Lecture 14

Given two, say, continuous, random vectors $X : \Omega \to \mathbb{R}^p$ and $Y : \Omega \to \mathbb{R}^q$, and an event $A \in \mathcal{F}$, the conditional expectation of X knowing A and the conditional independence of X and Y knowing A are defined the same way as without the condition, the densities $p_X(x|A)$, $p_Y(y|A)$, and $p_{X,Y}(x,y|A)$ given $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$ are merely replacing the densities $p_X(x)$, $p_Y(y)$, and $p_{X,Y}(x,y)$ in these definition. For simplicity, the definitions below are only provided for continuous random vectors, they can be naturally extended to more general distribution.

§1 Conditional expectation

1. **Definition:** Conditional expectation

令 ① event $A \in F$ 且 $P(A) > 0$

② random vector $X : \Omega \to R^p$

③ random vector $Y : \Omega \to R^q$, scalar $y \in R$

则 ① conditional expectation of X knowing A 被定义为:

$$E[X|A] = \int_{R^p} x \cdot p_X(x|A)\, dx$$

$$= \frac{1}{P(A)} \int_A X(w)\, d\, P(x)$$

$$= \frac{1}{P(A)} \int_{R^p} x \cdot p_X(x, A)\, dx$$

② conditional expectation of X knowing $Y = y$ 被定义为:

$$E[X|Y=y] = \int_{R^p} x \cdot p_{X,Y}(x|y)\, dx$$

③ 通常研究 w.s.F random vector:

$$E[X|Y] : \Omega \to R^p \quad (w \to E[X|Y=Y(w)])$$

2. **Lemma:** Law of total expectation

令 $X : \Omega \to R^p$, $Y : \Omega \to R^q$ 为 random vectors.

则 $E[E[Y|X]] = E[Y]$

证明:

$$E[E[Y|X]] = \int_{R^p} \left(\int_{R^q} y \cdot p_{Y,X}(y|x)\, dy \right) \cdot p_X(x)\, dx$$

$$= \int_{R^p} \int_{R^q} y \cdot p_{Y,X}(y|x) \cdot p_X(x)\, dy\, dx$$

$$= \int_{R^p} \int_{R^q} y \cdot p_{X,Y}(x,y)\, dy\, dx \quad \left(p_{Y,X}(y|x) = \frac{1}{p_X(x)} p_{X,Y}(x,y) \right)$$

$$= \int_{R^q} y \cdot \left(\int_{R^p} p_{X,Y}(x,y)\, dx \right) dy$$

$$= \int_{R^q} y \cdot p_Y(y)\, dy \quad (\text{marginal probability})$$

$$= E[Y]$$

3. **Lemma:** $E[Y \cdot f(X) | X]$

令 $X : \Omega \to R^p$, $Y : \Omega \to R^q$ 为 random vectors.

则 $E[Y \cdot f(X) | X] = f(X) \cdot E[Y|X]$ (注意等式两侧均为关于 w 的 functions)

证明:

对于 $\forall w \in \Omega$, 有

$$E[Y \cdot f(X) | X](w) = \int_{R^p} \int_{R^q} y \cdot f(x) \cdot p_{X,Y}((x,y) | X = X(w))\, dy\, dx$$

$$= \int_{\mathbb{R}^p}\int_{\mathbb{R}^q} y\cdot f(X(\omega))\cdot P_{X,Y}((x,y)\mid X=X(\omega))\,dy\,dx$$

(由于 $P_{X,Y}((x,y)\mid X=X(\omega))=0$ 若 $x\neq X(\omega)$)

$$= f(X(\omega))\cdot \int_{\mathbb{R}^q} y\cdot P_{X,Y}(y\mid X=X(\omega))\,dy$$

(由于 $P_{X,Y}((x,y)\mid X=X(\omega)) = \frac{P_{X,Y}(X(\omega),y)}{P_Y(y)} = P_{X,Y}(y\mid X=X(\omega))$)

$$= f(X)\cdot E[Y\mid X]$$

4. **Lemma:** 独立性与 conditional expectation

令 $X:\Omega\to R^p$, $Y:\Omega\to R^q$ 为 independent random vectors.

则 $E[Y\mid X] = E[Y]$

证明:

$$E[Y\mid X] = \int_{\mathbb{R}^q} y\cdot P_{X,Y}(y\mid x)\,dy$$
$$= \int_{\mathbb{R}^q} y\cdot P_Y(y)\,dy \quad (X,Y \text{ independent} \Rightarrow P_{X,Y}(y\mid x) = P_Y(y))$$
$$= E[Y]$$

e.g. **Example 5.28.** *Given four parameters $\mu,\nu,\sigma,\theta\in\mathbb{R}$, $\sigma,\theta>0$ and considering $X\sim\mathcal{N}(\mu,\sigma^2)$ and $Z\sim(\nu,\theta^2)$, the random variable $Y=X+Z$ satisfies:*

$$\mathbb{E}[Y]=\mu+\nu \qquad \mathbb{E}[Y\mid X] = X+\nu \quad (= E[X\mid X]+E[Z\mid X] = X+E[Z])$$
$$\mathbb{V}[Y]=\sigma^2+\theta^2 \qquad V(Y\mid X) = \mathbb{E}[(\underbrace{Y-X}_{=Z}-\nu)^2\mid X]=\theta^2,$$

one then denotes $Y\sim\mathcal{N}(X+\nu,\theta^2|X)$.

In the previous example it is interesting to see that the variance of Y, conditionally on X, $\mathbb{V}[Y|X]\equiv\mathbb{E}[(Y-\mathbb{E}[Y|X])^2\mid X]$, is independent of X. It is a simple consequence of the fact that $Y=X+Z$ and that X is independent with Z. We formalize below the notion of conditional independence.

