§3 条件分布

- 条件分布律
- 条件分布函数
- 条件概率密度

返回主目录

一、离散型随机变量的条件分布律

设 (X,Y) 是二维离散型随机变量，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：

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

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

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

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

为在$Y=y_j$条件下随机变量X的条件分布律。

条件分布律具有分布律的以下特性：

1^0 $P\{X=x_i \mid Y=y_j\}\geq 0$；

2^0 $\sum\limits_{i=1}^{\infty}P\{X=x_i \mid Y=y_j\}=\sum\limits_{i=1}^{\infty}\frac{p_{ij}}{p_{\bullet j}}=\frac{1}{p_{\bullet j}}\sum\limits_{i=1}^{\infty}p_{ij}=\frac{p_{\bullet j}}{p_{\bullet j}}=1.$

同样对于固定的 i, 若$P\{X=x_i\}>0$, 则称

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

为在 $X=x_i$ 条件下随机变量Y的条件分布律。

例 1

一射手进行射击，击中目标的概率为 p，射击到击中目标两次为止。设以 X 表示首次击 中目标所进行的射击次数，以 Y 表示总共进行 的射击次数，试求 X 和 Y 的联合分布律以及条件分布律。

解：

Y的取值是 2， 3， 4， …； X的取值是 1， 2， …，

并且 $X<Y$.

返回主目录

X,Y的联合分布律为

$$P\{X=m,\quad Y=n\}$$

$$=P\{\text{第}m\text{次射击时首次击中目标，并且共射击}n\text{次}\}$$

$$=P\{\text{第}m\text{次射击时首次击中目标，且第}n\text{次射击时第二次命中目标}$$

由独立性，可得

$$P\{X=m,\quad Y=n\}=q^{m-1}\cdot p\cdot q^{n-m-1}\cdot p\ =q^{n-2}\cdot p^2$$

$$(\text{其中}q=1-p)(n=2,\ 3,\ \cdots;\ \ m=1,\ 2,\ \cdots,\ \ n-1)$$

X的边缘分布律为

$$P\{X=m\}=\sum_{n=m+1}^{\infty}P\{X=m,Y=n\}=\sum_{n=m+1}^{\infty}p^2q^{n-2}$$

$$=p^2\sum_{n=m+1}^{\infty}q^{n-2}=p^2\bullet\frac{q^{m-1}}{1-q}=pq^{m-1},\quad m=1,2,\cdots$$

Y的边缘分布律为

$$P\{Y=n\}=\sum_{m=1}^{n-1}P\{X=m,Y=n\}=\sum_{m=1}^{n-1}p^2q^{n-2}=(n-1)p^2q^{n-2},$$

$n=2,3,\cdots$

在$Y=n$ 条件下随机变量 X 的条件分布律为

当 $n=2,3,\ldots$ 时，

$$P\{X=m\mid Y=n\}=\frac{p^2q^{n-2}}{(n-1)p^2q^{n-2}}=\frac{1}{n-1},\quad m=1,2,\cdots,n-1;$$

$$P\{X=m,\quad Y=n\}=q^{m-1}\cdot p\cdot q^{n-m-1}\cdot p=q^{n-2}\cdot p^2$$
$$(n=2,\ 3,\ \cdots;\ m=1,\ 2,\ \cdots,\ n-1)$$

在 $X=m$ 条件下随机变量Y的条件分布律为

当$m=1,2,3,\ldots$ 时,

$$P\{Y=n\mid X=m\}=\frac{P\{X=m,Y=n\}}{P\{X=m\}}=\frac{p^2q^{n-2}}{pq^{m-1}}$$

$$=pq^{n-m-1},\qquad n=m+1,m+2,\cdots$$

$$P\{X=m,\quad Y=n\}=q^{m-1}\cdot p\cdot q^{n-m-1}\cdot p\ =q^{n-2}\cdot p^2$$

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

二、条件分布函数

设（ X , Y ）是二维连续型随机变量，由于

$$P\{X=x_i\}=0,\quad P\{Y=y_j\}=0,$$

不能直接代入条件概率公式，我们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，极限

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

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

返回主目录

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

$$=\lim_{\varepsilon\to 0^+}\frac{F(x,y+\varepsilon)-F(x,y-\varepsilon)}{F_Y(y+\varepsilon)-F_Y(y-\varepsilon)}$$

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

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

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

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

称为在条件$Y = y$下X的条件分布函数,

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

称为随机变量X在$Y = y$的条件下的条件密度函数。

三、连续型随机变量的条件密度函数

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：

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

随机变量Y的边缘密度函数为：

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

则当$f_Y(y)>0$时，可得随机变量X在$Y=y$的条件下的条件密度函数为

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

当$f_X(x)>0$时，可得随机变量Y在$X=x$的条件下的条件密度函数为

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

条件密度函数的性质

性质 1. 对任意的x，有 $f_{X|Y}(x|y) \geq 0$

性质 2. $\int_{-\infty}^{+\infty} f_{X|Y}(x|y)dx = 1$

简言之，$f_{X|Y}(x|y)$是密度函数.

