第三节　　条件分布

一、离散型随机变量的条件分布

二、连续型随机变量的条件分布

三、小结

一、离散型随机变量的条件分布

问题 考虑一大群人,从其中随机挑选一个人,分别用 X 和 Y 记此人的体重和身高,则X 和 Y 都是随机变量,他们都有自己的分布.

现在如果限制 Y 取值为 1.6米, 在这个限值下求X的分布。

设 (X,Y) 是二维离散型随机变量，**其分布律为**

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

(X,Y)关于X和关于Y的边缘分布律为

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

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

设$p_{\bullet j}>0$，我们来考虑在事件 $\{Y=y_j\}$已发生的条件下事件$\{X=x_i\}$发生的概率 $P\{X=x_i|Y=y_j\}$

由条件概率公式, 可得

$$P\{X = x_i \mid Y = y_j\} = \frac{P\{X = x_i, Y = y_j\}}{P\{Y = y_j\}}$$

$$= \frac{p_{ij}}{p_{\bullet j}}, \quad i = 1,2,\cdots.$$

条件概率具有分布律的性质:

$1^\circ \quad P\{X = x_i \mid Y = y_j\} \geq 0;$

$$2^\circ \quad \sum_{i=1}^{\infty} P\{X = x_i \mid Y = y_j\} = \sum_{i=1}^{\infty} \frac{p_{ij}}{p_{\bullet j}} = \frac{1}{p_{\bullet j}} \sum_{i=1}^{\infty} p_{ij}$$

$$= \frac{p_{\bullet j}}{p_{\bullet j}} = 1.$$

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**

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

为在 $Y=y_j$ 条件下随机变量 X 的条件分布律.

同样，对于固定的 i，若 $P\{X=x_i\}>0$，**则称**

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

为在 $X=x_i$ 条件下随机变量 Y 的条件分布律 .

例1 在一汽车工厂中，一辆汽车有两道工序是由机器人完成的.其一是紧固3只螺栓,其二是焊接2处焊点. 以 X 表示螺栓紧固得不良的 数目,以Y 表**示由机器人焊接的不良焊点的数目. 据积累的资料**知 (X,Y) 具有分布律：

$Y \backslash X$	0	1	2	3	$P\{Y=j\}$
0	0.840	0.030	0.020	0.010	0.900
1	0.060	0.010	0.008	0.002	0.080
2	0.010	0.005	0.004	0.001	0.020
$P\{X=i\}$	0.910	0.045	0.032	0.013	1.000

(1) 求在 $X=1$ 的条件下, Y 的条件分布律；

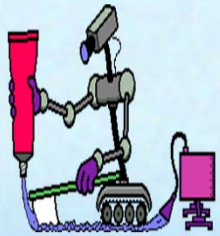
(2) 求在 $Y=0$ 的条件下, X 的条件分布律 .

解　边缘分布已经求出列在上表中.

在$X=1$的条件下,Y的条件分布律为

$$P\{Y=0|X=1\}=\frac{P\{X=1,Y=0\}}{P\{X=1\}}=\frac{0.030}{0.045},$$

$$P\{Y=1|X=1\}=\frac{P\{X=1,Y=1\}}{P\{X=1\}}=\frac{0.010}{0.045},$$



$$P\{Y=2 \mid X=1\}=\frac{P\{X=1,Y=2\}}{P\{X=1\}}=\frac{0.005}{0.045},$$

或写成

$Y=k$	0	1	2
$P\{Y=k \mid X=1\}$	$\frac{6}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

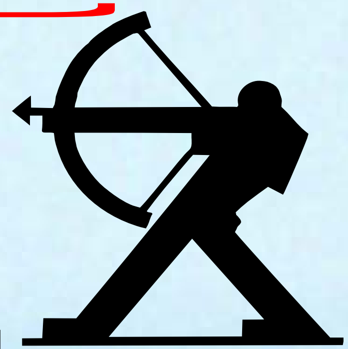
同样可得在 $Y=0$ 的条件下 X 的条件分布律为

$X=k$	0	1	2	3
$P\{X=k \mid Y=0\}$	$\frac{84}{90}$	$\frac{3}{90}$	$\frac{2}{90}$	$\frac{1}{90}$

例2一射手进行射击,击中目标的概率为$p(0 < p < 1)$,**射击直至击中目标两次为止**.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的联合分布律及条件分布律.

