# 第四章　随机变量的数字特征

## 条件频率函数的性质

① $P\{X=x_i \mid Y=y_j\} \geq 0 \quad (i=1,2,\cdots)$

② $\sum_{i=1}^{\infty} P\{X=x_i \mid Y=y_j\}=1$

这两条性质说明：
条件频率函数也是一种频率函数

## 条件密度函数的性质

① $f_{X|Y}(x \mid y) \geq 0$

② $\int_{-\infty}^{\infty} f_{X|Y}(u \mid y)du = 1$

从而可以定义条件期望(conditional expectation)和条件方差(conditional variance)

这两条性质说明：
条件密度函数也是一种密度函数

给定 $X=x$ 的情况下，$Y$的**条件期望**定义为

$$E(Y\mid X=x)=\sum_y y\underline{p_{Y|X}(y\mid x)}\qquad\text{（离散情形）}$$

$$E(Y\mid X=x)=\int y\underline{f_{Y|X}(y\mid x)dy}\qquad\text{（连续情形）}$$

更一般地, 函数 $h(Y)$ 的条件期望

$$E[h(Y)\mid X=x]=\sum_y h(y)p_{Y|X}(y\mid x)\qquad\text{（离散情形）}$$

$$E[h(Y)\mid X=x]=\int h(y)f_{Y|X}(y\mid x)dy\qquad\text{（连续情形）}$$

# §4 条件期望

例1

| $X \backslash Y$ | 0 | 1 | 2 | $p_{i\bullet}$ |
|---|---|---|---|---|
| 0 | 0.1 | 0.2 | 0.2 | 0.5 |
| 1 | 0.3 | 0.1 | 0.1 | 0.5 |
| $p_{\bullet j}$ | 0.4 | 0.3 | 0.3 | 1.0 |

$X$的条件分布律

| $X \backslash Y$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 1/4 | 2/3 | 2/3 |
| 1 | 3/4 | 1/3 | 1/3 |

$X$关于$Y$条件期望

| $Y$ | 0 | 1 | 2 |
|---|---|---|---|
| $E(X\|Y)$ | 3/4 | 1/3 | 1/3 |
| $p$ | 0.4 | 0.3 | 0.3 |

| $E(X\|Y)$ | 3/4 | 1/3 |
|---|---|---|
| $p$ | 0.4 | 0.6 |

| X \ Y | 0 | 1 | 2 | $p_{i\bullet}$ |
|---|---|---|---|---|
| 0 | 0.1 | 0.2 | 0.2 | 0.5 |
| 1 | 0.3 | 0.1 | 0.1 | 0.5 |
| $p_{\bullet j}$ | 0.4 | 0.3 | 0.3 | 1.0 |

$Y$的条件分布律

| X \ Y | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 1/5 | 2/5 | 2/5 |
| 1 | 3/5 | 1/5 | 1/5 |

