§3.3 条件分布

- 条件分布律
- 条件分布函数
- 条件概率密度

在第一章中，我们介绍了条件概率的概念 .

在事件*B*发生的条件下事件*A*发生的概率

$$P(A|B) = \frac{P(AB)}{P(B)}$$

推广到随机变量

设有两个随机变量 *X*, *Y*， 在给定 *Y* 取某个值的条件下，求 *X* 的概率分布.

这个分布就是条件分布.

1. 离散型随机变量的条件分布律

设 (X,Y) 是离散型随机变量，其分布律为

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

(X, Y)关于 X 和关于 Y 的边缘分布律分别为：

$$P(X=x_i)=p_{i\cdot}=\sum_{j=1}^{\infty}p_{ij},\quad i=1,2,\cdots$$

$$P(Y=y_j)=p_{\cdot j}=\sum_{i=1}^{\infty}p_{ij},\quad j=1,2,\cdots$$

重要

由条件概率公式自然地引出如下定义：

定义： 设(X,Y)是二维离散型随机变量，对于固定的j，若$P(Y=y_j)>0$, 则称

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

为在$Y=y_j$条件下随机变量X的条件分布律 **(Conditional Density)**.

同样对于固定的 i, 若 $P(X=x_i)>0$, 则称

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

为在 $X=x_i$ 条件下随机变量 Y 的条件分布律.

条件分布是一种概率分布，它具有概率分布的一切性质.

例如： $\boldsymbol{P(X=x_i \mid Y=y_j)\geq 0},$

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

例 设随机变量 X 在 1,2,3三个数中等可能地取值，另一个随机变量 Y 在1~X 中等可能地取一整数值，求Y=1时， X 的条件分布律。

X \ Y	1	2	3	$p_{i\cdot}$
1	$\frac{1}{3}$	0	0	$\frac{1}{3}$
2	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{3}$
3	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{3}$
$p_{\cdot j}$	$\frac{11}{18}$	$\frac{5}{18}$	$\frac{2}{18}$	1

联合分布与边缘分布

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

$$=\frac{P(X=i,\,Y=1)}{11/18}\qquad i=1,2,3$$

将表中第一列数据代入得条件分布

X	1	2	3
$P(X=i\mid Y=1)$	$6/11$	$3/11$	$2/11$

2. 连续型随机变量的条件分布

设 (X, Y) 是二维连续型随机变量，由于

$$P(X=x)=0,\ P(Y=y)=0$$

所以不能直接代入条件概率公式，先利用极限的方法来引入条件分布函数的概念。

定义：给定 y，设对于任意固定的正数ε，

$P(y-\varepsilon<Y\le y+\varepsilon)>0$, 若对于任意实数 x，

$$\lim_{\varepsilon\to 0^{+}} P(X\le x \mid y-\varepsilon<Y\le y+\varepsilon)$$

$$=\lim_{\varepsilon\to 0^{+}} \frac{P(X\le x, y-\varepsilon<Y\le y+\varepsilon)}{P(y-\varepsilon<Y\le y+\varepsilon)}$$

存在，则称其为在条件$Y=y$下X的条件分布函数，记为 $F_{X|Y}(x|y)=P(X\le x|Y=y)$.

由条件分布函数可以引出条件概率密度 (Conditional Density)

$$F_{X|Y}(x\mid y) = \lim_{\varepsilon\to 0^+}\frac{P\{X\le x, y-\varepsilon < Y\le y+\varepsilon\}}{P\{y-\varepsilon < Y\le y+\varepsilon\}}$$

$$= \lim_{\varepsilon\to 0^+}\frac{F(x,y+\varepsilon)-F(x,y-\varepsilon)}{F_Y(y+\varepsilon)-F_Y(y-\varepsilon)}$$ <负无穷概率分布为0，省略

$$= \frac{\lim\limits_{\varepsilon\to 0^+}[F(x,y+\varepsilon)-F(x,y-\varepsilon)]/2\varepsilon}{\lim\limits_{\varepsilon\to 0^+}[F_Y(y+\varepsilon)-F_Y(y-\varepsilon)]/2\varepsilon}$$

ξ趋近于0，相当于y处求导

$$= \frac{\dfrac{\partial F(x,y)}{\partial y}}{\dfrac{d}{dy}F_Y(y)} = \frac{\dfrac{\partial}{\partial y}\left(\displaystyle\int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)\,du\,dv\right)}{f_Y(y)} = \frac{\int_{-\infty}^{x} f(u,y)\,du}{f_Y(y)},$$

