{i=1}^{\infty}p_{ij}=\sum_{i=1}^{\infty}P(X=x_i,Y=y_j)$$

$$=\sum_{i=1}^{\infty}P(Y=y_j \mid X=x_i)P(X=x_i)$$

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

**例1** 把三个球等可能地放入编号为 1, 2, 3 的三个盒子中, 每盒可容球数无限. 记 $X$ 为落入 1 号盒的球数, $Y$ 为落入 2 号盒的球数，求

(1) 在 $Y = 0$ 的条件下，$X$ 的分布律;

(2) 在 $X = 2$ 的条件下，$Y$ 的分布律.

**解** 先求联合分布，

$$P(X=i,Y=j)=P(X=i)P(Y=j\mid X=i)$$

$$=C_3^i\left(\frac{1}{3}\right)^i\left(\frac{2}{3}\right)^{3-i}\cdot C_{3-i}^j\left(\frac{1}{2}\right)^j\left(\frac{1}{2}\right)^{3-i-j}$$

$$j=0,\cdots,3-i;\ i=0,1,2,3;$$

其联合分布与边缘分布如下表所示

| Y \ X ($p_{ij}$) | 0 | 1 | 2 | 3 | $p_{\cdot j}$ |
|---|---|---|---|---|---|
| **0** | $\frac{1}{27}$ | $\frac{1}{9}$ | $\frac{1}{9}$ | $\frac{1}{27}$ | $\frac{8}{27}$ |
| **1** | $\frac{1}{9}$ | $\frac{2}{9}$ | $\frac{1}{9}$ | 0 | $\frac{4}{9}$ |
| **2** | $\frac{1}{9}$ | $\frac{1}{9}$ | 0 | 0 | $\frac{2}{9}$ |
| **3** | $\frac{1}{27}$ | 0 | 0 | 0 | $\frac{1}{27}$ |
| $p_{i\cdot}$ | $\frac{8}{27}$ | $\frac{4}{9}$ | $\frac{2}{9}$ | $\frac{1}{27}$ | 1 |

$$(1)\quad P(X=i\mid Y=0)=\frac{P(X=i,\,Y=0)}{P(Y=0)}$$

$$=\frac{P(X=i,\,Y=0)}{8/27}\qquad i=0,1,2,3$$

将表中第一行数据代入得条件分布

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X=i\mid Y=0)$ | 1/8 | 3/8 | 3/8 | 1/8 |

(2) 当 $X=2$ 时，$Y$ 只可能取 0 与 1.

