**ECON 351* -- NOTE 1**

# Specification -- Assumptions of the Simple Classical Linear Regression Model (CLRM)

## 1. Introduction

**CLRM** stands for the **Classical Linear Regression Model**. The CLRM is also known as the ***standard* linear regression model**.

**Three sets of assumptions define the CLRM.**

**1.** Assumptions respecting the **formulation of the *population regression equation,* or PRE**.

*Assumption A1*

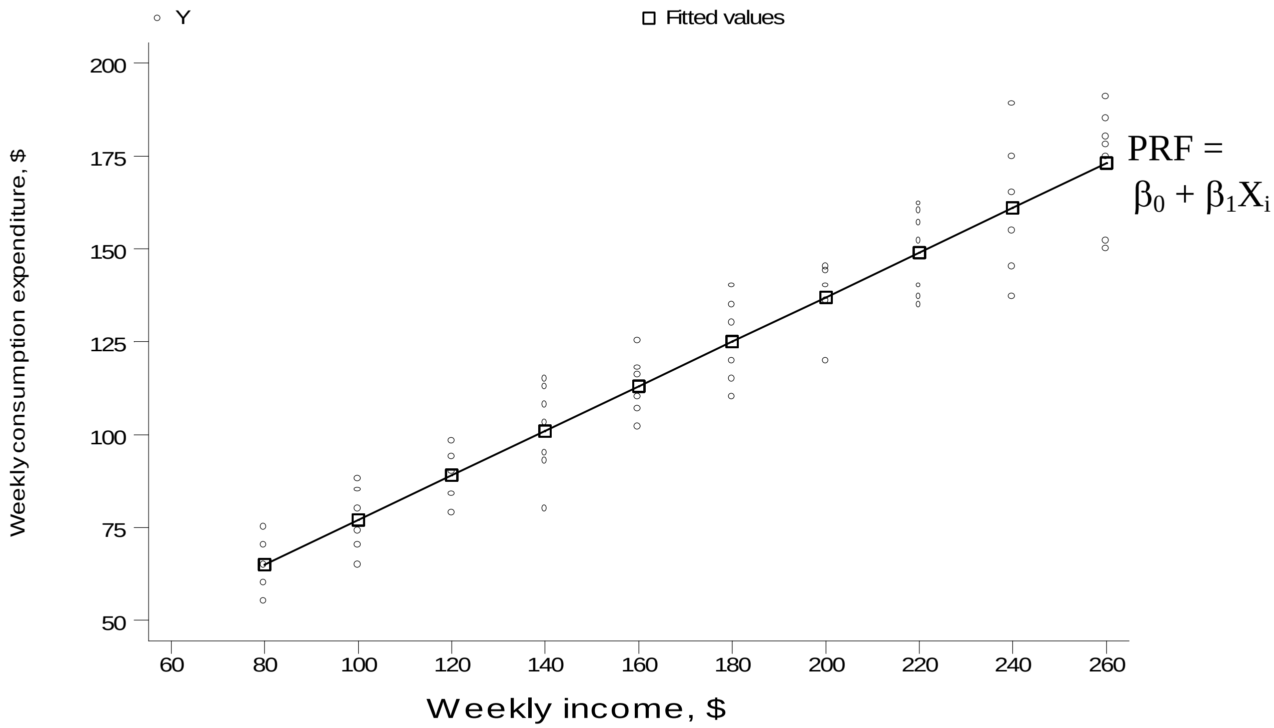
**2.** Assumptions respecting the **statistical properties of the *random error term* and the *dependent variable***.

*Assumptions A2-A4*

**3.** Assumptions respecting the **properties of the *sample data***.

*Assumptions A5-A8*

- **Figure 2.1 Plot of Population Data Points, Conditional Means E(Y|X), and the Population Regression Function PRF**




Recall that the solid line in Figure 2.1 is the population regression function, which takes the form $f(X_i) = E(Y_i \mid X_i) = \beta_0 + \beta_1 X_i$.

