§1.2 条件期望

现代概率论的核心

一 离散型随机变量的条件期望

案例1 某购物网站有A、B品牌的服饰，某购物节那天预计有n个人访问该网站，每个人购买A、购买B、两个都不买的概率分别为$p, q, 1-p-q$，

$$\boldsymbol{X} = \text{“购买A的人数”}$$

$$\boldsymbol{Y} = \text{“购买B的人数”}$$

求 $\boldsymbol{E(X|Y)}$

思考：怎么求呢？

一般方法

设（X，Y）为离散型随机变量，联合分布列为

$$P(X=x_i, Y=y_j)=p_{i,j} \qquad P(Y=y_k)=p_{\cdot k} = \sum_{i=1}^{+\infty} p_{ik}$$

求 $E(X|Y)$

- 先求条件分布 $P(X=\boxed{x_i}|Y=y_j)= \dfrac{p_{ij}}{p_{\cdot j}}$

- 再求期望 $E(X|Y=y_j)= \sum_{i=1}^{+\infty} x_i \dfrac{p_{ij}}{p_{\cdot j}}$

$$= \frac{1}{p_{\cdot j}} \sum_{i=1}^{+\infty} x_i p_{ij}$$

案例1(续) 某购物网站有A、B品牌的服饰，某购物节那天预计有n个人访问该网站，每个人购买A、购买B、两个都不买的概率分别为$p, q, 1-p-q$，

$$X = \text{“购买A的人数”}$$
$$Y = \text{“购买B的人数”}$$

三项分布

求 $E(X|Y)$

解：$P(X=i, Y=j) = \underbrace{\frac{n!}{i!\,j!\,(n-i-j)!} p^i q^j (1-p-q)^{n-i-j}}$

$P(Y=j) = \frac{n!}{j!\,(n-j)!} q^j (1-q)^{n-j}$ $\quad Y \sim B(n, q)$

$P(X=i \mid Y=j) = \underbrace{\frac{(n-j)!}{i!\,(n-i-j)!} \left(\frac{p}{1-q}\right)^i \left(1-\frac{p}{1-q}\right)^{n-i-j}}$, $\quad \frac{(n-j)!}{i!\,(n-i-j)!} = C_{n-j}^{i}$

$i = 0, 1, \cdots, n-j$

$(X \mid Y=j) \sim B(\underbrace{n-j, \frac{p}{1-q}})$

$E(X \mid Y=j) = \underbrace{(n-j) \times \frac{p}{1-q}}$

$E(X|Y) = \underbrace{(n-Y) \times \frac{p}{1-q}} = \underline{g(Y)}$

❑ $E(X|Y)=g(Y)$ 为随机变量

$$E(X|Y=y_j)=g(y_j)=\frac{1}{p_{\cdot j}}\sum_i x_i p_{ij}$$

$\forall\, \mathbf{D}\in\mathscr{B}$

$$E[E(X|Y)I_D(Y)] = E[\,g(Y)\,I_D(Y)] = \sum_j g(y_j)\, I_D(y_j)\, p_{\cdot j}$$

$$g(y_j) = \frac{1}{p_{\cdot j}}\sum_i x_i p_{ij}$$

$$= \sum_j \sum_i x_i\, I_D(y_j)\, p_{ij}$$

$$= E(X I_D(Y))$$

可以得到什么结论？ $E[E(X|Y)I_D(Y)] = E(XI_D(Y))$

取 $D=(-\infty,\infty)$ ，有 $E(E(X|Y)) = E(X)$ 全数学期望公式

二 连续型随机变量的条件期望

案例 X——血液中的甘油三脂含量

Y——BMI指数 $(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$

求 $E(X|Y)$

设 $(X,Y)\sim f(x,y)$, $f_X(x|Y=y)=\dfrac{f(x,y)}{f_Y(y)}$ 怎么求 $E(X|Y)$?

$$E(X|Y=y)=\int_{-\infty}^{+\infty}x\,f_X(x|Y=y)\,dx$$

$$=\frac{1}{f_Y(y)}\int_{-\infty}^{+\infty}x\,f(x,y)\,dx$$

$$\triangleq g(y)$$

r.v.

$$\Rightarrow E(X|Y)=g(Y)$$

☆ $\forall D \in \mathscr{B}$

$E(X|Y) = g(Y)$, $\quad g(y) := \dfrac{1}{f_Y(y)} \int_{-\infty}^{\infty} x f(x,y)dy$

$$E\big(E(X|Y)I_D(Y)\big) = E[g(Y)I_D(Y)] = \int_{-\infty}^{+\infty} g(y) I_D(y) f_Y(y) dy$$

$$= \int_{-\infty}^{+\infty} \left(\frac{1}{\cancel{f_Y(y)}} \int_{-\infty}^{+\infty} x f(x,y) dx \right) I_D(y) \cancel{f_Y(y)} dy$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x I_D(y) f(x,y) dx dy$$

$$= E(X I_D(Y))$$

可以得到什么结论?