将表中第三列数据代入下式

$$P(Y=j \mid X=2)=\frac{P(X=2, Y=j)}{2/9}, \quad j=0,1$$

得 $Y$ 的条件分布

| $Y$ | 0 | 1 |
| --- | --- | --- |
| $P(Y=j \mid X=2)$ | 1/2 | 1/2 |

**例2** 已知一射手每次击中目标概率为 $p\,(0<p<1)$, 射击进行到击中两次为止. 令 $X$ 表示首次击中目标所需射击次数, $Y$ 表示总共射击次数. 求 $(X,Y)$ 的联合分布律、条件分布律和边缘分布律.

**解** 由题设知 $X \sim G(p)$, $Y \sim P(2, p)$

故 $X$ 与 $Y$ 的边缘分布律分别为

$$P(X=m)=p(1-p)^{m-1}, \qquad m=1,2,\cdots$$

$$P(Y=n)=(n-1)p^2(1-p)^{n-2}, \quad n=2,3,\cdots$$

$(X, Y)$ 的联合分布律为

$$P(X=m, Y=n)$$

$$=P(X=m)P(Y=n \mid X=m)$$

$$=p(1-p)^{m-1} \cdot p(1-p)^{n-m-1}$$

$$=p^2(1-p)^{n-2}$$

$$m=1,2,\cdots,n-1;\ n=2,3,\cdots$$

$$(m=1,2,\cdots;\ n=m+1,m+2,\cdots)$$

当 $Y=n\ (n=2,3,\cdots)$ 时, $X$ 的条件分布律为

$$P(X=m \mid Y=n)=\frac{P(X=m,Y=n)}{P(Y=n)}$$

$$=\frac{p^2(1-p)^{n-2}}{(n-1)p^2(1-p)^{n-2}}=\frac{1}{n-1}$$

$$m=1,2,\cdots,n-1$$

当$X=m\ (m=1,2,\cdots)$时, $Y$的条件分布律为

$$P(Y=n\,|\,X=m)=\frac{P(X=m,Y=n)}{P(X=m)}$$

$$=\frac{p^2(1-p)^{n-2}}{p(1-p)^{m-1}}=p(1-p)^{n-m-1}$$

$$n=m+1,m+2,\cdots$$

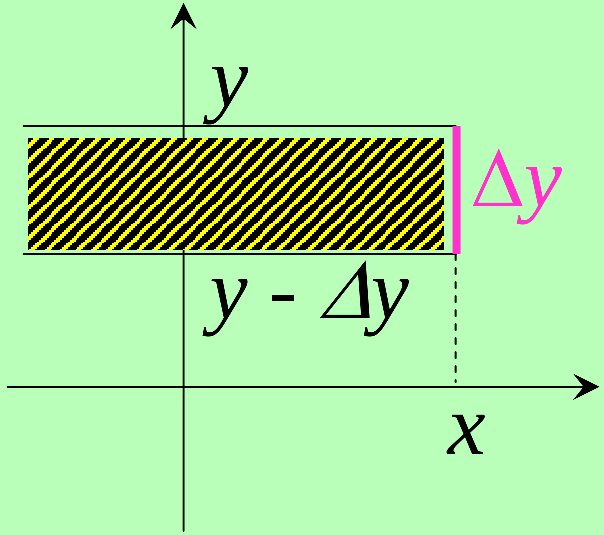
## 二维连续型随机变量的条件分布和条件密度

当$X$连续时，条件分布不能用$P(X=x_i \mid Y=y_j)$来定义，因为$P(X=x_i \mid Y=y_j)\equiv 0$，而应该用$P(X\le x \mid Y=y)$来定义.

设 $\Delta y > 0$

$$P(X \le x \mid y - \Delta y < Y \le y)$$

$$= \frac{P(X \le x, y - \Delta y < Y \le y)}{P(y - \Delta y < Y \le y)}$$

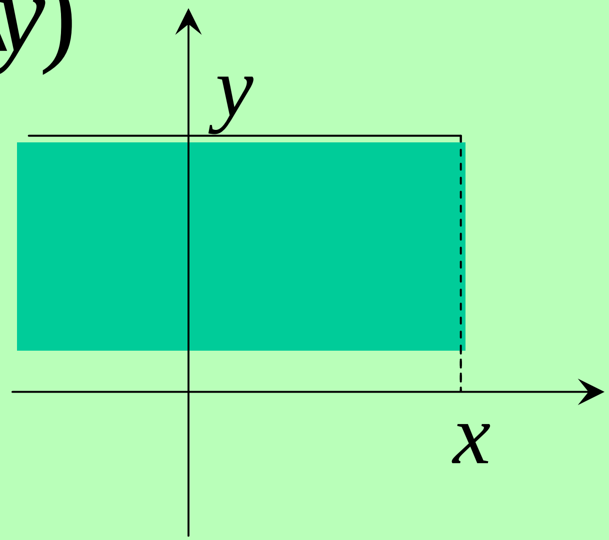



$$= \frac{F(x, y) - F(x, y - \Delta y)}{F_Y(y) - F_Y(y - \Delta y)}$$

$$= \frac{[F(x, y - \Delta y) - F(x, y)] / (-\Delta y)}{[F_Y(y - \Delta y) - F_Y(y)] / (-\Delta y)}$$

$$\lim_{\Delta y\to 0+}\frac{[F(x,y-\Delta y)-F(x,y)]/(-\Delta y)}{[F_Y(y-\Delta y)-F_Y(y)]/(-\Delta y)}$$

$$=\frac{\dfrac{\partial F(x,y)}{\partial y}}{\dfrac{dF_Y(y)}{dy}}=\frac{\int_{-\infty}^{x} f(u,y)du}{f_Y(y)}$$




$f(x,y)$连续

$f_Y(y)\neq 0$,连续

$$\overset{\text{def.}}{=}P(X\le x\mid Y=y)$$

**定义** 若 $f(x,y)$ 在点 $(x,y)$ 连续，$f_Y(y)$ 在点 $y$ 处连续且 $f_Y(y)>0$，则称

$$\frac{\dfrac{\partial F(x,y)}{\partial y}}{\dfrac{dF_Y(y)}{dy}} = \frac{\int_{-\infty}^{x} f(u,y)du}{f_Y(y)} = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)} du$$

为 $Y=y$ 时，$X$ 的条件分布函数，记作

$$F_{X|Y}(x|y) = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)} du$$

称 $$\boxed{f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}}$$

为 $Y=y$ 的条件下 $X$ 的条件 p.d.f.