在条件$Y=y$下X的**条件分布函数**

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

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

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

称为随机变量X在条件$Y=y$下的**条件密度函数**

重要

定义

$f_Y(y) > 0$时，X在$Y = y$条件下的条件密度函数为

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

$f_X(x) > 0$时，Y在$X = x$条件下的条件密度函数为

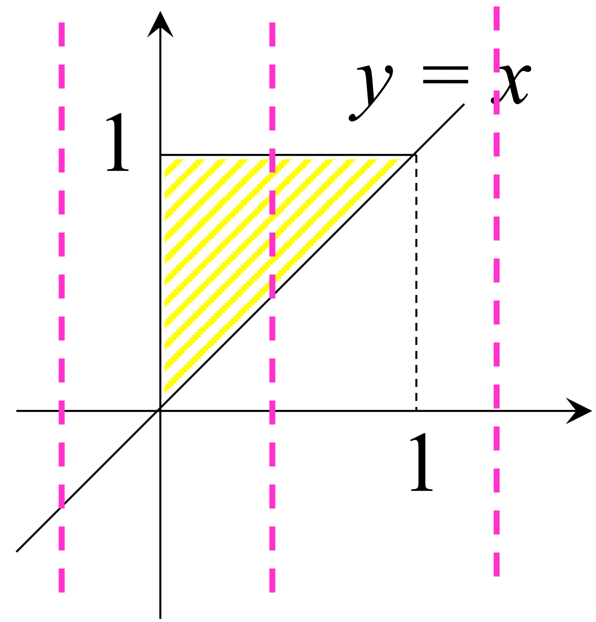
$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

例 设

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

求 $f_{X|Y}(x|y), f_{Y|X}(y|x)$

解

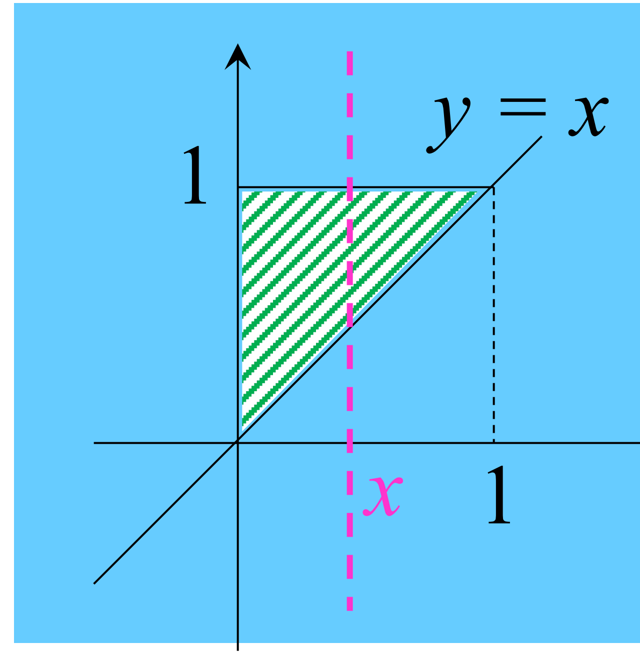


$$f_X(x)\square\begin{cases}\int_x^1 8xy\,dy, & 0\le x\le 1\\ 0, & 其他\end{cases}$$

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

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

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



当$0 < y \le 1$时，

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

当$0 \le x < 1$时，

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

例 $f_{Y|X}(y|x) \square \begin{cases} \dfrac{2y}{1-x^2}, & x \le y \le 1 \\ 0, & \text{其他} \end{cases}$ $\quad f_X(x) \square \begin{cases} 4x(1-x^2), & 0 \le x \le 1 \\ 0, & \text{其他} \end{cases}$

