ST790 Homework 2 Solution

1. (3.2)

Denote $x^T = (x_1, x_2, \cdots, x_n)$ and $y^T = (y_1, y_2, \cdots, y_n)$. The design matrix is $X = (1_n, x, x^2, x^3)$. The estimator of the linear function $a^T\beta$ is $a^T\hat{\beta} = a^T(X^TX)^{-1}X^Ty$. Its estimated standard error is

$$\widehat{se}(a^T\hat{\beta}) = (\hat{\sigma}^2 a^T(X^TX)^{-1}a)^{\frac{1}{2}}.$$

where $\hat{\sigma}^2 = (y^T(I_n - X(X^TX)^{-1}X^T)y)/(n-4)$. The 95% confidence interval is then $[a^T\hat{\beta} \pm z_{0.975}\widehat{se}(a^T\hat{\beta})]$.

For the second approach, the 95% confidence set for β is

$$C_\beta^{(2)} = \{\beta | (\hat{\beta} - \beta)^T X^T X(\hat{\beta} - \beta) \leq \hat{\sigma}^2 \chi^2_{4,0.95}\},$$

and the induced confidence interval for $a^T\beta$ is $[\min\{a^T\beta | \beta \in C_\beta^{(2)}\}, \max\{a^T\beta | \beta \in C_\beta^{(2)}\}]$.

The first approach treats a as known and estimates the variance of $a^T\hat{\beta}$ conditional on a while the second approach constructs a global confidence set of β without any knowledge of a. Thus the second approach is more conservative and will yield a wider interval.

2.

Write X, $\tilde{X}$, $h(X)$, Y and ϵ as the stack of X_i^T, $\tilde{X}_i^T$, $h(X_i)$, Y_i and ϵ_i. Let $A = diag\{A_1, A_2, \cdots, A_n\}$ and $\Pi = diag\{p(X_1), p(X_2), \cdots, p(X_n)\}$. The posited model is

$$Y = (\tilde{X} \quad (A - \Pi)\tilde{X}) \begin{pmatrix} \gamma \\ \beta \end{pmatrix} + \epsilon_{n\times 1} =: X_* \begin{pmatrix} \gamma \\ \beta \end{pmatrix} + \epsilon.$$

The least square estimator for $(\gamma^T, \beta^T)^T$ is

$$\begin{pmatrix} \hat{\gamma} \\ \hat{\beta} \end{pmatrix} = (X_*^T X_*)^{-1} X_*^T Y = \left(\frac{1}{n} X_*^T X_*\right)^{-1} \begin{pmatrix} \frac{1}{n}\tilde{X}^T \\ \frac{1}{n}\tilde{X}^T(A - \Pi) \end{pmatrix} Y,$$

where

$$\frac{1}{n} X_*^T X_* = \frac{1}{n} \begin{pmatrix} \tilde{X}^T\tilde{X} & \tilde{X}^T(A-\Pi)\tilde{X} \\ \tilde{X}^T(A-\Pi)\tilde{X} & \tilde{X}^T(A-\Pi)(A-\Pi)\tilde{X} \end{pmatrix} =: \begin{pmatrix} \hat{\Sigma}_{n,11} & \hat{\Sigma}_{n,12} \\ \hat{\Sigma}_{n,12} & \hat{\Sigma}_{n,22} \end{pmatrix}.$$

Since $\mathbb{E}(A - \Pi | \tilde{X}) = 0$, the off-diagonal blocks $\hat{\Sigma}_{n,12} \to_p 0$ when $n \to \infty$ and hence

$$\left(\frac{1}{n} X_*^T X_*\right)^{-1} \to_p \begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix},$$

where Σ_{11} and Σ_{22} are the expectations of $\hat{\Sigma}_{n,11}$ and $\hat{\Sigma}_{n,22}$. Similarly, we have $\mathbb{E}[\tilde{X}^T(A - \Pi)h(X)|\tilde{X}] = 0$, which indicates the orthogonality between $h(X)$ and $(A - \Pi)\tilde{X}\beta$, and hence it can be easily verified that

$$\begin{aligned}
\mathbb{E}[\tilde{X}^T(A - \Pi)Y] =& \mathbb{E}[\tilde{X}^T(A - \Pi)(h(X) + A\tilde{X}\beta_0 + \epsilon)] \\
=& \mathbb{E}[\tilde{X}^T(A - \Pi)(h(X) + (A - \Pi)\tilde{X}\beta_0 + \Pi\tilde{X}\beta_0 + \epsilon)] \\
=& 0 + n\Sigma_{22}\beta_0 + 0 + 0.
\end{aligned}$$

Therefore $\tilde{X}^T(A - \Pi)Y/n \to_p \Sigma_{22}\beta_0$ as $n \to \infty$. Put it all together, we have $\hat{\beta} \to_p \Sigma_{22}^{-1}\Sigma_{22}\beta_0 = \beta_0$.