类似地，称 $F_{Y|X}(y|x)=\int_{-\infty}^{y}\frac{f(x,v)}{f_X(x)}dv$

为 $X=x$ 的条件下 $Y$ 的条件分布函数;

称 $$\boxed{f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}}$$

为 $X=x$ 的条件下 $Y$ 的条件 p.d.f.

## 注意

- $F_{X|Y}(x|y)$, $f_{X|Y}(x|y)$ 仅是 $x$ 的函数, $y$是常数, 对每一 $f_Y(y)>0$ 的 $y$ 处, 只要符合定义的条件, 都能定义相应的函数.

  $F_{Y|X}(y|x)$, $f_{Y|X}(y|x)$ 相仿论述.

- 类似于乘法公式:

$$f(x,y)=f_X(x)f_{Y|X}(y|x) \qquad f_X(x)>0$$

$$=f_Y(y)f_{X|Y}(x|y) \qquad f_Y(y)>0$$

类似于全概率公式

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_{-\infty}^{+\infty} f_{X|Y}(x|y)f_Y(y)dy$$

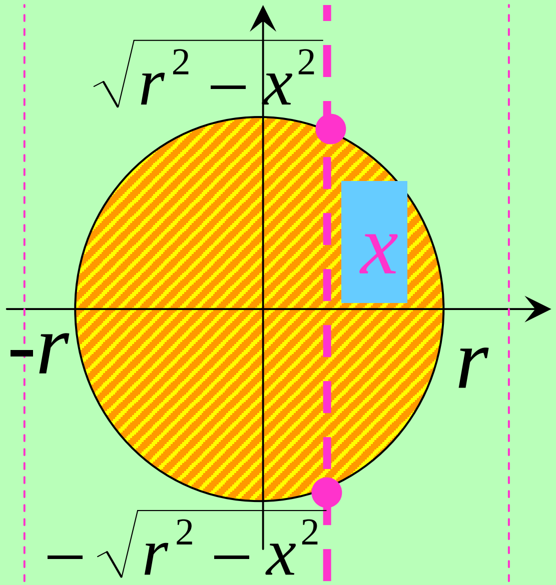
$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_{-\infty}^{+\infty} f_{Y|X}(y|x)f_X(x)dx$$

类似于Bayes公式

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}$$

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\frac{f_{X|Y}(x|y)f_Y(y)}{f_X(x)}$$

**例3** 已知$(X,Y)$服从圆域 $x^2+y^2\le r^2$ 上的均匀分布,求$f_{X|Y}(x|y), f_{Y|X}(y|x)$.




**解**

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

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

$$=\begin{cases}\displaystyle\int_{-\sqrt{r^2-x^2}}^{+\sqrt{r^2-x^2}}\frac{1}{\pi r^2}dy, & -r<x<r\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}\dfrac{2\sqrt{r^2-x^2}}{\pi r^2}, & -r<x<r\\ 0, & \text{其他}\end{cases}$$

