The Classical Linear Regression Model (CLRM)

The CLRM makes the following assumptions:

A-1: The regression model is *linear in the parameters* as in

$$Y_i = B_1 + B_2X_{2i} + B_3X_{3i} + \cdots + B_kX_{ki} + u_i$$

it may or may not be linear in the variables *Y* and the *X*s.

A-2: The regressors are assumed to be fixed or **nonstochastic** in the sense that their values are fixed in repeated sampling. This assumption may not be appropriate for all economic data, but as we will show later, if ***X*** and *u* are *independently distributed* the results based on the classical assumption discussed below hold true provided our analysis is conditional on the particular *X* values drawn in the sample. However, if *X* and *u* are *uncorrelated*, the classical results hold true asymptotically (i.e.in large samples.)[1]

A-3: Given the values of the *X* variables, the expected, or mean, value of the error term is zero. That is,[2]

$$E(u_i|\mathrm{X}) = 0 \qquad (1.8)$$

where, for brevity of expression, ***X*** (the bold *X*) stands for all *X* variables in the model. In words, the **conditional expectation** of the error term, given the values of the *X* variables, is zero. Since the error term represents the influence of factors that may be essentially random, it makes sense to assume that their mean or average value is zero.
As a result of this *critical* assumption, we can write (1.2) as:

$$\begin{aligned} E(Y_i|\mathrm{X}) &= \boldsymbol{BX} + E(u_i|\mathrm{X}) \\ &= \boldsymbol{BX} \qquad (1.9) \end{aligned}$$

which can be interpreted as the model for *mean* or *average* value of Y_i conditional on the ***X*** values. **This is the population (mean) regression function** (**PRF**) mentioned earlier. In regression analysis our main objective is to estimate this function. If there is only one *X* variable, you can visualize it as the (population) regression line. If there is more than one *X* variable, you will have to imagine it to be a curve in a multi-dimensional graph. The estimated PRF, the sample counterpart of Eq. (1.9), is denoted by $\hat{Y}_i = bx$. That is, $\hat{Y}_i = bx$ is an estimator of $E(Y_i|X)$.

[1] Note that independence implies no correlation, but no correlation does not necessarily imply independence.
[2] The vertical bar after u_i is to remind that the analysis is conditional on the given values of *X*.

A-4: The variance of each u_i given the values of $\boldsymbol{X}$, is constant, or **homoscedastic** (*homo* means equal and *scedastic* means variance). That is,

$$var(u_i|X) = \sigma^2 \qquad (1.10)$$

Note: There is no subscript on σ^2.

A-5: There is no correlation between two error terms. That is, there is no **autocorrelation**. Symbolically,

$$\text{Cov}\,(u_i, u_j \mid \boldsymbol{X}) = 0 \quad i \neq j \qquad (1.11)$$

where Cov stands for covariance and *i* and *j* are two different error terms. Of course, if *I* = *j*, Eq. (1.11) will give the variance of u_i given in Eq. (1.10).

A-6: There are no perfect linear relationships among the *X* variables. This is the assumption of **no multicollinearity**. For example, relationships like $X_5 = 2X_3 + 4X_4$ are ruled out.

A-7: The regression model is *correctly specified*. Alternatively, there is no **specification bias** or **specification error** in the model used in empirical analysis. It is implicitly assumed that the number of observations, *n*, is greater than the number of parameters estimated.
Although it is not a part of the CLRM, it is assumed that the error term follows the **normal distribution** with zero mean and (constant) variance σ^2. Symbolically,

A-8: $u_i \sim N(0, \sigma^2)$ (1.12)

On the basis of Assumptions A-1 to A-7, it can be shown that the method of **ordinary least squares** (**OLS**), the method most popularly used in practice, provides estimators of the parameters of the PRF that have several *desirable statistical properties*, such as:

1- The **estimators are *linear, that is, they are linear functions of the dependent variable Y*. Linear estimators are easy to understand and deal with compared to nonlinear estimators.**

2- The **estimators are *unbiased, that is, in repeated applications of the method, on average, the estimators are equal to their true values.***

3- In the class of linear unbiased estimators, OLS estimators have minimum variance. As a result, the true parameter values can be estimated with least possible uncertainty; an unbiased estimator with the least variance is called an ***efficient estimator.***

In short, under the assumed conditions, OLS estimators are **BLUE**: **best linear unbiased estimators**. This is the essence of the well-known **Gauss–Markov theorem**, which provides a theoretical justification for the method of least squares.

With the added Assumption **A-8**, it can be shown that the OLS estimators are themselves normally distributed. As a result, we can draw inferences about the true values of the population regression coefficients and test statistical hypotheses. With the added assumption of normality, the OLS estimators are **best unbiased estimators** (**BUE**) in the entire class of unbiased estimators, whether linear or not. With normality assumption, CLRM is known as the **normal classical linear regression model** (**NCLRM**).

References:

- Gujarati, D., ***Econometrics by Example***, 2012. McGraw Hill.