STAT151A Quiz 4 solutions

Please write your full name and email address:

This question will take the residuals of the training data to be random, and will consider variablity under sampling of the training data. The regressors for both the training data and test data will be taken as fixed.

The inverse of a 2x2 matrix is given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

The OLS estimator is given by $\hat{\beta} = (\boldsymbol{X}^\intercal \boldsymbol{X})^{-1} \boldsymbol{X}^\intercal \boldsymbol{Y}$.

You have 20 minutes for this quiz.

There are three parts, (a), (b), and (c), each weighted equally..

(a)

For this quiz, we will assume that $y_n = \beta^\intercal \boldsymbol{x}_n + \varepsilon_n$ for some β, and that the residuals ε_n are IID with $\mathbb{E}[\varepsilon_n] = 0$ and $\mathbb{E}[\varepsilon_n^2] = 1$, but not necessarily normal.

Let $\boldsymbol{x}_n = (1, z_n)^\intercal$, where $\frac{1}{N}\sum_{n=1}^N z_n = 0$ and $\frac{1}{N}\sum_{n=1}^N z_n^2 = \delta > 0$. That is, assume we are regressing on a **constant** and a **single mean-zero regressor**. For this question, take the regressors to be fixed (not random).

Find the limiting distribution of $\sqrt{N}(\hat{\beta} - \beta)$ as $N \to \infty$.

Students can use the fact that

$$\sqrt{N}(\hat{\beta} - \beta) \to \mathcal{N}\left(\mathbf{0}, \sigma^2 \left(\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \boldsymbol{x}_n \boldsymbol{x}_n^\intercal\right)^{-1}\right)$$

without proving it. We have that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \boldsymbol{x}_n \boldsymbol{x}_n^\intercal = \begin{pmatrix} 1 & 0 \\ 0 & \delta \end{pmatrix}$$

so

$$\sqrt{N}(\hat{\beta} - \beta) \to \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} 1 & 0 \\ 0 & \delta^{-1} \end{pmatrix}\right)$$

(b)

Define the expected prediction error

$$y_{\text{new}} := \beta^\intercal \boldsymbol{x}_{\text{new}} + \varepsilon_{\text{new}} \quad \text{and} \quad \hat{y}_{\text{new}} := \hat{\beta}^\intercal \boldsymbol{x}_{\text{new}}.$$

Under the conditions given in part (a), find the limiting distribution of

$$\sqrt{N}(\hat{y}_{\text{new}} - \mathbb{E}\left[y_{\text{new}}\right])$$

as $N \to \infty$, as a function of $\boldsymbol{x}_{\text{new}}$. That is, the limiting distribution will depend on $\boldsymbol{x}_{\text{new}}$, so please make the dependence explicit.

You may use your answer from part (a).

Since $\hat{y}_{\text{new}} = \hat{\boldsymbol{\beta}}^\intercal \boldsymbol{x}_{\text{new}}$ and $\mathbb{E}\left[y_{\text{new}}\right] = \beta^\intercal \boldsymbol{x}_{\text{new}}$,

$$\sqrt{N}(\hat{y}_{\text{new}} - \mathbb{E}\left[y_{\text{new}}\right]) = \sqrt{N}\boldsymbol{x}_{\text{new}}^\intercal(\hat{\beta} - \beta) \to \mathcal{N}\left(0, \boldsymbol{x}_{\text{new}}^\intercal \begin{pmatrix} 1 & 0 \\ 0 & \delta^{-1} \end{pmatrix} \boldsymbol{x}_{\text{new}}\right) = \mathcal{N}\left(0, 1 + \delta^{-1} z_{\text{new}}^2\right)$$

(c)

Assume the conditions and definitions given in (a) and (b). Assume that $\delta \ll 1$ (that is, δ is much smaller than 1.)

Find the limiting distribution of $\sqrt{N}(\hat{y}_{\text{new}} - \mathbb{E}\left[y_{\text{new}}\right])$ when

- $\boldsymbol{x}_{\text{new}} = (1, 0)$ and
- $\boldsymbol{x}_{\text{new}} = (1, 1)$.

Which of the two is larger?

You may use your answer from parts (a) and (b).

Plugging in, we see that the distributions are

- $\mathcal{N}(0, 1)$
- $\mathcal{N}\left(0, 1 + \delta^{-1}\right)$.

The second has larger variance.