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

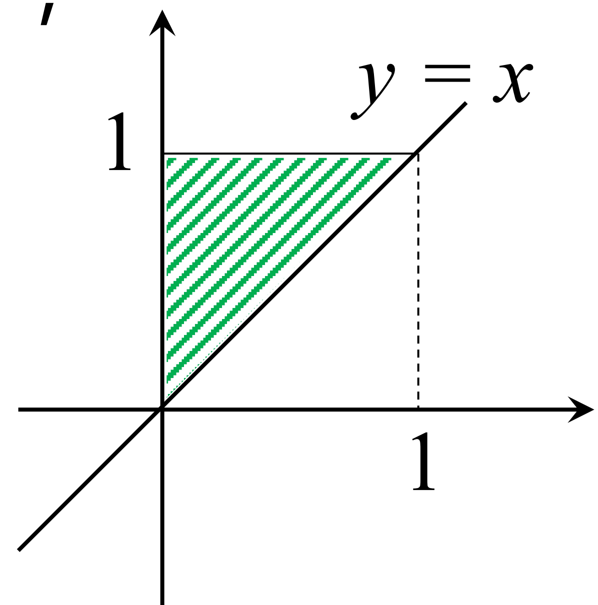
解

当 $f_X(x) > 0$ 时，即 $0 < x < 1$ 时，

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

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

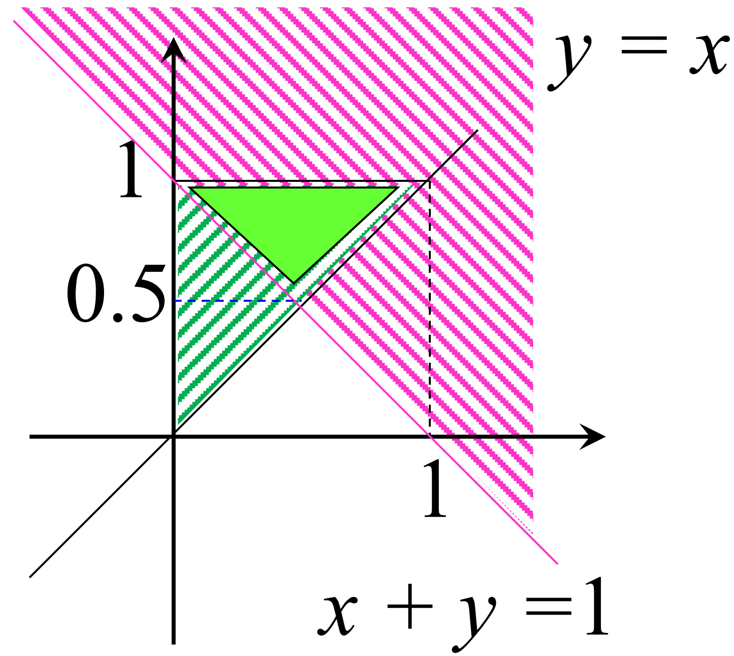
当 $f_X(x) = 0$ 时，$f(x,y) = 0$



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

(1)

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



$$\square \iint f(x,y)dxdy \square \int_{0.5}^{1} dy \int_{1-y}^{y} 8xydx \quad \square \frac{5}{6}$$

能不能用面积比值？为什么？

(2)

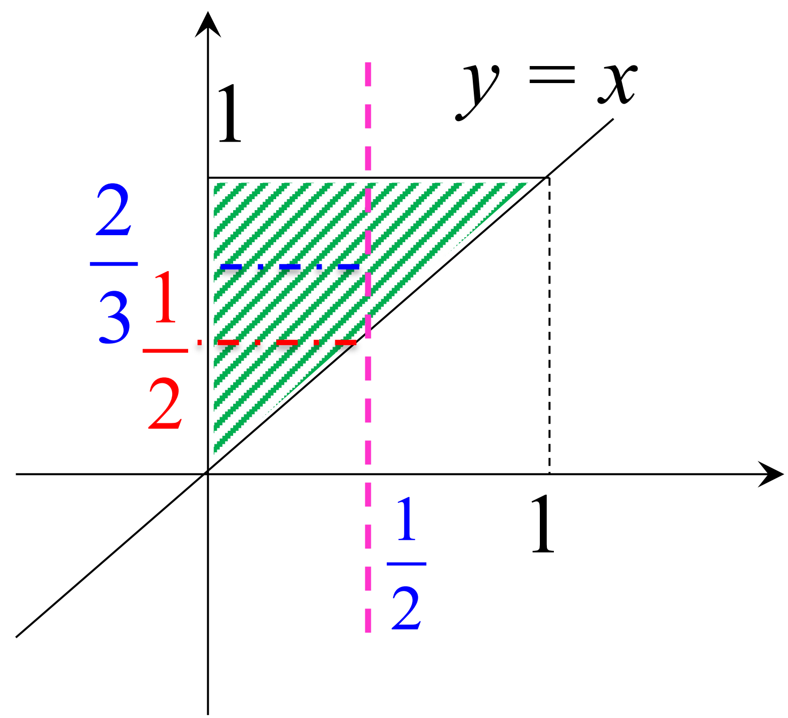
$$f_{Y|X}(y|x) = \begin{cases} \dfrac{2y}{1-x^2}, & {\color{red}x \le y} \le 1 \\ 0, & \text{其他} \end{cases}$$

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

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

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

$$= \int_{1/2}^{2/3} \frac{8y}{3} dy = \frac{7}{27}$$



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

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，

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

总结

1. 离散型条件分布
2. 连续型条件分布
   a. 极限定义
   b. 求条件密度
      画图确定密度非0区域
      边缘化
      条件密度公式（注意定义域）