§2 Conditional independence

1. **Definition:** conditional independence

① 令 events $A, B, C\in\mathcal{F}$

则称 A is independent of B conditionally on C 当且仅当

$$P(A\cap B\mid C) = P(A\mid C)\cdot P(B\mid C)$$

② 令 random vectors $X:\Omega\to R^p$, $Y:\Omega\to R^q$, $Z:\Omega\to R^r$

则称 X is independent of Y conditionally on Z 当且仅当

对 $\forall$ measurable mapping $f:R^p\to R$, $g:R^q\to R$, 有

$$E[f(X)g(Y)\mid Z] = E[f(X)\mid Z]\cdot E[g(Y)\mid Z]$$

注: Z 也可以替换为 event $A\in\mathcal{F}$

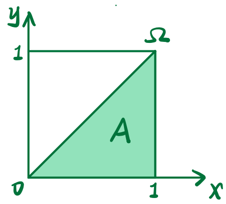
e.g. 令 $\Omega=[0,1]\times[0,1]$, $P=\lambda_2\cdot 1_\Omega$ 其中 λ_2 为 Lebesgue measure $\lambda_2(A)=\iint_{(x,y)\in A}dx\,dy$

考虑 $X:\Omega\to[0,1]$ 为 $(\omega_1,\omega_2)\to\omega_1$

$Y:\Omega\to[0,1]$ 为 $(\omega_1,\omega_2)\to\omega_2$

A 为 event $Y\le X$

则:



① $p_{X,Y}((x,y)|A) = \dfrac{p_{X,Y}((x,y),A)}{P(A)} = \dfrac{\overset{=1}{p_{X,Y}((x,y))}\cdot 1_{(x,y)\in A}}{\underset{=\frac{1}{2}}{P(A)}} = 2\cdot 1_{(x,y)\in A}$

② $p_X(x|A) = \int p_{X,Y}((x,y)|A)\,dy = 2\cdot\int 1_{(x,y)\in A}\,dy = 2\cdot\int_0^x dy = 2\cdot x$

③ $E[X|A] = \int_0^1 x\cdot p_X(x|A)\,dx = \int_0^1 x\cdot 2x\,dx = \frac{2}{3}$

④ $p_Y(y|A) = \int p_{X,Y}((x,y)|A)\,dx = 2\cdot\int 1_{(x,y)\in A}\,dx = 2\cdot\int_y^1 dy = 2\cdot(1-y)$

⑤ $E[Y|A] = \int_0^1 y\cdot p_Y(y|A)\,dy = \int_0^1 y\cdot 2(1-y)\,dy = \frac{1}{3}$

⑥ $E[XY|A] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\cdot p_{X,Y}\,dxdy = 2\cdot\int_0^1\int_0^x xy\,dydx = 2\cdot\int_0^1 \frac{1}{2}x^3\,dx = \frac{1}{4}$

$\Rightarrow$ $E[XY|A] \neq E[X|A]E[Y|A]$

$\Rightarrow$ X, Y not independent conditionally on A

2. **Lemma:** $f(X)$ 与 Y independent conditionally on X

令 ① random vectors $X:\Omega\to R^p$, $Y:\Omega\to R^q$, (可能 dependent)

② $\forall$ measurable mapping $f: R^p\to R$

则 $f(X)$ 与 Y independent conditionally on X

证明:

对 $\forall$ measurable mapping $g: R\to R$, $h: R^q\to R$, 有

$$E[g(f(X))\cdot h(Y)|X] = g(f(X))\cdot E[h(Y)|X] = E[g(f(X))|X]\cdot E[h(Y)|X]$$

利用 Lemma: $E[f(X)\cdot Y|X] = f(X)\cdot E[Y|X]$

$\Rightarrow$ $f(X)$ 与 Y independent conditionally on X

3. **Lemma:** 三个独立变量间的 conditional independence

令 ① random vectors $X:\Omega\to R^p$, $Y:\Omega\to R^q$, $Z:\Omega\to R^r$

② X, Y, Z mutually independent

则 Y 与 Z independent conditionally on X

证明:

对 $\forall$ measurable mapping $f: R^q\to R$, $g: R^r\to R$, 有

$$E[f(Y)g(Z)|X] = E[f(Y)g(Z)] = E[f(Y)]E[g(Z)] = E[f(Y)|X]\cdot E[g(Z)|X]$$

$\Rightarrow$ Y 与 Z independent conditionally on X

注: 该 Lemma 可拓展为: $\forall$ measurable mappings $\phi: R^p\to R^q$, $\psi: R^p\to R^r$, 有

$Y+\phi(X)$ 与 $Z+\psi(X)$ independent conditionally on X

4. Lemma: Conditionally independent 的随机变量的 variance

令 ① random vectors $X:\Omega\to R$, $Y:\Omega\to R$, $Z:\Omega\to R$

② Y 与 Z independent conditionally on X

则 $V[Y+Z|X] = V[Y|X] + V[Z|X]$

证明:

$$
\begin{aligned}
V[Y+Z|X] &= E[(Y+Z-E[Y+Z|X])^2|X] \\
&= E[(Y-E[Y|X]+Z-E[Z|X])^2|X] \\
&= E[(Y-E[Y|X])^2|X] + E[(Z-E[Z|X])^2|X] + \\
&\quad 2E[(Y-E[Y|X])(Z-E[Z|X])|X] \\
&= E[(Y-E[Y|X])^2|X] + E[(Z-E[Z|X])^2|X] + \\
&\quad 2\underbrace{E[(Y-E[Y|X])|X]}_{=0}\,\underbrace{E[(Z-E[Z|X])|X]}_{=0} \quad \text{(利用注中的结论)} \\
&= V[Y|X] + V[Z|X]
\end{aligned}
$$