chapter 4: The Classical Model

Lecture 2

1 Variance of the OLS Estimators

Under the Classical Assumptions (except for VII), the variance of the OLS estimators can be expressed in terms of the unknown variance of the error term, $VAR(\epsilon_i)$, which we usually denote by σ^2. For the regression with K=1, the variance of the OLS estimator for the slope coefficient is

$$VAR(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{N}(X_i - \bar{X})^2} \tag{1}$$

For the purpose of interpreting this, it is better to multiply and divide the denominator by N, so that we have the sample size multiplied by the sample variance of X in the denominator.

$$VAR(\hat{\beta}_1) = \frac{\sigma^2}{N[\sum_{i=1}^{N}(X_i - \bar{X})^2)/N]} \tag{2}$$

If $\sigma^2 = VAR(\epsilon_i)$ increases, then $VAR(\hat{\beta}_1)$ increases. If the sample size N increases, then $VAR(\hat{\beta}_1)$ decreases. If the sample variance of X increases, then $VAR(\hat{\beta}_1)$ decreases. Do these statements make sense? why?

2 The Gauss-Markov Theorem

The OLS estimators have the minimum variance among the unbiased linear estimators.

An estimator which is linear, unbiased and has the minimum variance among all linear unbiased estimators is called the *Best Linear Unbiased Estimator* (BLUE).

Gauss-Markov Theorem: Under the Classical Assumptions I-VI, the Ordinary Least Squares (OLS) estimator of $\hat{\beta}_k$ is the Best Linear Unbiased Estimators (BLUE).

Example (The Two-Point Estimator): As an example of an alternative estimator to OLS when K=1, consider fitting a line that passes through two points you choose, (X_i, Y_i) and (X_j, Y_j). This line specifies intercept and slope coefficients. These coefficients can be considered as estimators for β_0 and β_1. We will focus on the slope coefficient. Let the slope coefficient denoted by $\hat{\beta}^*{}_1$. We call this a two-points estimator. Note that

$$\hat{\beta}_1^* = \frac{Y_j - Y_i}{X_j - X_i}, \tag{3}$$

when $X_j > X_i$. This is an alternative estimator for the OLS estimator for the coefficient $\hat{\beta}_1$. Under the Classical Assumptions, the two-point estimator is a linear unbiased estimator with the variance

$$VAR(\hat{\beta}^*{}_1) = \frac{2\sigma^2}{(X_j - X_i)^2} \tag{4}$$

The OLS use all data points, while an two-point estimator only uses two data points. So the OLS estimator is likely to be more precise than an two-point estimator. This intuition means that

$$VAR(\hat{\beta}_1) \leq VAR(\hat{\beta^*}_1). \tag{5}$$

The Gauss-Markov Theorem tells us that the OLS estimator has a smaller variance than any other linear unbiased estimator. This is consistent with the intuition for the two-point estimator.

Exercise (Demand for an R.E.M. CD): Let Y be the demand for an R.E.M. CD and X be the price of the CD. Imagine that the expected demand is given by

$$E(Y|X) = 93 - 5X \tag{6}$$

At each level of the price, actual demand depends on whether or not it rains. The probability that it rains is $1/2$. If it rains, demand falls from the expected value by 7 copies; and if it does not rain, demand rises from the expected value by 7 copies. We have data for three levels of the price, \$3, \$10, and \$15. The expected demand is as in Table 4.A.

Table 4.A Expected Demand for an R.E.M. CD

Observation i (1)	Price X_i (2)	The expected number of copies demanded $E(Y_i \mid X_i)$ (3)
1	3	78
2	12	33
3	15	18

(a) Compute the variance of the error term, $VAR(\epsilon_i)$.

(b) Compute the variance of the OLS estimator for the slope coefficient, $VAR(\hat{\beta}_1)$.

(c) Compute the variance of the two-point estimator when you choose i=2, and j=3.

(d) Compute the variance of the two-point estimator when you choose i=1, and j=2.

(e) Compute the variance of the two-point estimator when you choose i=1, and j=3.

(f) Compare your results for (c), (d), and (e). Which two-point estimator has the maximum variance among these? Which two-point estimator has the minimum variance among these?

(g) Give an intuitive explanation for your answer to (f).

(h) Compare your results for (b), (c), (d), and (e). Which estimator has the minimum variance among these?

(i) Is your answer to (h) consistent with the Gauss-Markov Theorem?