目 录

第二章　随机变量及其分布

§2.6　条件分布和随机变量的独立性

§2.6.1　条件分布

一个随机变量(或向量)的条件概率分布，就是在给定(或已知)某种条件(某种信息)下该随机变量(向量)的概率分布。

1. 离散型随机变量的条件分布

设(X,Y)为二维离散型随机变量，其全部的可能取值为$\{(x_i,y_j):i,j=1,2,\cdots\}$。记其联合分布律为

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

若对给定的事件$\{Y=y_j\}$，其概率$P(Y=y_j)>0$，则称

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

为在给定$Y=y_j$的条件下X的条件分布律(概率函数)。类似的，若$P(X=x_i)>0$，则称

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

为在给定条件$X=x_i$下Y的条件分布律。

例 2.6.1. 设二维随机向量(X_1,X_2)的联合分布律如下所示:

X_1 \ X_2	-1	0	5	行和$p_{i\cdot}$
1	0.17	0.05	0.21	0.43
3	0.04	0.28	0.25	0.57
列和$p_{\cdot j}$	0.21	0.33	0.46	1.00

试求当$X_2=0$时，X_1的条件分布律。

解: 由联合分布律先算出两个边缘分布律$p_{i\cdot}$与$p_{\cdot j}$并填入表中，由此进一步算出条件分布律为：

$$P\{X_1=1|X_2=0\}=\frac{0.05}{0.33}=\frac{5}{33}$$

$$\text{而}\quad P\{X_1=3|X_2=0\}=\frac{0.28}{0.33}=\frac{28}{33}.$$

例 2.6.2. 设$X=(X_1,X_2,\cdots,X_n)\sim M(N;p_1,p_2,\ldots,p_n)$，试求$X_1$在给定$X_2=k$的条件下的条件分布律。

解：由于易知$(X_1,X_2)\sim M(N;p_1,p_2,1-p_1-p_2)$，即其联合分布律为

$$P(X_1=i,X_2=j)=\frac{N!}{i!j!(N-i-j)!}p_1^ip_2^j(1-p_1-p_2)^{N-i-j},\quad 0\le i,j\le N\ \&\ 0\le i+j\le N.$$

并且$X_2\sim B(N,p_2)$.

因此

$$\begin{aligned}
P(X_1=i|X_2=k) &= \frac{P(X_1=i,X_2=k)}{P(X_2=k)}\\
&= \frac{N!}{i!k!(N-i-k)!}p_1^ip_2^k(1-p_1-p_2)^{N-i-k}\Big/C_N^kp_2^k(1-p_2)^{N-k}\\
&= \frac{(N-k)!}{i!(N-k-i)!}\left(\frac{p_1}{1-p_2}\right)^i\left(1-\frac{p_1}{1-p_2}\right)^{N-k-i},\ i=0,1,\cdots,N-k.
\end{aligned}$$

即X_1在给定$X_2=k$的条件下服从二项分布$B(N-k,p_1/(1-p_2))$.

2. 连续型随机变量的条件分布

设(X,Y)有概率密度$f(x,y)$，我们考虑在给定$y\le Y\le y+\epsilon$的条件下X的条件分布函数(设$P\{y\le Y\le y+\epsilon\}>0$)

$$\begin{aligned}
P(X\le x|y\le Y\le y+\epsilon) &= \frac{P(X\le x,y\le Y\le y+\epsilon)}{P(y\le Y\le y+\epsilon)}\\
&= \int_{-\infty}^x\int_y^{y+\epsilon}f(u,v)dvdu\Big/\int_y^{y+\epsilon}f_Y(y)dy\\
&= \int_{-\infty}^x\frac{\int_y^{y+\epsilon}f(u,v)dv}{\int_y^{y+\epsilon}f_Y(y)dy}du
\end{aligned}$$

对上式两端关于x求导并令$\epsilon\to 0$, 可求得X在给定条件$Y=y$下的条件概率密度为

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

记为

$$X|y\sim f_{X|Y}(x|y).$$

类似地有Y在给定$X=x$的条件下的条件概率密度:

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

记为
$$Y|x \sim f_{Y|X}(y|x).$$

例 2.6.3. 设(X,Y)服从二元正态分布$N(a,b,\sigma_1^2,\sigma_2^2,\rho)$，试求$X|Y=y$的条件概率密度。

解:
$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f(x,y)}{f_Y(y)} \\ &= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}\exp\{-\frac{[x-(a+\rho\sigma_1\sigma_2^{-1}(y-b))]^2}{2\sigma_1^2(1-\rho^2)}\} \end{aligned}$$
即$X|Y=y\sim N(a+\rho\sigma_1\sigma_2^{-1}(y-b),\sigma_1^2(1-\rho^2))$。同理有：$Y|X=x\sim N(b+\rho\sigma_1^{-1}\sigma_2(x-a),\sigma_2^2(1-\rho^2))$。

