Homework 10 - Solutions

Intermediate Statistics - 10705/36705

Problem 1. Recall that:

\[ E[|Y - g(X)|] = E \{ E[|Y - g(X)| \mid X] \} \]

The idea is to choose \( c \) such that \( E[|Y - c| \mid X = x] \) is minimized. Now define:

\[ r(c) = E[|Y - c| \mid X = x] = \int |y - c| p_{Y \mid X = x}(y) \, dy \]

The function \( h_y(c) = |y - c| \) is differentiable everywhere except when \( y = c \). Thus for \( c \neq y \)

\[ h_y'(c) = \begin{cases} 1 & c > y \\ -1 & c < y \end{cases} = I(c > y) - I(c < y) \]

Since \( Y \) is continuous and has a density function, \( \mathbb{P}(Y = c) = 0 \). So to minimize \( r(c) \) we can differentiate under the integral sign and set the derivative equal to 0 to obtain:

\[ r'(c) = \int h_y'(c) p_{Y \mid X = x}(y) \, dy = \int_{-\infty}^{c} p_{Y \mid X = x}(y) \, dy - \int_{c}^{\infty} p_{Y \mid X = x}(y) \, dy = 2 \int_{-\infty}^{c} p_{Y \mid X = x}(y) \, dy - 1 = 0 \]

\[ \Leftrightarrow \int_{-\infty}^{c} p_{Y \mid X = x}(y) \, dy = \frac{1}{2} \]

so that \( c = m(x) \), which is the median of \( p_{Y \mid X = x}(y) \). It is a minimum since \( r'(c) < 0 \) for \( c < m(x) \) and \( r'(c) > 0 \) for \( c > m(x) \). Since \( m \) minimizes \( E[|Y - g(X)| \mid X = x] \) at every \( x \), for any \( g \) we get

\[ E[|Y - g(X)| - |Y - m(X)||X = x]] \geq 0 \]

which implies

\[ R(g) - R(m) = E[|Y - g(X)| - |Y - m(X)|] = E \{ E[|Y - g(X)| - |Y - m(X)||X] \} \geq 0 \]
Problem 2. We will show that $\hat{R}$ converges in quadratic mean to $R$ conditional on $D_1$. First note that:

$$E[\hat{R}|D_1] = \frac{1}{n}E[(Y_{n+1} - \hat{g}(X_{n+1}))^2|D_1] = E[|Y - \hat{g}(X)|^2|D_1] = R$$

i.e. $E[\hat{R} - R|D_1] = 0$. Next assuming that $|\hat{g}| \leq C$, and since $R$ is a deterministic sequence we obtain:

$$\nabla[\hat{R} - R|D_1] = \nabla[\hat{R}|D_1] = \frac{1}{n}\nabla[(Y_{n+1} - \hat{g}(X_{n+1}))^2|D_1]$$

$$\leq \frac{1}{n}E[(Y_{n+1} - \hat{g}(X_{n+1}))^2|D_1]$$

$$\leq \frac{(B + C)^2}{n} \rightarrow 0$$

By the bias-variance decomposition we obtain $(\hat{R} - R)|D_1 \overset{q.m.}{\rightarrow} 0 \Rightarrow (\hat{R} - R)|D_1 \overset{P}{\rightarrow} 0$.

Problem 3. Saying that we “randomly assign” means $P( X_i = 0) = P( X_i = 1) = \frac{1}{2}$. Notice that

$$E[X_iY_i] = E[E[X_iY_i|X_i}] = E[X_iY_i|X_i = 0]P(X_i = 0) + E[X_iY_i|X_i = 1]P(X_i = 1)$$

$$= E[Y_i|X_i = 1]P(X_i = 1) = E[Y|X = 1]/2$$

Similarly we can get $E[(1 - X_i)Y_i] = E[Y|X = 0]/2$. Thus, by the Weak Law of Large Numbers we have:

$$\frac{2}{n}\sum_{i=1}^{n} X_iY_i \overset{P}{\rightarrow} 2E[XY] = E[Y|X = 1]$$

and

$$\frac{2}{n}\sum_{i=1}^{n} (1 - X_i)Y_i \overset{P}{\rightarrow} 2E[(1 - X)Y] = E[Y|X = 0]$$

By the CMT $\hat{\alpha}$ is a consistent estimator of $\alpha = E[Y|X = 1] - E[Y|X = 0]$. Also from Lecture Notes 17 section 2 we know that:

$$\alpha = E[Y|X = 1] - E[Y|X = 0]$$

$$= E[Y(1)|X = 1] - E[Y(0)|X = 0] \quad \text{(since } Y = XY(1) + (1 - X)Y(0))$$

$$= E[Y(1)] - E[Y(0)] = \theta$$

Where the last equality follows from the independence of $X$ and $Y(1), Y(0)$. So $\hat{\alpha}$ is a consistent estimator of $\theta$.  

**Problem 2 Old version.** Notice that

\[ \mathbb{E}[\hat{R}] = \mathbb{E}\{\mathbb{E}[^\hat{R}|D_1]\} = \mathbb{E}\{\mathbb{E}[|Y - \hat{g}(X)|^2D_1]\} = R \]

i.e. \( \mathbb{E}[\hat{R} - R] = 0. \)

Moreover, since \( R \) is a deterministic sequence of \( n \)

\[ \mathbb{V}[^\hat{R} - R] = \mathbb{V}(\hat{R}) = \mathbb{E}[\mathbb{V}(\hat{R}|D_1)] + \mathbb{V}(\mathbb{E}[\hat{R}|D_1]) \]

By assuming \( |\hat{g}| \leq C, \hat{g} \xrightarrow{P} g^∗ \), where \( g^∗ \) is a deterministic function, and \( \mathbb{E}[\hat{R}|D_1] = \mathbb{E}[|Y - h(X)|^2|h = \hat{g}] \xrightarrow{P} \mathbb{E}[|Y - g^∗(X)|^2] = c \)

- \( \mathbb{E}[\mathbb{V}(\hat{R}|D_1)] \leq \frac{(B+C)^2}{n} \xrightarrow{n \to \infty} 0 \), as shown in the other solution
- \( \mathbb{V}(\mathbb{E}[\hat{R}|D_1]) \xrightarrow{n \to \infty} 0 \)

Therefore, \( \hat{R} - R \xrightarrow{q.m.} 0 \Rightarrow \hat{R} - R \xrightarrow{P} 0. \)