解 由题意知 X 取 m 且 Y 取 n时, 有

$$P\{X = m, Y = n\} = p \cdot p \cdot \underbrace{(1-p)\cdot(1-p)\cdots(1-p)}_{(n-2)\text{个}}$$



即得 X 和 Y 的联合分布律为

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

其中$q = 1 - p, n = 2,3,\cdots;\quad m = 1,2,\cdots,n-1.$

现在求条件分布律.

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

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

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

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

$$P\{Y=n\}=\sum_{m=1}^{n-1}P\{X=m,Y=n\}$$

$$=\sum_{m=1}^{n-1}p^2q^{n-2}=(n-1)p^2q^{n-2},\quad n=2,3,\cdots.$$

所以当 $n=2,3,\cdots$ 时,

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

$$=\frac{1}{n-1}, \quad m=1,2,\cdots,n-1,$$

当 $m=1,2,\cdots$,时,

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

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

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

二、连续型随机变量的条件分布

定义 设二维随机变量 (X,Y) 的概率密度为 $f(x,y)$，(X,Y) 关于 Y 的边缘概率密度为 $f_Y(y)$. **若**对于固定的 y, $f_Y(y)>0$,则称 $\dfrac{f(x,y)}{f_Y(y)}$ 为在 $Y=y$ 的条件下 X 的条件概率密度, **记为** $f_{X|Y}(x|y)=\dfrac{f(x,y)}{f_Y(y)}$.

称 $\int_{-\infty}^{x} f_{X|Y}(x|y)\,\mathrm{d}x=\int_{-\infty}^{x}\dfrac{f(x,y)}{f_Y(y)}\,\mathrm{d}x$ 为在 $Y=y$ 的条

件下X 的条件分布函数,记为$P\{X \le x | Y = y\}$或

$F_{X|Y}(x|y)$, **即**

$$F_{X|Y}(x|y) = P\{X \le x | Y = y\} = \int_{-\infty}^{x} \frac{f(x,y)}{f_Y(y)} \mathrm{d}x.$$

类似地, 可以定义 $f_{Y|X}(y|x) = \dfrac{f(x,y)}{f_X(x)}$和

$$F_{Y|X}(y|x) = \int_{-\infty}^{y} \frac{f(x,y)}{f_X(x)} \mathrm{d}y.$$

请同学们思考

为什么不能用条件概率的定义来直接定义条件分布函数 $F_{X|Y}(x|y)$?

答 条件分布是指在一个随机变量取某个确定值的条件下,另一个随机变量的分布,即

$$F_{X|Y}(x|y)=P\{X\le x|Y=y\}.$$

由于$P\{Y=y\}$可能为零(连续型时一定为零).故直接用条件概率来定义时,会出现分母为零.

