UPPSALA UNIVERSITET

Assignment 1: Interactions I

In this assignment, we want you to look at interaction terms. Section 1 in this text specifies the initial DGP. In section 2, we want you to formulate the interpretation of a model with interaction term, and to look at the correlation that emerges when two variables that have the same sign are interacted. In section 3, we want you show that omitting the interaction term from the initial DGP leads to omitted variable bias. In section 4, you should show that centering reduces collinearity problems that interaction models have given the initial DGP. In section 5, you should demonstrate that these collinearity problem increase when the two interaction variables are highly correlated.
Remember to install and load the `car,mvtnorm` packages

1 Assume that the true data-generating process has the form

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 \times X_2 + u_i$$

where

- u_i has a normal distribution N(0,σ)
- X_1, X_2 are drawn from normal distributions $N(3, 1)$ (with mean = 3 and standard deviation = 1)
- $X_1 \times X_2$ is a multiplicative interaction term

UPPSALA UNIVERSITET

Assignment 1: Interactions II

- Set the true values for the β parameters to be an arbitrary combination of the values $(-1, -0.5, 0.5, 1)$, and the size of the sample to $N = 500$

2 Draw a realization from this DGP, fit two linear models – one including all the four terms and one omitting the interaction term

- Present the results in a joint regression table and interpret the results.
- Plot the predicted value for Y as a function of $X1$ for $X2 = 2$ for $X2 = 4$ and interpret the figures
- Present the joint distribution of X_1 and the multiplicative interaction term $X_1 \times X_2$ as three-dimensional density plots (one perspective plot and one contour plot). Describe and discuss.

3 Run a MC analysis to explore the consequences of omitting the interaction term from the model specification when fitting a regression model. Present results in terms of bias for the estimated parameter $\hat{\beta}_1$

UPPSALA UNIVERSITET

Assignment 1: Interactions III

4. Extend the MC analysis further by varying the correlation between X_1 and X_2 along the lines of the collinearity MC experiment from Chapter 5 in Carsey and Harden (2014). Again, focus on the fully specified model and concentrate on the collinearity issues in interaction models. For the centered and non-centered solution, report the following plots, discuss:
   1. The standard deviation of $\hat{\beta}_1$ as a function of `mc.level` – are estimates efficient?
   2. MSE for the β_1 estimates as a function of `mc.level` – are estimates biased?
   3. The mean variance inflation factor (VIF) for β_1 across simulations for each `mc.level`
   4. The estimated $\hat{\beta}_1$ against $\hat{\beta}_3$ for each of the repeated samples for `mc.level` = 0, 0.5 or .99. Choose either the centered or non-centered solution.
   5. Histograms of $\hat{\beta}_1$, $\hat{\beta}_3$, and the sum $\hat{\beta}_1 + \hat{\beta}_3$ for `mc.level` = 0, 0.5 or .99. Choose either the centered or non-centered solution. What happens here?