例 2.6.4. 设X,Y服从单位圆上的均匀分布，试求$f_{X|Y}(x|y)$和$f_{Y|X}(y|x)$。

解: 由题设知(X,Y)的联合概率密度为
$$f(x,y)=\begin{cases} \frac{1}{\pi}, & x^2+y^2\le 1 \\ 0, & \text{其它} \end{cases}$$
易知
$$f_X(x)=\begin{cases} \frac{2}{\pi}\sqrt{1-x^2}, & -1\le x\le 1 \\ 0, & \text{其它} \end{cases}$$
所以
$$f_{X|Y}(x|y)=\begin{cases} \frac{1}{2\sqrt{1-y^2}}, & -\sqrt{1-y^2}\le x\le\sqrt{1-y^2} \\ 0, & \text{其它} \end{cases}$$
只需要把x,y互换，就可以得到$f_{Y|X}(y|x)$。

3. 更一般情形

无论离散型还是连续型条件分布，上述(X,Y)中的X和Y皆可推广到高维。例如: 设$(X_1,X_2,\cdots,X_n)\sim f(x_1,x_2,\ldots,x_n)$，且$(X_1,\cdots,X_k)\sim g(x_1,\ldots,x_k)$，则可定义在$(X_1,\cdots,X_k)=(x_1,\ldots,x_k)$的条件下，$(X_{k+1},\cdots,X_n)$的条件密度为：
$$h(x_{k+1},\ldots,x_n|x_1,\ldots,x_k)=\frac{f(x_1,\ldots,x_n)}{g(x_1,\ldots,x_k)},\quad \text{其中 } g(x_1,\ldots,x_k)>0.$$

注: 若记$(X_1,\cdots,X_k)=\boldsymbol{X}$，$(X_{k+1},\cdots,X_n)=\boldsymbol{Y}$，$(x_1,\ldots,x_k)=\boldsymbol{x}$，$(x_{k+1},\ldots,x_n)=\boldsymbol{y}$，则上式还可表示为:
$$h(\boldsymbol{y}|\boldsymbol{x})=\frac{f(\boldsymbol{x},\boldsymbol{y})}{g(\boldsymbol{x})},\ g(\boldsymbol{x})>0$$

§2.6.2 随机变量的独立性

若条件分布等于无条件分布，或者说条件分布与“条件”无关，例如，设$f_{X|Y}(x|y) = g(x)$，则可推出$g(x) = f_1(x)$，从而得到：

$$f(x, y) = f_1(x)f_2(y), \quad (x, y) \in \mathbb{R}^2$$

此时我们称X与Y是(相互)独立的。更一般的定义如下：

定义 2.6.1. 称离散型随机变量$X_1, \cdots, X_n$相互独立, 若它们的联合分布律等于各自的边缘分布律的乘积, 即

$$P(X_1 = x_1, \ldots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n),$$

其中$(x_1, \ldots, x_n)$ 为$(X_1, X_2, \cdots, X_n)$ 的值域中的任意一点.

定义 2.6.2. 称连续型随机变量$X_1, \cdots, X_n$相互独立, 若它们的联合密度等于各自的边缘密度的乘积, 即

$$f(x_1, \ldots, x_n) = f_1(x_1) \cdots f_n(x_n), \quad \forall \ (x_1, \ldots, x_n) \in \mathbb{R}^n$$

注: 更一般地, 有下面的的定义:

定义 2.6.3. 设$X_1, \cdots, X_n$为n个随机变量, 如果它们的联合分布函数等于各自边缘分布函数的乘积, 即

$$F(x_1, \cdots, x_n) = F_1(x_1) \cdots F_n(x_n), \quad \forall \ (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$$

则称随机变量$X_1, \cdots, X_n$相互独立.

在离散型和连续型两种情况下, 可以证明本定义分别与定义2.6.1和定义2.6.2等价.

例 2.6.5. 如果随机变量$X_1, \cdots, X_n$相互独立, 则容易证明其中任何一部分随机变量也相互独立. 然而一般来说, 仅由某一部分独立却无法推出$X_1, \cdots, X_n$相互独立. 如见下例:

例 2.6.6. 若ξ, η相互独立,都服从*-1*和*1*这两点上的等可能分布，而$\zeta = \xi\eta$。则ζ, ξ, η两两独立但不相互独立。

例 2.6.7. 设$(X,Y)\sim N(a,b,\sigma_1^2,\sigma_2^2,\rho)$，则$X$与$Y$相互独立的充要条件是$\rho=0$。

例 2.6.8. 设(X,Y)服从矩形$D=[a,b]\times[c,d]$上的均匀分布，则X与Y相互独立。

例 2.6.9. 设(X,Y)服从单位圆上的均匀分布，则X与Y不独立。

例 2.6.10. 设有n个事件：$A_1,\ A_2,\ \cdots,\ A_n$，对于每个事件A_i，定义：$X_i=I_{A_i}$ (A_i的示性函数), $i=1,\ 2,\ \cdots,\ n$，则可证明：$A_1\ ,A_2,\ \cdots,\ A_n$ 独立$\Longleftrightarrow X_1,\ X_2,\ \cdots,\ X_n$ 独立。