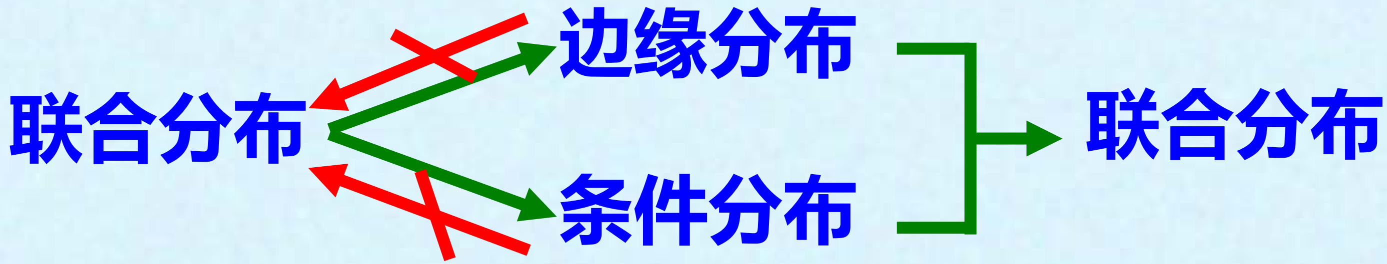
条件分布函数与条件密度函数的关系

$$F_{X|Y}(x|y)=\int_{-\infty}^{x} f_{X|Y}(x|y)\,\mathrm{d}x = P\{X \le x | Y = y\}$$

$$F_{Y|X}(y|x)=\int_{-\infty}^{y} f_{Y|X}(y|x)\,\mathrm{d}y = P\{Y \le y | X = x\}$$

说明

联合分布、边缘分布、条件分布的关系如下



例3 设 G 是平面上的有界区域，其面积为 A. 若二维随机变量 (X,Y) 具有概率密度

$$f(x,y)=\begin{cases}\dfrac{1}{A}, & (x,y)\in G,\\ 0, & \text{其他}.\end{cases}$$

则称(X,Y)在G上服从均匀分布．**现设二维随机变量在圆域** $x^2+y^2\le 1$ 上服从均匀分布，**求条件概率**密度 $f_{X|Y}(x|y)$.

解 由假设随机变量 (X,Y) 具有概率密度

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

且有边缘概率密度

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

$$=\begin{cases}\dfrac{1}{\pi}\displaystyle\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\mathrm{d}\,x=\dfrac{2}{\pi}\sqrt{1-y^2}, & -1\le y\le 1,\\ 0, & \text{其他}.\end{cases}$$

于是当$-1<y<1$时, 有

$$f_{X|Y}(x|y)$$

$$=\begin{cases} \dfrac{\frac{1}{\pi}}{\frac{2}{\pi}\sqrt{1-y^2}} = \dfrac{1}{2\sqrt{1-y^2}}, & -\sqrt{1-y^2} \le x \le \sqrt{1-y^2}, \\ 0, & \text{其他}. \end{cases}$$

当$y=0$和$y=\dfrac{1}{2}$时$f_{X|Y}(x|y)$的图形分别如图$3-6$, $3-7$所示.

