Regression Analysis Tutoring4

Seung Bong Jung

Seoul National University

November 1, 2021

Model comparison test

- Suppose we are considering following model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma^2),$$

  where $x_2 = 0$ or 1.
- If $x_{2i} = 0$, we have $y_i = \beta_0 + \beta_1 x_{1i} + \epsilon_i$. Otherwise, $y_i = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) x_{1i} + \epsilon_i$
- Thus, if we are to test $H_0 : \beta_2 = \beta_3 = 0$ versus H_1 : not H_0, we're testing whether two give regression models are equal or not.
- Similarly, if we're testing $H_0 : \beta_3 = 0$ versus H_1 : not H_0, we're testing whether two regression models are parallel. Also, for the test $H_0 : \beta_2 = 0$ versus H_1 : not H_0, we're comparing the intercepts of two models.

Model comparison test

- In general, fundamental principle for hypothesis testing in regression analysis is that we reject H_0 if the following F-statistic is too large:

$$F \equiv \frac{(\mathsf{SSE}^0 - \mathsf{SSE}^1)/(\mathsf{df}^0 - \mathsf{df}^1)}{\mathsf{SSE}^1/\mathsf{df}^1} \sim F(\mathsf{df}^0 - \mathsf{df}^1, \mathsf{df}^1)$$

  where SSE^0 denotes the SSE under reduced model and SSE^1 the SSE under under full model. Note that $\mathsf{SSE}^0/\sigma^2 \sim \chi^2(\mathsf{df}^0)$ and $\mathsf{SSE}^1/\sigma^2 \sim \chi^2(\mathsf{df}^1)$. This principle is valid provided that the definition of F-distribution is satisfied.

Model comparison test

- Write $\mathbf{X} = (\mathbf{X}_0, \mathbf{X}_1)$. For the matrix $\mathbf{A}$, denote the column space of $\mathbf{A}$ by $\mathcal{C}(\mathbf{A})$. Then clearly, $\mathcal{C}(\mathbf{X}_0) \subseteq \mathcal{C}(\mathbf{X})$. Hence, $\mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{X}_0 = \mathbf{X}_0$.
- Following the notations in the previous slide and using this fact, one can prove followings:
- For $H_0 : \beta_2 = \beta_3 = 0$, under H_0, $F \sim F(2, n-4)$.
- For $H_0 : \beta_3 = 0$, under H_0, $F \sim F(1, n-4)$.
- For $H_0 : \beta_2 = 0$, under H_0, $F \sim F(1, n-4)$