$$E[E(X|Y)I_D(Y)] = E(X I_D(Y))$$ ☆☆

取 $D = (-\infty, \infty)$ ，有

$$E(E(X|Y)) = E(X) \longleftarrow$$ 全数学期望公式

作业

若$(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$，证明

$$E(Y\mid X)=\mu_2+\rho\frac{\sigma_2}{\sigma_1}(X-\mu_1)$$

- 对于一般的随机变量，如何定义条件期望呢?
- 条件期望的数学本质是什么呢?
- 条件期望有什么性质呢? 为什么说条件期望是最好的预测呢?

定义 设(X,Y)为任意随机变量，随机变量$Z = E(X|Y)$称为X关于Y的**条件期望**，if

- Z为Y的函数； 可测函数
- $E[ZI_D(Y)] = E[XI_D(Y)], \ \forall D \in \mathscr{B}$

$X - E(X|Y)$ 不包含 Y 的信息 充分利用 任何信息

$E(X - E(X|Y)) = 0$

解释

解释：$E[(X-Z)I_D(Y)] = 0$

$E[(X - E(X|Y))I_D(Y)] = 0$

$Cov(X - E(X|Y), I_D(Y)) = 0$

$E(Z\,h(Y)) = E(X\,h(Y))$

$h(y)$为任意有界可测函数

命题 $E(X\,|Y)$ 是唯一的 ← 证法

四 条件期望的性质

1) $E[E(X|Y)]=$ $\underline{E(X)}$

案例 某购物网站有A、B品牌的服饰，某购物节那天预计有N个人访问该网站，N服从参数为λ的泊松分布。每个人购买A、购买B、两个都不买的概率分别为$p, q, 1-p-q$，

$X=$“购买A的人数” $Y=$“购买B的人数”

求 $E(XY)$ $Cov(X,Y)=?$

2) $$E\left(\sum_{i=1}^{n} c_i X_i \middle| Y\right) = \sum_{i=1}^{n} c_i E(X_i|Y)$$

3) $$E\left(g(X)h(Y) \middle| Y\right) = h(Y)E(g(X)|Y)$$

4) 若 (X,Y) 相互独立，则 $E\left(X \middle| Y\right) = E(X)$

证明. ① $E(X)$ 为 Y 的函数

② $$E\left((X-E(X))h(Y)\right) = E(X-EX)E(h(Y))$$

$$E(X-EX) = 0$$

$$= 0$$

5) 设X,Y是任意二阶矩存在的随机变量，$g(x)$是任意有界可测函数，且$g(Y)$的二阶矩存在，则

哪一个大?

$$E\left[(X-g(Y))^2\right] \geq E\left[(X-E(X|Y))^2\right]$$

证明 $E[(X-g(Y))^2] = E[(X-E(X|Y)+E(X|Y)-g(Y))^2]$

$= E[(X-E(X|Y))^2] + E[(E(X|Y)-g(Y))^2]$

$+2E[E((X-E(X|Y))|Y)(E(X|Y)-g(Y))] = 0$

$E((X-E(X|Y))|Y) = E(X|Y)-E(X|Y)=0$

$\geq E[(X-E(X|Y))^2]$

#

五 关于多维随机变量的条件期望 X, $Y_1, Y_2, \ldots, Y_n$

$\mathcal{G} = \sigma(Y_1, \cdots, Y_n)$

定义 随机变量 $Z = E(X|Y_1, \ldots, Y_n)$ 称为 X 关于 $Y_1, \ldots, Y_n$ 的**条件期望**，if

- ❑ Z 为 $Y_1, \ldots, Y_n$ 的函数；

信息充分利用

- ❑ $E[Zh(Y_1, \ldots, Y_n)] = E[Xh(Y_1, \ldots, Y_n)]$，$h$为任意有界可测函数

性质

1) $E\left(\sum_{i=1}^{n} c_i X_i \middle| Y_1, \ldots, Y_n\right) = \sum_{i=1}^{n} c_i E(X_i | Y_1, \cdots, Y_n)$

2) $E\left(g(X)h(Y_1, \ldots, Y_n) \middle| Y_1, \ldots, Y_n\right) = h(Y_1, \cdots, Y_n) E(g(X) | Y_1, \cdots, Y_n)$

