§4 条件期望

条件频率函数的性质

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

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 $f_{X|Y}(x \mid y) \ge 0$

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

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

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

更一般地，函数 $h(Y)$ 的条件期望

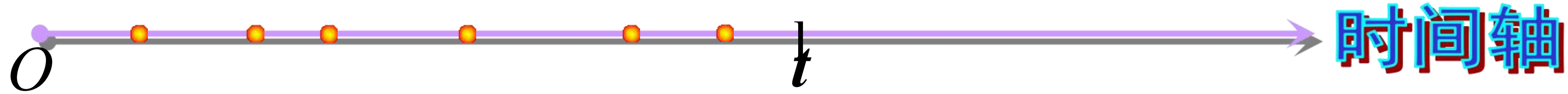
$$E[h(Y)\mid X=x]=\sum_y h(y)\,p_{Y|X}(y\mid x) \qquad \text{（离散情形）}$$

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

例（泊松流）

考虑 $[0,1]$ 区间上均值为 λ 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.

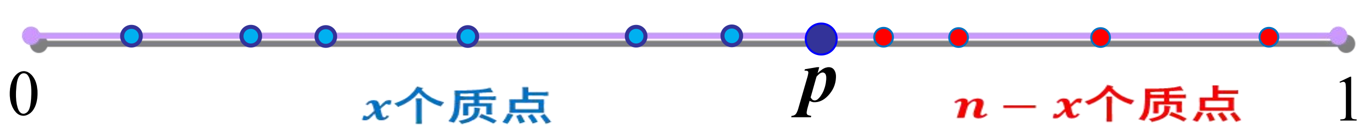
回顾 在泊松流中, 记时间间隔 $(0,t]$中出现的质点数为X



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 λ 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

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

从而 X 的条件期望为 $\boldsymbol{np}$.

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.

$E[E(Y \mid X)]$　$D[E(Y \mid X)]$

定理 $E(Y) = E[E(Y \mid X)].$

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

定理 $\boxed{E(Y)=E[E(Y\mid X)].}$ **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).

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

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

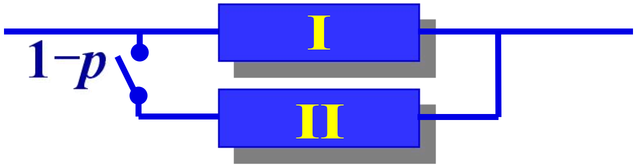
$$=\sum_y y\underline{\sum_x p_{Y|X}(y\mid x)p_X(x)}$$

$$=\sum_y y p_Y(y)$$

$$=E(Y).$$

注 Y 的期望可以通过先以 X 为条件, 计算 $E(Y\mid X)$, 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 μ.



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

因此

$$E(T) = E(T \mid X = 1)P\{X = 1\} + E(T \mid X = 0)P\{X = 0\}$$
$$= 2\mu(1 - p) + \mu p = \mu(2 - p).$$

$$\boxed{E(Y)=E[E(Y\mid X)].}\qquad \boxed{D(Y)=D[E(Y\mid X)]+E[D(Y\mid X)].}$$

例 (随机和) $T=\sum_{i=1}^{N}X_i.$ (关心 T 的期望和方差)

其中 N 是r.v., 具有有限期望和方差, $X_i(i=1,\dots,N)$ 具有相同的均值 $E(X)$和方差$D(X)$, 且 X_i 与 N 独立.

实际背景

保险公司在给定时段内收到N个索赔, 每个索赔额度用r.v. X_i 来刻画.

进入商场的顾客数为N, 第i个顾客的消费数额用r.v. X_i 来刻画.

单一服务队列的工作数目为N, 第i个工作的服务时间用r.v. X_i 来刻画.

解 由于 $E(T \mid N=n)=nE(X)$, $E(T \mid N)=NE(X)$,

从而 $E(T)=E[E(T \mid N)]=E[NE(X)]=E(N)E(X)$.

即: 完成 N 个工作的平均时间是随机数 N 的平均值乘以完成一个工作的平均时间.

由于 $D[E(T \mid N)]=D[NE(X)]=[E(X)]^2 D(N)$,

又 $D(T \mid N=n)=D(\sum_{i=1}^{n} X_i)=nD(X)$,

$D(T \mid N)=ND(X)$, $E[D(T \mid N)]=E(N)D(X)$,

从而 $D(T)=D[E(T \mid N)]+E[D(T \mid N)]$

$$=[E(X)]^2 D(N)+E(N)D(X).$$

即: 总时间 T 的不确定性来源于 N 的随机性和 X 的随机性. 【若固定 $N=n$, 则 $D(T)=nD(X)$.】

例 假设某时间区间内保险索赔数的期望等于900, 标准差等于30, 索赔数是期望为900的泊松随机变量.

假设平均索赔额度是1000美元, 标准差是500美元.

则: 所有索赔额度 T 具有

期望 $\mu = E(T) = 1000 \times 900 = 900000$ 美元,

方差 $D(T) = 1000^2 \times 900 + 900 \times 500^2 = 1.125 \times 10^9$,

标准差为 $\sigma = 33541$ 美元.

保险公司计划总索赔额度: $\mu \pm n\sigma$. (Chebyshev不等式).

注: 若总的索赔额是不变量, 即固定 $N=900$, 那么总索赔额度的方差为 900×500^2, 标准差为 15000美元.

因此, 索赔数的波动大大地提高了总索赔额度的不确定性.

课后作业

P120：67、77，补充题1，2

1. 如果X和Y是两独立的随机变量，证明：$E(X|Y=y)=E(X)$.

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$Cov(X,Y)$，ρ_{XY}；

(2) 计算$E(X|Y=y)$ 和$E(Y|X=x)$；

(3) 推导随机变量$E(X|Y)$和$E(Y|X)$的概率密度.