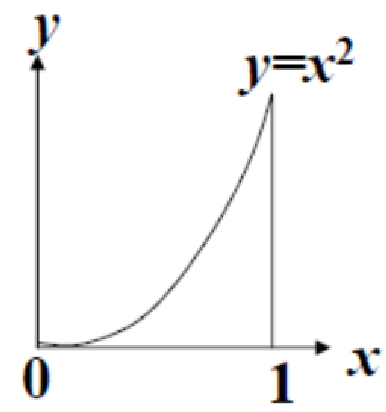
$Y$关于$X$条件期望

| $X$ | 0 | 1 |
|---|---|---|
| $E(Y\|X)$ | 6/5 | 3/5 |
| $p$ | 0.5 | 0.5 |

| $E(Y\|X)$ | 6/5 | 3/5 |
|---|---|---|
| $p$ | 0.5 | 0.5 |

例3 设$(X,Y)$服从$D=\{(x,y)|0<x<1,0<y<x^2\}$上的均匀分布

$$f(x,y)=\begin{cases}3, & 0<x<1,0<y<x^2\\ 0, & \text{其他}\end{cases}$$




$$f_X(x)=\begin{cases}3x^2, & 0<x<1\\ 0, & \text{其他}\end{cases}$$

当$0<x<1$时

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

$m_{Y|X}(x)=E(Y|X=x)=x^2/2$

得 $$E(Y|X)=\frac{1}{2}X^2$$

$$f_Y(y)=\begin{cases}3(1-\sqrt{y}), & 0<y<1\\ 0, & \text{其他}\end{cases}$$

当$0<y<1$时

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

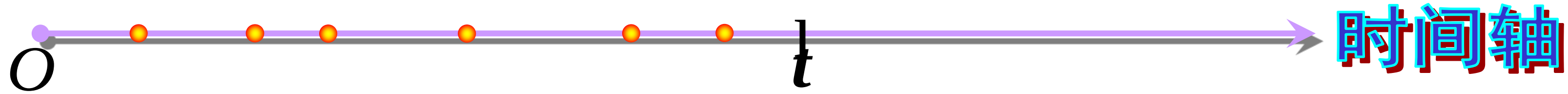
$$m_{X|Y}(y)=E(X\mid Y=y)=\frac{1}{2}(1+\sqrt{y})$$

得 $E(X\mid Y)=\dfrac{1}{2}(1+\sqrt{Y})$

**例（泊松流）**

考虑 $[0,1]$ 区间上均值为 $\lambda$ 的泊松流，令 $N$ 是$[0,1]$上点的个数. 对于 $p<1$, 令 $X$ 是 $[0, p]$ 上点的个数. 计算给定 $N=n$ 的情况下，$X$ 的条件分布和条件期望.

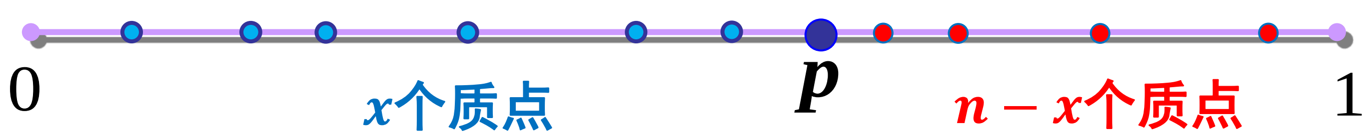
**回顾** 在泊松流中,记时间间隔 $(0,t]$ 中出现的质点数为 $X$




则 $X \sim P(\lambda t)$，即有 $P\{X=k\}=\dfrac{(\lambda t)^k}{k!}e^{-\lambda t}, k=0,1,2,\cdots$

其中参数 $\lambda>0$, 称为**泊松强度**.

**析** 本例中 $P\{X=x, N=n\}$ 是指：



**例（泊松流）**

考虑 $[0,1]$ 区间上均值为 $\lambda$ 的泊松流，令 $N$ 是$[0,1]$上点的个数. 对于 $p<1$, 令 $X$ 是 $[0,p]$ 上点的个数. 计算给定 $N=n$ 的情况下，$X$ 的条件分布和条件期望.

**解** 联合分布

$$P\{X=x,N=n\}=\frac{(p\lambda)^x e^{-p\lambda}}{x!}\frac{[(1-p)\lambda]^{n-x}e^{-(1-p)\lambda}}{(n-x)!},$$

而 $N\sim P(\lambda)$，即 $P\{N=n\}=\dfrac{\lambda^n e^{-\lambda}}{n!}$,

因此

$$P\{X=x|N=n\}=\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}\sim \boldsymbol{b(n,p)},$$

从而 $X$ 的条件期望为$np$.

**例（二元正态分布）** 设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$,

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\{-\frac{1}{2(1-\rho^2)}\times[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}]\}$$

则经过计算可得

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

即：二维正态分布，给定 $X$ 时 $Y$ 的条件密度是一维正态分布．因此

给定 $X=x$ 时 $Y$ 的条件期望为 $\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$,

条件方差为 $\sigma_2^2(1-\rho^2)$.

注

- 条件期望是 $x$ 的线性函数.
- 条件方差随着 $|\rho|$ 的增加而减小.

**理解条件期望 $E(Y\mid X)$：**

**假设对于$X$范围内的任意 $x$, 有$E(Y\mid X=x)$都存在.**

**是$X$的函数,**
**从而是r.v., 记为$E(Y\mid X)$**

**可以对条件期望 $E(Y\mid X)$再求期望和方差.**

$E[E(Y\mid X)]$ $D[E(Y\mid X)]$

# 条件期望的性质

性质:

