Jared Yu

Module 3 Discussion Reply

I had a similar thought process initially in trying to understand this discussion problem. Based off the supplemental lecture, I also wrote out the same formula for y, except centering the regressors with their sample means. I was interested in making the same conclusion regarding how Model 1 and 2 were the same, given that $\alpha_0 = \beta_0 + \beta_1\bar{x}_1 + \beta_2\bar{x}_2$. The issue that I was unsure about however is regarding the error terms, ε and η. I understand that they both have mean zero with constant variance, but I was unsure whether it is possible to just interchange them as such. For example, what if the variance was constant but of different magnitude between the two models? This question made me try to find an alternative approach towards understanding the problem. In my result, I did see also that in general, $\hat{\beta}_1 = \hat{\alpha}_1$ and $\hat{\beta}_2 = \hat{\alpha}_2$. However, $\hat{\beta}_0 \neq \hat{\alpha}_0$, with the exception of some special conditions, such as if $\bar{x}_1 = \bar{x}_2 = 0$.

Regarding the issue of the regressors being random or fixed, I notice that you referenced a section from the textbook. I recall reading that line too, but I did not use it directly. Instead, I tried to argue that in the least-squares normal equations, to solve for $(\mathbf{X}'\boldsymbol{X})^{-1}$ requires that the resulting inverse matrix exists. In the steps, it seems that it is necessary that a certain denominator is nonzero, otherwise the inverse is not calculatable. The book states that the requirement is for linear independence amongst the regressors. If the fixed data is simulated, this is controllable. However, if it is random data, then it is not possible to determine beforehand.