对于条件密度函数 $f_{Y|X}(y|x)$也有类似的性质.

例 2

§3条件分布

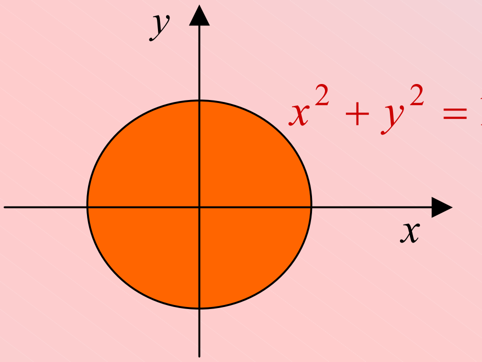
设二维随机变量$(X,\ Y)$服从圆域：$x^2+y^2\le 1$

上的均匀分布，试求条件密度函数$f_{X|Y}(x|y)$.

解：

二维随机变量$(X,\ Y)$的联合密度函数为

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



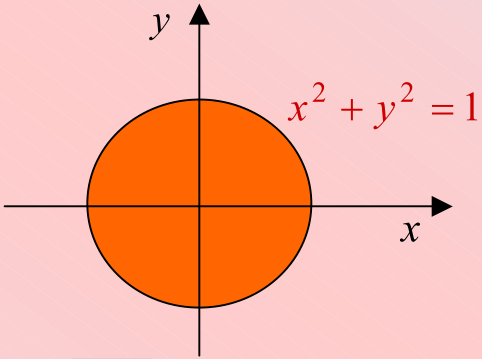
例 2（续）

由此得，当$-1\le y\le 1$时，

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,\ y)dx=\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\frac{1}{\pi}dx=\frac{2}{\pi}\sqrt{1-y^2}$$

所以，随机变量Y的密度函数为

$$f_Y(y)=\begin{cases}\dfrac{2}{\pi}\sqrt{1-y^2} & -1\le y\le 1\\ 0 & 其它\end{cases}$$



由此得，当$-1<y<1$时，$f_Y(y)>0$

例 2（续）

因此当 $-1<y<1$ 时，

$$f_{X|Y}(x|y)=\frac{f(x,\quad y)}{f_Y(y)}=\frac{\frac{1}{\pi}}{\frac{2}{\pi}\sqrt{1-y^2}}=\frac{1}{2\sqrt{1-y^2}}$$

所以，

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

即当 $-1<y<1$ 时，X 在 $Y=y$ 下的条件分布是区间 $\left[-\sqrt{1-y^2},\quad \sqrt{1-y^2}\right]$ 上的均匀分布.

例 3

设二维随机变量$(X,\ Y)$服从二元正态分布：

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

则$(X,\ Y)$的联合密度函数为

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

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

例 3（续）

又随机变量Y的边缘密度函数为

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

因此，对任意的y，$f_Y(y) > 0$，　有

$$f_{X|Y}(x|y) = \frac{f(x,\quad y)}{f_Y(y)} = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2(1-r^2)}}$$

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

例 3（续）

这表明，二元正态分布的条件分布是一元正态分布：

$$N\left(\mu_1 + r\frac{\sigma_1}{\sigma_2}(y-\mu_2),\quad \sigma_1^2(1-r^2)\right)$$

例 4

设随机变量 X 服从区间 $(0, 1)$ 上的均匀分布，当 $0<x<1$ 时，随机变量 Y 在 $X=x$ 的条件下服从区间 $(x, 1)$ 上的均匀分布．试求随机变量 Y 的密度函数．

解：

随机变量 X 的密度函数为

$$f_X(x)=\begin{cases}1 & 0<x<1\\ 0 & \text{其它}\end{cases}$$

返回主目录

例 4（续）

又由题设，知当$0<x<1$时，随机变量Y在条件$X=x$下的条件密度函数为

$$f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{1-x} & x<y<1\\ 0 & \text{其它}\end{cases}$$

所以，由公式

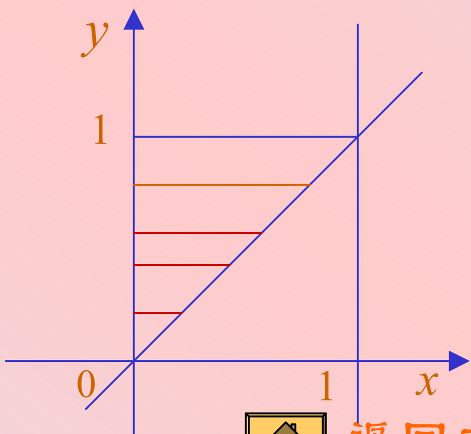
$$f_{Y|X}(y|x)=\frac{f(x,\ y)}{f_X(x)}$$