(1) $E(c|X)=c$

(2) $E(aY_1+bY_2|X)=aE(Y_1|X)+bE(Y_2|X)$

(3) 若$X$与$Y$相互独立, 则$E(Y|X)=E(Y)$

(1) $E(c|X)=c$

证 $P\{Y=c\}=1$

$P\{Y=c|X=x\}=1$

$E(Y|X=x)=c$

(2) $E(aY_1+bY_2|X)=aE(Y_1|X)+bE(Y_2|X)$

证 设$X,Y_1,Y_2$的联合密度为$f(x,y_1,y_2)$

$$E(aY_1+bY_2|X=x)$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(ay_1+by_2)f_{(Y_1,Y_2)|X}(y_1,y_2|x)dy_1dy_2$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(ay_1+by_2)\frac{f(x,y_1,y_2)}{f_X(x)}dy_1dy_2$$

$$= a\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y_1 \frac{f(x,y_1,y_2)}{f_X(x)} dy_1 dy_2 + b\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y_2 \frac{f(x,y_1,y_2)}{f_X(x)} dy_1 dy_2$$

$$= a\int_{-\infty}^{+\infty} y_1 \frac{f_{(X,Y_1)}(x,y_1)}{f_X(x)} dy_1 + b\int_{-\infty}^{+\infty} y_2 \frac{f_{(X,Y_2)}(x,y_2)}{f_X(x)} dy_2$$

$$= a\int_{-\infty}^{+\infty} y_1 f_{Y_1|X}(y_1 \mid x) dy_1 + b\int_{-\infty}^{+\infty} y_2 f_{Y_2|X}(y_2 \mid x) dy_2$$

$$= aE(Y_1 \mid X = x) + bE(Y_2 \mid X = x)$$

(3) 若$X$与$Y$相互独立, 则$E(Y|X)=E(Y)$

证 $f_{Y|X}(y|x)=f(x,y)/f_X(x)=f_Y(y)$

$$E(Y \mid X = x) = \int_{-\infty}^{+\infty} y f_{Y|X}(y \mid x) dy = \int_{-\infty}^{+\infty} y f_Y(y) dy = E(Y)$$

**性质4** $$E(Y) = E[E(Y \mid X)].$$

$$D(Y) = D[E(Y \mid X)] + E[D(Y \mid X)].$$

**性质** $\boxed{E(Y)=E[E(Y\mid X)].}$

**全期望公式,**
**law of total expectation**

**证** 只证明离散情形(连续情形的证明与之类似).

$$E[E(Y\mid X)]=\sum_x E(Y\mid X=x)p_X(x)$$

$$=\sum_x[\sum_y yp_{Y|X}(y\mid x)]p_X(x)$$

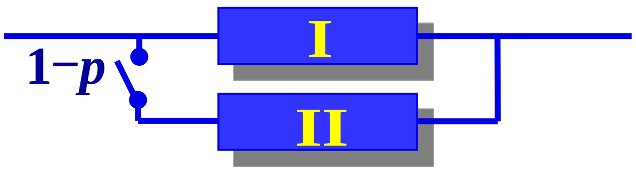
$$=\sum_y y\underline{\sum_x p_{Y|X}(y\mid x)p_X(x)}$$

$$=\sum_y yp_Y(y)$$

$$=E(Y).$$

**注** $Y$ 的期望可以通过先以 $X$ 为条件, 计算 $E(Y|X)$, 然后再对其关于 $X$ 取期望得到 (对条件期望**加权**).

**例** 假设在系统中，元件和备件的平均寿命都是 $\mu$.




如果元件失效，系统自动用其备件替代，但替换出错的概率为 $p$.

求整个系统的平均寿命.

**解** 令 $T$ 是系统的寿命.

$$X=\begin{cases}1, & \text{备件替代成功}\\ 0, & \text{备件替代不成功}\end{cases}$$

则 $\quad E(T\mid X=1)=2\mu,\qquad E(T\mid X=0)=\mu,$

因此

$$E(T)=E(T\mid X=1)P\{X=1\}+E(T\mid X=0)P\{X=0\}$$

$$=2\mu(1-p)+\mu p=\mu(2-p).$$

# P120：67、77、补充题

1. 如果$X$和$Y$是两个连续的随机变量，证明：$E[E(X|Y)]=E(X)$.

2. 设随机变量$(X,Y)$的概率密度为

$$f(x,y)=\begin{cases} ke^{-(x+y)}, & 0\leqslant y\leqslant x, \\ 0, & \text{其他}. \end{cases}$$

(1) 计算$Cov(X,Y)$, $\rho_{XY}$;

(2) 计算$E(X|Y=y)$ 和$E(Y|X=x)$;

(3) 推导随机变量$E(X|Y)$的概率密度.