Daniel Rudy
Intermediate Quantitative Methods

Problem Set 2

Problem 1 Writeup (other parts addressed in R file):

Is OLS the appropriate estimator? Why or why not? Explain in terms of the Gauss-Markov assumptions.

Yes, OLS is the appropriate estimator in this case, because all of the Gauss-Markov assumptions are met, so by the Gauss-Markov Theorem, OLS is BLUE. The data for the regressors were generated randomly and independently according to either a normal or Bernoulli distribution, which produce normally distributed error terms. The final model combines these variables linearly via addition. Full explanations by assumption follow:

Linearity: Yes, The model here is clearly linear as all independent variables have a linear relationship with the independent variable, and their addition maintains linearity.

Full Rank: Yes, As all of the regressors are generated independently, they are linearly independent.

Exogeneity: Yes, The error terms in a normal distribution (X_1,X_2) are normally distributed. Bernoulli distributions also have error terms with a normal distribution (X_3). Because of this, they will have mean 0, and will not be correlated with the regressors.

Homoskedasticity and Nonautocorrelation: Yes, Homoskedasticity is demonstrated by each error term being normally distributed. They are nonautocorrelated because they are generated independently based on (but not correlated to) their corresponding regressor, and therefore do not depend on each other.

Fixed or Random Data Generation: Yes, The data in this case were randomly generated.

Normality: Yes, As mentioned above, the error terms are normally distributed by their construction. Because their corresponding regressors are randomly generated to be either normally or Bernoulli distributed, the error terms will have a normal distribution.