

STAT 151A Lecture 16

Henry Liev

3 October 2025

Example 0.1

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

Write as $L\beta = c$

$$\begin{pmatrix} \vec{0} & \mathbb{I}_p \end{pmatrix} \beta = \begin{pmatrix} \vec{0} \end{pmatrix}$$

F -statistic numerator: $(L\hat{\beta} - c)^\top [L(\mathbf{X}^\top \mathbf{X})^{-1} L^\top]^{-1} (L\hat{\beta} - c)$

$$(\mathbf{X}^\top \mathbf{X})^{-1} = V \in \mathbb{R}^{(p+1) \times (p+1)}$$

$$LVL^\top = \begin{pmatrix} \vec{0} & \mathbb{I}_p \end{pmatrix} V \begin{pmatrix} \vec{0}^\top \\ \mathbb{I}_p \end{pmatrix} = V \text{ without first row and column}$$

Method 0.2 (Incremental F -test)

$$L\beta = c$$

$$\begin{pmatrix} \vec{0} & \mathbb{I}_k & \vec{0}_{(k \times (p+1-k))} \end{pmatrix} \beta = \begin{pmatrix} \vec{0} \end{pmatrix}$$

Subset of V where we remove the first row and column and the last $(p+1) - k$ rows and columns

Method 0.3 (Confidence Interval)

For $\hat{\beta}_j$

Under NLM $Y = \mathbf{X}\beta + \varepsilon, \varepsilon \sim \mathcal{N}[0, \sigma^2 \mathbb{I}_n]$

$$\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (\mathbf{X}^\top \mathbf{X})^{-1})$$

Obvious move: $\hat{\beta}_j \pm z_{\alpha/2} \cdot \sigma \sqrt{(\mathbf{X}^\top \mathbf{X})_{jj}^{-1}}$

Over all random samples $Y | \mathbf{X}$, this random interval should contain the true β_j 95% of the time if the normal linear model holds.

Actually we don't know σ , need to use $\hat{\sigma}$, as a result we need to use the t -distribution

Actual CI reported by lm:

$$\hat{\beta}_j \pm (t_{n-p-1, 1-\alpha/2}) \cdot \hat{\sigma} \sqrt{(\mathbf{X}^\top \mathbf{X})_{jj}^{-1}}$$

Method 0.4 (Prediction Interval)

We observe $(y_1 \in \mathbb{R}, \mathbf{X} \in \mathbb{R}^{p+1}), \dots, (y_n, \vec{x}_n)$ and fit a regression

Then we observe a new \vec{x}_{n+1}

How do we predict the value of y_{n+1}

$$\hat{y}_{n+1} = (\vec{x}_{n+1})^\top \hat{\beta}$$

What is the uncertainty in your prediction?

2 ways to answer this:

- (1) Confidence interval for $\vec{x}_{n+1} \hat{\beta}$ conditional on X
- (2) Prediction interval for y_{n+1}

Start with (1): Under NLM

$$\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (\mathbf{X}^\top \mathbf{X})^{-1})$$

$$x_{n+1} \sim \mathcal{N}(x_{n+1} \beta, \sigma^2 x_{n+1} (\mathbf{X}^\top \mathbf{X})^{-1} x_{n+1}^\top)$$

$$x_{n+1} \in \mathbb{R}^{p+1}$$

$$x_{n+1}^\top \hat{\beta} \pm (T_{n-p-1})_{1-\alpha/2} \cdot \hat{\sigma} \sqrt{x_{n+1} (\mathbf{X}^\top \mathbf{X})^{-1} x_{n+1}^\top}$$

(2): $y_{n+1} = x_{n+1} \beta + \varepsilon_{n+1}$ under NLM

$$\text{Prediction error: } y_{n+1} - \hat{y}_{n+1} = x_{n+1} (\beta - \hat{\beta}) + \varepsilon_{n+1}$$

Confidence interval gets the estimation error but not the fundamental noise

Prediction interval:

$$x_{n+1}^\top \hat{\beta} \pm (t_{n-p-1})_{1-\alpha/2} \cdot \sqrt{\hat{\sigma}^2 (x_{n+1}^\top (\mathbf{X}^\top \mathbf{X})^{-1} x_{n+1} + \hat{\sigma}^2)} = (t_{n-p-1})_{1-\alpha/2} \cdot \hat{\sigma} \sqrt{1 + (x_{n+1}^\top (\mathbf{X}^\top \mathbf{X})^{-1} x_{n+1})}$$