Probability of the Error Term

***Assumption**

In a simple or classical linear regression model of the form $Y = \beta_1 + \beta_2X + U_i$, ($i = 1,2,3,..,n$); we know that U_i is called the error term or the disturbance is a stochastic or random variable. This stochastic variable is assumed to follow a **normal** probability distribution.

***Characteristics**

Due to the **normal** probability distribution assumption of the error term, we have the following characteristics:

- Y and the sampling distribution of the parameters of the regression are also normally distributed and tests can be conducted on the significance of the parameters.
- The expected value of the error term or its mean equals to zero: $E(U_i) = 0$
- The variance of the error term is constant in each period and for all values of X: $E(U_i^2) = \sigma^2$. This feature ensures that estimates of the regression equation coefficients are efficient and tests of hypothesis are not biased. We can express the error term as

  $U_i \sim N(0, \sigma^2)$

  where the symbol ~ means 'distributed as' and N stands for 'normal distribution', the terms in parentheses representing two parameters of the normal distribution, namely the mean and the variance.
- The value which the error term assumes in one period is uncorrelated to its value in any other period: $E(U_i, U_j) = 0$, for $i \neq j$; $i,j = 1, 2, 3, \ldots, n$. This ensures that the average value of Y depends only on X and not on U, and its required in order that we get efficient estimates of the regression coefficients and unbiased tests of significance.
- The explanatory variable assumes fixed values that can be obtained in repeated samples, so that the explanatory variable is also uncorrelated with the error term: $E(X_i, U_i) = 0$.

***Reasons for normality assumption of the error term**

1. The **Central Limit Theorem** provides a theoretical justification for the assumption of normality of the error term. According to the **Central Limit Theorem,** if there are a large number of independent and identically distributed random variables, then, with a

few exceptions, the distribution of their sum tends to a normal distribution as the number of such variables increases indefinitely.

2. The normal distribution is a simple distribution involving only two parameters the mean and the variance.
3. One property of normal distribution is that any linear function of normally distributed variables is itself normally distributed. Since one of the properties of the OLS estimators is that the estimators are linear function of the error term, hence normality assumption of the error term becomes relevant.
4. If we deal with a small or finite data with less than 100 observations, the normality assumption helps to derive the exact probability distribution of the OLS estimators and use the t, *F* and χ^2 statistical tests for regression models.
5. In large samples, t and *F* statistics have approximately the t and *F* probability distributions so that the tests that are based on the assumption that the error term is normally distributed have validity in application

*N.B. Explanations and all theoretical disposition have been taken from a combination of the following references:

References:

1. Gujarati, D. N. Porter, D.C., Gunasekar, S. (2009), *Basic econometrics*. (Fifth ed.) McGraw-Hill Education (India).
2. Salvatore, D., & Reagle, D. (2011). *Schaum's outline of statistics and econometrics* (2nd ed.). McGraw-Hill Education.