For **each population value $X_i$ of X**, there is **a *conditional* distribution of population Y values** and a corresponding ***conditional* distribution of population random errors u**, where

(1) **each population value of u for X = $X_i$** is

$$u_i \mid X_i = Y_i - E(Y_i \mid X_i) = Y_i - \beta_0 - \beta_1 X_i,$$

and

(2) **each population value of Y for X = $X_i$** is

$$Y_i \mid X_i = E(Y_i \mid X_i) + u_i = \beta_0 + \beta_1 X_i + u_i.$$

## 2. Formulation of the Population Regression Equation (PRE)

**<u>Assumption A1</u>: The population regression equation, or PRE, takes the form**

$$\mathbf{Y = \beta_0 + \beta_1 X + u} \qquad \textbf{\textit{or}} \qquad \mathbf{Y_i = \beta_0 + \beta_1 X_i + u_i} \qquad \textbf{(A1)}$$

- The PRE (A1) gives the value of the **regressand (dependent variable) Y** for each value of the **regressor (independent variable) X**. The "i" subscripts on Y and X are used to denote individual population or sample values of the dependent variable Y and the independent variable X.

- The PRE (A1) states that **each value $\mathbf{Y_i}$** of the dependent variable Y can be written as **the *sum* of two parts**.

  1. A linear function of the independent variable X that is called the ***population regression function*** (or ***PRF***).

     The PRF for $Y_i$ takes the form

     $$\mathbf{f(X_i) = \beta_0 + \beta_1 X_i}$$

     where

     $\beta_0$ and $\beta_1$ are **regression *coefficients*** (or parameters), the true population values of which are unknown, and

     $X_i$ is the value of the regressor X corresponding to the value $Y_i$ of Y.

  2. A ***random error term*** $\mathbf{u_i}$ (also called a stochastic error term).

     Each random error term $u_i$ is the difference between the observed $Y_i$ value and the value of the population regression function for the corresponding value $X_i$ of the regressor X:

     $$\mathbf{u_i = Y_i - f(X_i) = Y_i - (\beta_0 + \beta_1 X_i) = Y_i - \beta_0 - \beta_1 X_i}$$

     The random error terms are ***unobservable*** because the true population values of the regression coefficients $\beta_0$ and $\beta_1$ are unknown.

The PRE (A1) incorporates ***three*** **distinct assumptions**.

$$Y_i = \beta_0 + \beta_1 X_i + u_i \qquad \textbf{(A1)}$$

**A1.1: <u>Assumption of an Additive Random Error Term</u>.**

⇒ **The random error term $u_i$ enters the PRE *<u>additively</u>*.**

Technically, this assumption means that the partial derivative of $Y_i$ with respect to $u_i$ equals 1: i.e.,

$$\frac{\partial Y_i}{\partial u_i} = 1 \qquad \text{for all i } (\forall \text{ i}).$$

**A1.2: <u>Assumption of Linearity-in-Parameters or Linearity-in-Coefficients</u>.**

⇒ **The PRE is *<u>linear</u>* in the population regression coefficients $\beta_j$ (j = 0, 1).**

This assumption means that the partial derivative of $Y_i$ with respect to each of the regression coefficients is a function only of known constants and/or the regressor $X_i$; it is not a function of any unknown parameters.

$$\frac{\partial Y_i}{\partial \beta_j} = f_j(X_i) \quad j = 0, 1 \quad \text{where } f_j(X_i) \textbf{ contains } \textbf{\textit{no unknown parameters}}.$$

**A1.3: <u>Assumption of Parameter or Coefficient Constancy</u>.**

⇒ **The population regression coefficients $\beta_j$ (j = 0, 1) are (unknown) *<u>constants</u>* that do not vary across observations.**