例 4（续）

$$f(x,\quad y)=f_X(x)f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{1-x} & 0<x<y<1\\ 0 & \text{其它}\end{cases}$$

所以，当$0<y<1$时，

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,\quad y)dx=\int_0^y\frac{1}{1-x}dx=-\ln(1-y)$$



所以，随机变量Y的密度函数为

$$f_Y(y)=\begin{cases}-\ln(1-y) & 0<y<1\\ 0 & \text{其它}\end{cases}$$

§3条件分布

例 5

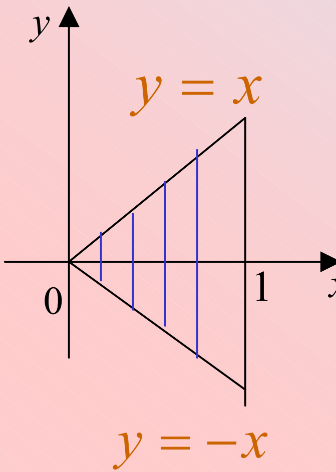
设随机变量 (X,Y)的概率密度为

$$f(x,y)=\begin{cases}1, |y|<x, 0<x<1,\\ 0, 其它.\end{cases}$$

试求：(1) $f_X(x), f_Y(y)$；(2) $f_{X|Y}(x\mid y), f_{Y|X}(y\mid x)$；

(3) $P\{X>\frac{1}{2}\mid Y>0\}$.

解：



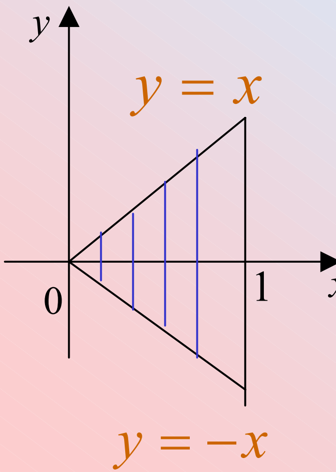
(1) $$f_X(x)=\int_{-\infty}^{\infty}f(x,y)dy=\begin{cases}\int_{-x}^{x}dy=2x, 0<x<1,\\ 0, \qquad 其它.\end{cases}$$

§3条件分布

例 5（续）

$$f_Y(y)=\int_{-\infty}^{\infty} f(x,y)dx=\begin{cases}\int_y^1 dx=1-y,\ 0\le y<1,\\ \int_{-y}^1 dx=1+y,\ -1\le y<0,\\ 0, \qquad\qquad 其它.\end{cases}$$

$$=\begin{cases}1-|y|, & |y|<1\\ 0, & 其它.\end{cases}$$



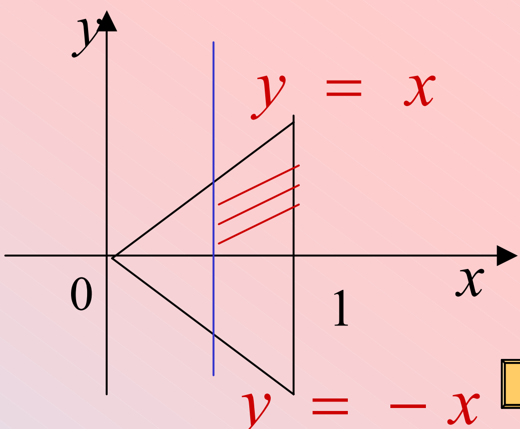
(2) 当 $|y|<1$, $f_{X|Y}(x\,|\,y)=\dfrac{f(x,y)}{f_Y(y)}=\begin{cases}\dfrac{1}{1-|y|}, & |y|<x<1\\ 0, & 其它。\end{cases}$

例 5（续）

当$0 < x < 1$，$f_{Y|X}(x \mid y) = \dfrac{f(x,y)}{f_X(x)} = \begin{cases} \dfrac{1}{2x}, & |y| < x \\ 0, & \text{其它。} \end{cases}$

$$(3).\ P\{X > \frac{1}{2} \mid Y > 0\} = \frac{P\{X > \frac{1}{2}, Y > 0\}}{P\{Y > 0\}} \qquad x = \frac{1}{2}$$

$$= \frac{(1 + \frac{1}{2}) \times \frac{1}{2} \div 2}{\frac{1}{2} \times 1 \times 1} = \frac{3}{4}$$



返回主目录