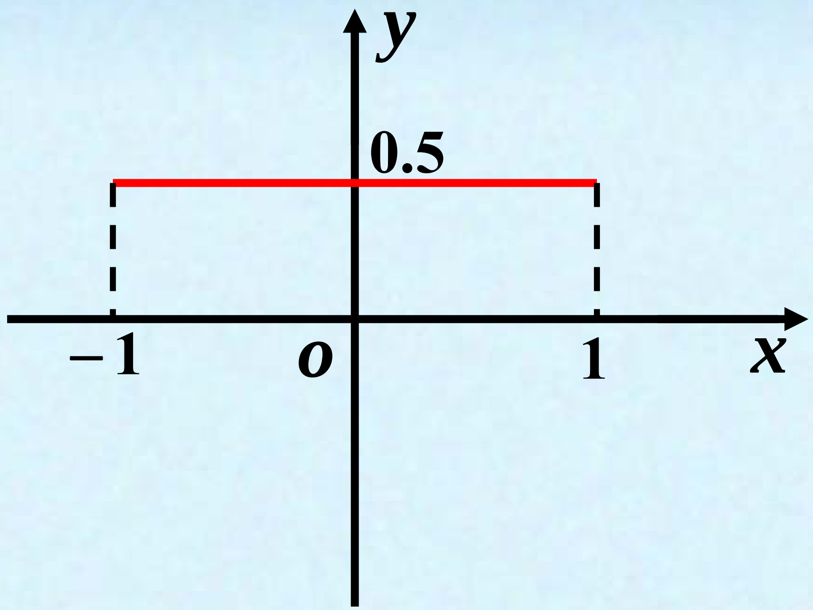


图3—6

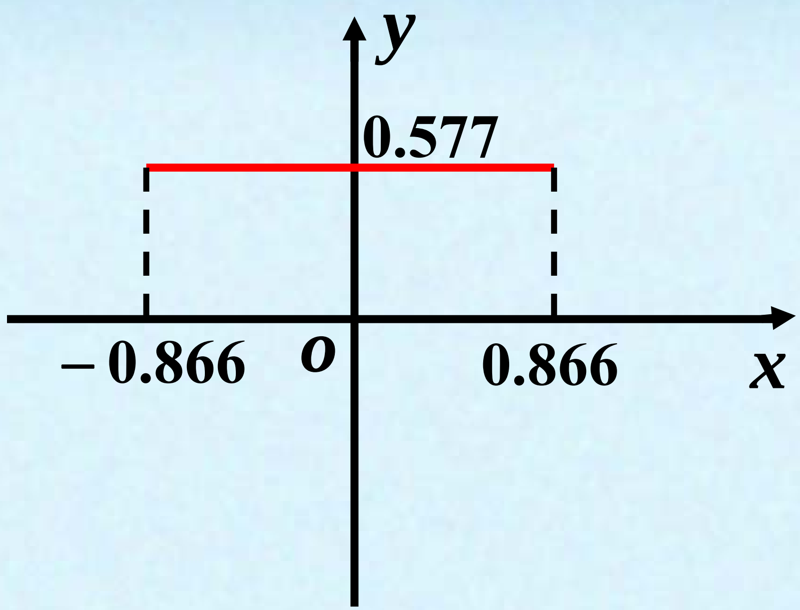


图3—7

例4 设数 X 在区间 $(0,1)$ 上随机地取值，当观察到 $X = x\ (0 < x < 1)$ 时，数 Y 在区间 $(x, 1)$ 上随机的取值，求 Y 的概率密度 $f_Y(y)$.

解 按题意 X 具有概率密度

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

对于任意给定的值 $x(0 < x < 1)$，在 $X = x$ 的条件下，Y 的条件概率密度为

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

由(3.4)式得X 和 Y 的联合概率密度为

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

$$=\begin{cases}\dfrac{1}{1-x}, & 0<x<y<1,\\ 0, & \text{其他}.\end{cases}$$

于是得关于Y的边缘概率密度为

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

$$=\begin{cases}\displaystyle\int_0^y \frac{1}{1-x}\,\mathrm{d}x=-\ln(1-y), & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

三、小 结

1. 设 (X,Y) 是二维离散型随机变量，$p_{ij}(i,j=1,2\cdots)$ 为其联合分布律，在给定 $Y=y_j$ 条件下随机变量 X 的条件分布律为

$$P\{X=x_i|Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}$$

$$=\frac{p_{ij}}{p_{\bullet j}},$$

在给定 $X = x_i$ 条件下随机变量 Y 的条件分布律为

$$P\{Y = y_j \mid X = x_i\} = \frac{P\{X = x_i, Y = y_j\}}{P\{X = x_i\}}$$

$$= \frac{p_{ij}}{p_{i\bullet}},$$

其中 $i, j = 1, 2, \cdots.$

2. 设 (X,Y) 是二维连续型随机变量,则有

$$F_{X|Y}(x|y)=\int_{-\infty}^{x} f_{X|Y}(x|y)\,\mathrm{d}x$$
$$=\int_{-\infty}^{x}[f(x,y)/f_Y(y)]\,\mathrm{d}x.$$

$$F_{Y|X}(y|x)=\int_{-\infty}^{y} f_{Y|X}(y|x)\,\mathrm{d}y$$
$$=\int_{-\infty}^{y}[f(x,y)/f_X(x)]\,\mathrm{d}y.$$

作业：11,13,15