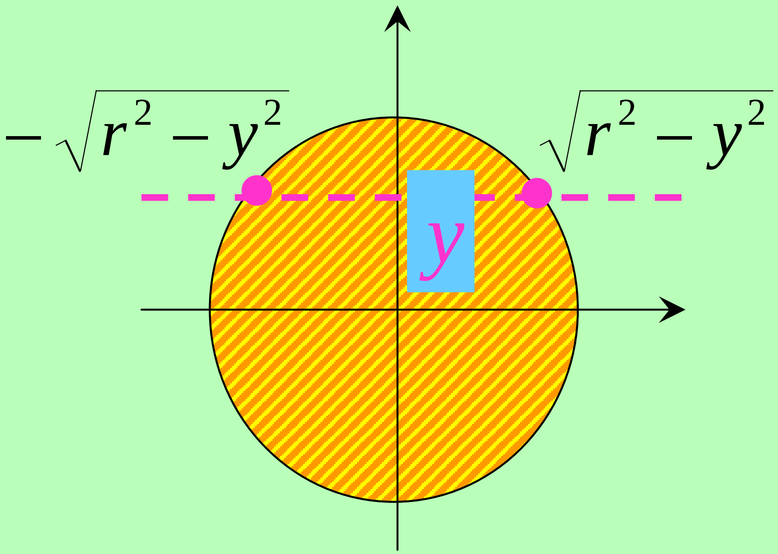
同理，

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

$$= \begin{cases} \dfrac{2\sqrt{r^2 - y^2}}{\pi r^2}, & -r < y < r \\ 0, & \text{其他} \end{cases}$$

边缘分布不是均匀分布！

当 $-r<y<r$ 时，

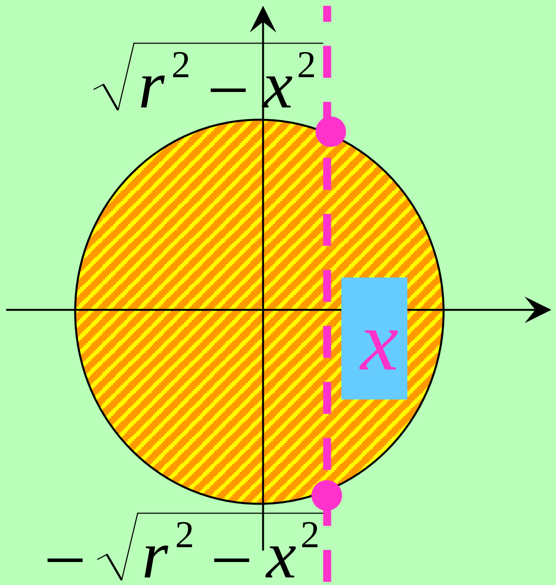



$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}$$

$$=\begin{cases}\dfrac{1}{2\sqrt{r^2-y^2}}, & -\sqrt{r^2-y^2}<x<\sqrt{r^2-y^2}\\ 0, & \text{其他}\end{cases}$$

— 这里 $y$ 是常数，当 $Y=y$ 时，$X\sim U\left(-\sqrt{r^2-y^2},\sqrt{r^2-y^2}\right)$

当 $-r < x < r$ 时,




$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

$$= \begin{cases} \dfrac{1}{2\sqrt{r^2 - x^2}}, & -\sqrt{r^2 - x^2} < y < \sqrt{r^2 - x^2} \\ 0, & \text{其他} \end{cases}$$

— 这里 $x$ 是常数，当 $X = x$ 时，

$$Y \sim U\left(-\sqrt{r^2 - x^2}, \sqrt{r^2 - x^2}\right)$$

**例4** 已知$(X,Y)\sim N\left(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho\right)$

求 $f_{X|Y}(x|y)$

**解**

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}$$

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

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

$$f_{X|Y}(x|y)\sim N\left(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2),\sigma_1^2(1-\rho^2)\right)$$

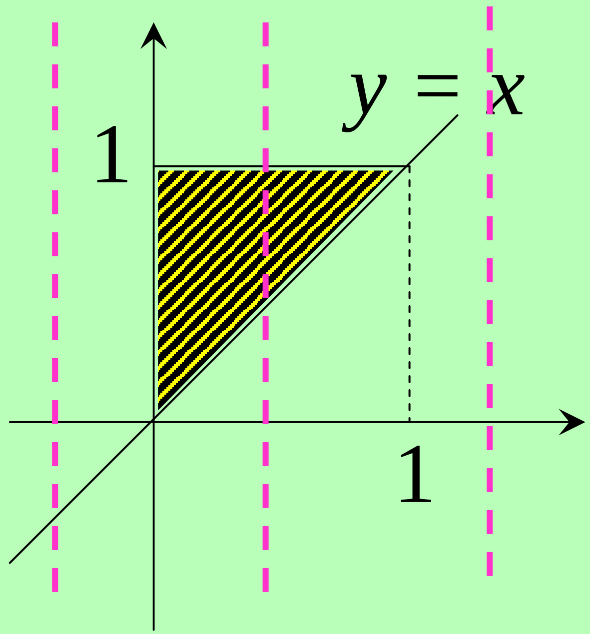
同理，

$$f_{Y|X}(y|x)\sim N\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1),\sigma_2^2(1-\rho^2)\right)$$