3) 若X与$Y_1, \ldots, Y_n$相互独立，则 $E\left(X \middle| Y_1, \ldots, Y_n\right) = E(X)$

4)（塔式法则） 对 $1 \le m \le n$，则

$$E\left(X \mid Y_1,\ldots,Y_m\right)=E\left[E\left(X \mid Y_1,\ldots,Y_n\right) \mid Y_1,\ldots,Y_m\right]$$

证明：① $E[E(X|Y_1,\cdots,Y_n)|Y_1,\cdots,Y_m]$ 为 $Y_1,\cdots,Y_m$ 的函数

② $\forall\ h(y_1,\cdots,y_m)$ 有界可测函数

$$E\left[\left(X-E(X|Y_1,\cdots,Y_n)\right)h(Y_1,\cdots,Y_m)\right]=0$$

$$= E[Xh(Y_1,\cdots,Y_m)]-E[E(Xh(Y_1,\cdots,Y_m)|Y_1,\cdots,Y_n)]$$

$$E\left(g(X)g(Y_1,\ldots,Y_n) \mid Y_1,\ldots,Y_n\right)=E\left(g(X) \mid Y_1,\ldots,Y_n\right)g(Y_1,\ldots,Y_n)$$

$$\Rightarrow E(X-E(X|Y_1,\cdots,Y_n)|Y_1,\cdots,Y_m)=0 \quad E(Xh(Y_1,\cdots,Y_m))$$

\#

六 关于一般σ域的条件期望

信息

$(\Omega,\mathscr{F},P)$，ξ为随机变量，满足$E|\xi|<\infty$， $\mathscr{G}\subset\mathscr{F}$

✪ η关于$\mathscr{G}$可测—— $\forall x\in\mathbb{R},\ \{\eta\le x\}\in\mathscr{G}$

定义 随机变量 $E(\xi|\mathscr{G})$ 称为ξ关于$\mathscr{G}$的**条件期望**，if

- ❑ $E(\xi|\mathscr{G})$ 关于 $\mathscr{G}$ 可测；

- ❑ 对应任意 $\mathscr{G}$ 可测的有界随机变量η

$$E[\xi\eta]=E[E(\xi|\mathscr{G})\eta]$$

信息充分利用

性质　1） $E(c_1\xi_1+c_2\xi_2\mid\mathscr{G})=$ $c_1E(\xi_1\mid\mathscr{G})+c_2E(\xi_2\mid\mathscr{G})$

2） 若 η 为 $\mathscr{G}$ 可测的有界随机变量，则

$$E(\xi\eta\mid\mathscr{G})=\eta E(\xi\mid\mathscr{G})$$

3） 若 ξ 关于 $\mathscr{G}$ 独立，则

$$E(\xi\mid\mathscr{G})=E(\xi)$$

4） 若 $\mathscr{G},\mathscr{H}$ 为 $\mathscr{F}$ 的子σ域，且 $\mathscr{G}\subset\mathscr{H}$，则 ← 塔氏法则

$$E(\xi\mid\mathscr{G})=E(E(\xi\mid\mathscr{H})\mid\mathscr{G})$$

作业 1.18

补充题 1某购物网站有A、B品牌的服饰，某购物节那天预计有$\boldsymbol{N}$个人访问该网站，$\boldsymbol{N}$服从参数为$\boldsymbol{\lambda}$的泊松分布。每个人购买A、购买B、两个都不买的概率分别为$\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{1}-\boldsymbol{p}-\boldsymbol{q}$，

$$\boldsymbol{X}=\text{“购买A的人数”}$$
$$\boldsymbol{Y}=\text{“购买B的人数”}$$

1）求 $\boldsymbol{E}(\boldsymbol{X}), \boldsymbol{E}(\boldsymbol{Y})$

2）求 $\boldsymbol{E}(\boldsymbol{Y}\mid\boldsymbol{X})$ ← 挑战

补充题2

1）如果我们定义条件方差：$\boldsymbol{Var(Y|X)} = E\left[(Y - E(Y|X))^2 \middle| X\right]$，
证明：$\boldsymbol{Var(Y)} \geq \mathbf{Var}(\boldsymbol{Y|X})$.

2）设$(\boldsymbol{X},\boldsymbol{Y})$服从二维正态分布，证明：$\mathbf{Var}(\boldsymbol{Y|X}) = \boldsymbol{\rho}^2\mathbf{Var}(\boldsymbol{Y})$.

3） 继续1）的推广，证明：$\boldsymbol{Var(Y|X_1)} \geq \mathbf{Var}(\boldsymbol{Y|X_1,X_2})$.