625.661 Statistical Models and Regression

Module 3 Discussion Questions

H.M. James Hung

Please complete all problems.

A set of n subjects give data on a response variable y and two regressors, x_1 and x_2. The data are fitted to two different models

Model 1: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$
Model 2: $y = \alpha_0 + \alpha_1(x_1 - \bar{x}_1) + \alpha_2(x_2 - \bar{x}_2) + \eta$

where $\bar{x}_k$ is the simple average of x_k values from this set of data, k = 1, 2; ε and η are random errors with mean zero.

Discuss with mathematical arguments whether the ordinary least-squares estimator of β_i under Model 1 is equal to the ordinary least-squares estimator of α_i under Model 2, for $i = 0, 1, 2$. State the assumptions in your discussion. Discuss whether or not the regressors are random or non-random change your discussion.

From Model 2,

$$\begin{aligned} \boldsymbol{E(y \mid x_1, x_2) } &= \boldsymbol{\alpha_0 + \alpha_1(x_1 - \bar{x}_1) + \alpha_2(x_2 - \bar{x}_2)} \\ &= \boldsymbol{\alpha_0 - \alpha_1\bar{x}_1 - \alpha_2\bar{x}_2 + \alpha_1 x_1 + \alpha_2 x_2} \end{aligned}$$

The difference is between $\boldsymbol{\beta_0}$ and $\boldsymbol{\alpha_0 - \alpha_1\bar{x}_1 - \alpha_2\bar{x}_2}$. We then can see from minimizing the corresponding sum of squares to get to three normal equations for each model in order to solve for the estimators of the three regression coefficients. After the derivations from solving the three normal equations for each model, we can prove that the estimators of $\boldsymbol{\alpha_1}$ and $\boldsymbol{\beta_1}$ are identical and so are the estimators of $\boldsymbol{\alpha_2}$ and $\boldsymbol{\beta_2}$. But $\boldsymbol{\alpha_0}$ and $\boldsymbol{\beta_0}$ are different. These results remain when $\boldsymbol{x}$'s are random.