**例5** 设

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

求 $f_{X|Y}(x|y)$, $f_{Y|X}(y|x)$



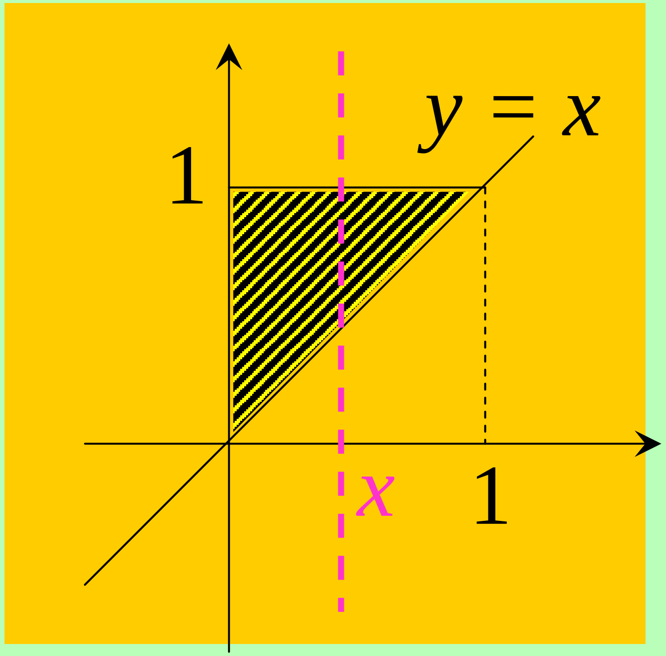

**解**

$$f_X(x)=\begin{cases}\int_x^1 8xydy, & 0\le x\le 1\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}4x(1-x^2), & 0\le x\le 1\\ 0, & \text{其他}\end{cases}$$

$$f_Y(y)=\begin{cases}\int_0^y 8xy\,dx, & 0\le y\le 1\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}4y^3, & 0\le y\le 1\\ 0, & \text{其他}\end{cases}$$




当 $0<y<1$ 时,

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\begin{cases}\dfrac{2x}{y^2}, & 0\le x\le y\\ 0, & \text{其他}\end{cases}$$

当 $0<x<1$ 时,

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\begin{cases}\dfrac{2y}{1-x^2}, & x\le y\le 1\\ 0, & \text{其他}\end{cases}$$

**例6** 已知

$$f_{Y|X}(y|x)=\begin{cases}\dfrac{2y}{1-x^2}, & x\le y\le 1\\ 0, & 其他\end{cases}$$

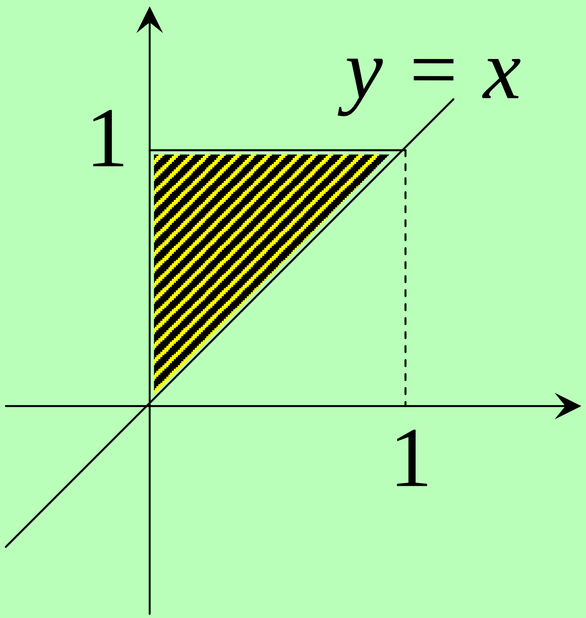
$$f_X(x)=\begin{cases}4x(1-x^2), & 0\le x\le 1\\ 0, & 其他\end{cases}$$

