《概率论与数理统计》第八周要点

复习要点：§3 条件分布与 §4 相互独立的随机变量

一、条件分布

1. 离散型随机变量的条件分布律 设 (X,Y) 为二维离散型随机变量，其联合分布律为

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad p_{ij}\geq 0,\quad \sum_i\sum_j p_{ij}=1.$$

边缘分布律为

$$p_{i\cdot}=\sum_j p_{ij}=P\{X=x_i\},\qquad p_{\cdot j}=\sum_i p_{ij}=P\{Y=y_j\}.$$

若 $p_{\cdot j}>0$，则定义在 $Y=y_j$ 条件下 X 的条件分布律为：

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

同理，若 $p_{i\cdot}>0$，则

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

性质：

$$\sum_i P\{X=x_i|Y=y_j\}=1,\quad \sum_j P\{Y=y_j|X=x_i\}=1.$$

例 1： 在一汽车厂，机器人分别完成“焊接”和“紧固”两道工序。设 X 表示焊接不良点数，Y 表示螺栓紧固不良数，其联合分布如下表所示（略）。由表中边缘概率 $P(X=i),P(Y=j)$，可得条件分布：

$$P\{Y=k|X=1\}=\frac{P\{X=1,Y=k\}}{P\{X=1\}},\qquad P\{X=k|Y=0\}=\frac{P\{X=k,Y=0\}}{P\{Y=0\}}.$$

说明：条件分布反映了在一个变量取定条件下另一个变量的分布情况。

—

例 2：射击问题 射手独立射击，命中率为 p，直到击中两次为止。设 $X=$ 第一次命中的射击次数，$Y=$ 总射击次数。则

$$P\{X=m,Y=n\}=p^2(1-p)^{n-2},\quad n>m\geq 1.$$

边缘分布：

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

条件分布：

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

该例说明：条件分布可由联合分布与边缘分布直接求得。

—

2. 连续型随机变量的条件概率密度 设 (X,Y) 为二维连续型随机变量，联合密度为 $f(x,y)$，边缘密度为

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

若 $f_Y(y)>0$，则定义：

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

称为在 $Y=y$ 条件下 X 的条件概率密度函数。

性质：

$$f_{X|Y}(x|y)\geq 0,\quad \int_{-\infty}^{+\infty}f_{X|Y}(x|y)\,dx=1.$$

并有关系式：

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

条件分布函数：

$$F_{X|Y}(x|y)=P\{X\leq x|Y=y\}=\int_{-\infty}^{x}f_{X|Y}(t|y)\,dt.$$

—

例 3：二维均匀分布 设 (X,Y) 在单位圆域 $x^2+y^2<1$ 上均匀分布，密度为

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

边缘密度：

$$f_Y(y)=\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\frac{1}{\pi}\,dx=\frac{2\sqrt{1-y^2}}{\pi},\quad |y|<1.$$

条件密度：

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

说明：在圆域上均匀分布的 (X,Y)，在给定 $Y=y$ 时，X 在区间 $(-\sqrt{1-y^2},\sqrt{1-y^2})$ 上服从均匀分布。

—

例 4：随机区间问题 $X\sim U(0,1)$，给定 $X=x$，则 Y 在 $(x,1)$ 上均匀分布。条件密度：

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

联合密度：

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

边缘密度：

$$f_Y(y)=\int_0^y\frac{1}{1-x}\,dx=-\ln(1-y),\quad 0<y<1.$$

—

二、相互独立的随机变量

1. 定义 设 (X,Y) 的联合分布函数为 $F(x,y)$，若对所有 x,y 有

$$F(x,y)=F_X(x)\,F_Y(y),$$

则称 X 与 Y **相互独立**。

2. 等价刻画 - 对离散型随机变量：

$$P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{Y=y_j\},\quad \forall i,j.$$

- 对连续型随机变量：

$$f(x,y)=f_X(x)\,f_Y(y),\quad \forall x,y.$$

3. 基本性质

- 若 X,Y 独立，则 $f_{X|Y}(x|y)=f_X(x)$，即条件分布与 Y 无关；
- 独立 $\Rightarrow$ 不相关，但不相关不一定独立；
- 独立性是多维分布的重要结构性质。

—

三、小结

- 条件分布：刻画“给定一个变量取值后，另一个变量的分布”；
- 条件密度公式：$f_{X|Y}(x|y) = f(x,y)/f_Y(y)$；
- 常见例题：二维均匀分布、随机区间模型；
- 相互独立的判断标准：联合分布可分解为边缘分布的乘积。