求 $P(X+Y\ge 1),P(Y<0.5),P\left(Y<\dfrac{2}{3}\middle|X=\dfrac{1}{2}\right)$

**解** 当$f_X(x) > 0$时，即 $0 < x < 1$ 时，

$$f(x,y) = f_{Y|X}(y|x) f_X(x)$$

$$= \begin{cases} 8xy, & x \le y \le 1 \\ 0, & \text{其他} \end{cases}$$



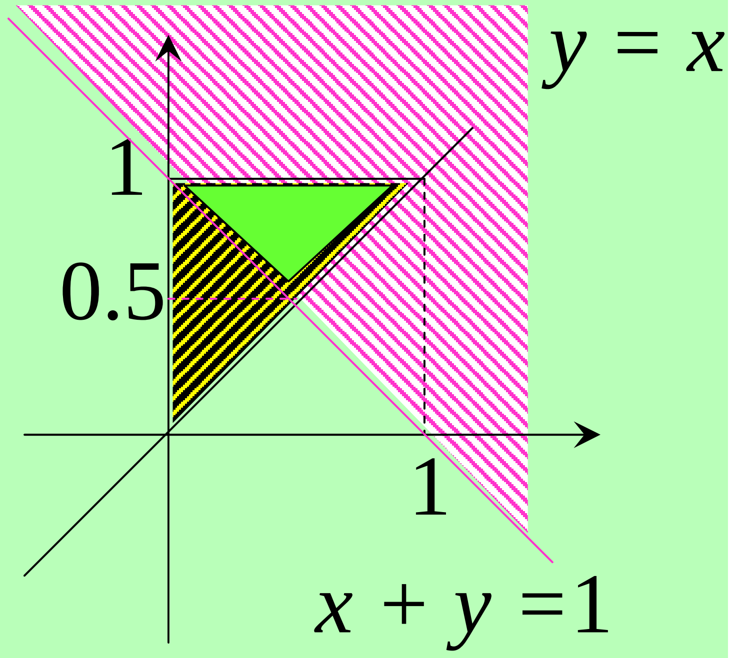

当$f_X(x) = 0$时，$f(x,y) = 0$

故

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

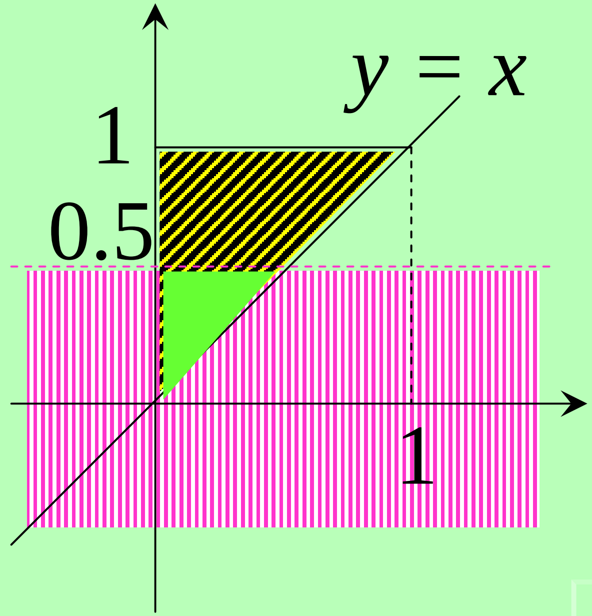
$$P(X+Y\geq 1)$$

$$=\int_{0.5}^{1}dy\int_{1-y}^{y}8xydx=\frac{5}{6}$$




$$P(Y<0.5)$$

$$=\int_{0}^{1/2}dy\int_{0}^{y}8xydx=\frac{1}{16}$$

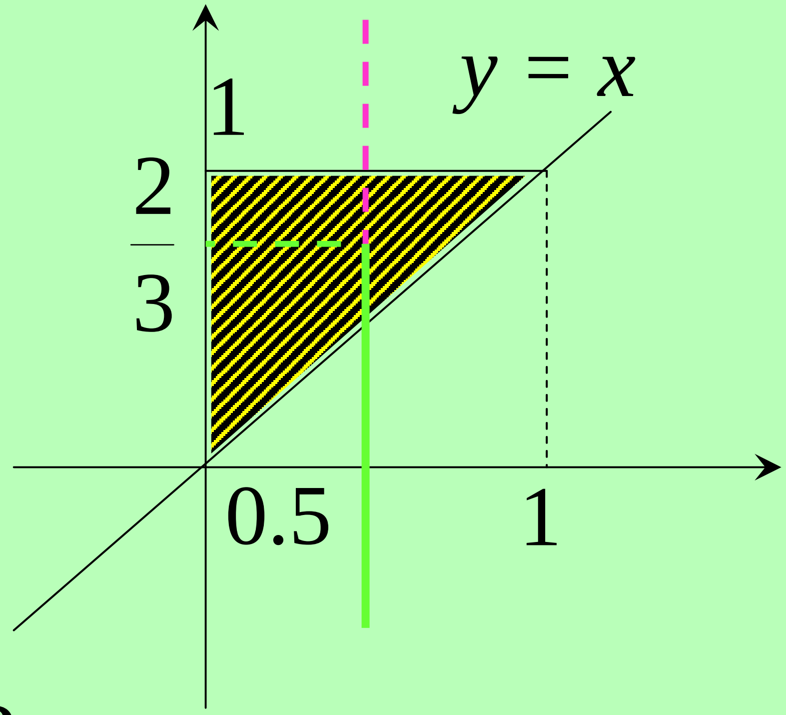


$$P\left(Y<\frac{2}{3}\middle|X=\frac{1}{2}\right)$$

$$=\int_{-\infty}^{2/3} f_{Y|X}\left(y\,\middle|\,\frac{1}{2}\right)dy$$

$$=\int_{1/2}^{2/3}\frac{2y}{1-(0.5)^2}dy=\int_{1/2}^{2/3}\frac{8y}{3}dy$$

$$=\frac{7}{27}$$



# 由脚印估计罪犯身高

公安人员根据收集到的罪犯脚印，通过公式

$$\boxed{\text{身高}=\text{脚印长度}\times 6.876}$$

算出罪犯的身高. 这个公式是如何推导出来的?

设一个人身高为$X$，脚印长度为$Y$. 显然，两者之间是有统计关系的，故应作为二维随机变量$(X,Y)$来研究.

由于影响人类身高与脚印的随机因素是大量的、相互独立的，且各因素的影响又是微小的，可以叠加的. 故由中心极限定理知$(X,Y)$可以近似看成服从二维正态分布 $N(u_1,\sigma_1^2,u_2,\sigma_2^2;\rho)$.

其中参数 $u_1,\sigma_1^2; u_2,\sigma_2^2; \rho$ 因区域、民族、生活习惯的不同而有所变化，但它们都能通过统计方法而获得.

现已知罪犯的脚印长度为 $y$，要估计其身高就需计算条件期望，条件密度为

$$f_{X|Y}(x\,|\,y) = \frac{f(x,y)}{f_Y(y)}$$

$$
=\frac{\sqrt{2\pi}\sigma_2}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot\frac{\exp\{\frac{1}{2(1-\rho^2)}[\frac{(x-u_1)^2}{\sigma^2}-2\rho\frac{(x-u_1)(y-u_2)}{\sigma_1\sigma_2}+\frac{(y-u_2)^2}{\sigma_2^2}]\}}{\exp\{\frac{(y-u_2)^2}{2\sigma_2^2}\}}
$$

这正是正态分布 $N(u_1+\rho\frac{\sigma_1}{\sigma_2}(y-u_2),\sigma_1^2(1-\rho^2))$ 的密度函数，因此

$$
E(X\mid Y=y)=u_1+\rho\frac{\sigma_1}{\sigma_2}(y-u_2)
$$

如果按中国人的相应参数代入上式，即可得出以脚印长度作自变量的身高近似公式.

作业 P.133习题三

16　　17

第8周

# 问题

设随机变量 Z 服从参数为 1 的指数分布，引入随机变量：

$$X=\begin{cases}0 & Z\le 1\\ 1 & Z>1\end{cases}\qquad Y=\begin{cases}0 & Z\le 2\\ 1 & Z>2\end{cases}$$

求 $(X,Y)$ 的联合分布律和联合分布函数.