This assumption means that the regression coefficients $\beta_0$ and $\beta_1$ do not vary across observations -- i.e., do not vary with the observation subscript "i". Symbolically, if $\beta_{ji}$ is the value of the j-th regression coefficient for observation i, then assumption A1.3 states that

$$\beta_{ji} = \beta_j = \textbf{ a } \textbf{\textit{constant}} \ \forall \textbf{ i} \quad (j = 0, 1).$$

## 3. Properties of the Random Error Term

**<u>Assumption A2</u>: The Assumption of *Zero Conditional Mean Error***

**The *conditional* mean, or *conditional* expectation, of the random error terms $u_i$ for any given value $X_i$ of the regressor X is equal to zero:**

$$E(u|X) = 0 \quad or \quad E(u_i|X_i) = 0 \;\; \forall\, i \;\; \text{(for all } X_i \text{ values)} \qquad \text{(A2)}$$

This assumption says two things:

1. the ***conditional*** **mean of the random error term u is the *same* for all population values of X** -- i.e., it does not depend, either linearly or nonlinearly, on X;
2. the common **conditional population mean of u** for all values of X is ***zero***.

**<u>Implications of Assumption A2</u>**

- **<u>Implication 1 of A2</u>.** Assumption A2 implies that the ***unconditional*** **mean** of the population values of **the random error term u equals *zero***:

$$E(u|X) = 0 \quad \Rightarrow \quad E(u) = 0 \qquad \text{(A2-1)}$$

***or***

$$E(u_i|X_i) = 0 \;\Rightarrow\; E(u_i) = 0 \quad \forall\, i. \qquad \text{(A2-1)}$$

This implication follows from the so-called **law of iterated expectations**, which states that $E[E(u|X)] = E(u)$. Since $E(u|X) = 0$ by A2, it follows that $E(u) = E[E(u|X)] = E[0] = 0$.

The logic of (A2-1) is straightforward: If the conditional mean of u for each and every population value of X equals zero, then the mean of these zero conditional means must also be zero.

- **<u>Implication 2 of A2: the Orthogonality Condition</u>.** Assumption A2 also implies that **the population values $X_i$ of the regressor X and $u_i$ of the random error term u have *zero covariance* -- i.e., the population values of X and u are *uncorrelated***:

$$\mathbf{E(u \mid X) = 0 \quad \Rightarrow \quad Cov(X,u) = E(Xu) = 0} \qquad \textbf{(A2-2)}$$

***or***

$$\mathbf{E(u_i \mid X_i) = 0 \quad \Rightarrow \quad Cov(X_i, u_i) = E(X_i u_i) = 0 \quad \forall\ i} \qquad \textbf{(A2-2)}$$

**1.** The equality $\text{Cov}(X_i, u_i) = E(X_i u_i)$ in (A2-2) follows from the definition of the covariance between $X_i$ and $u_i$, and from assumption (A2):

$$\begin{aligned}
\text{Cov}(X_i, u_i) &\equiv E\{[X_i - E(X_i)][u_i - E(u_i)]\} && \text{by definition} \\
&= E\{[X_i - E(X_i)]u_i\} && \text{since } E(u_i) = 0 \text{ by A2 and A2-1} \\
&= E[X_i u_i - E(X_i)u_i] && \\
&= E(X_i u_i) - E(X_i)E(u_i) && \text{since } E(X_i) \text{ is a constant} \\
&= E(X_i u_i) && \text{since } E(u_i) = 0 \text{ by A2 and A2-1.}
\end{aligned}$$

**2. Implication (A2-2)** states that the **population random error terms $u_i$** have ***zero covariance with***, or are ***uncorrelated with***, the **corresponding population regressor values $X_i$**. This assumption means that there exists ***no <u>linear</u> association* between $u_i$ and $X_i$**.

Note that ***zero covariance* between $X_i$ and $u_i$** implies ***zero correlation* between $X_i$ and $u_i$**, since the **simple *correlation coefficient* between $X_i$ and $u_i$**, denoted as $\rho(X_i, u_i)$, is defined as

$$\rho(X_i, u_i) \equiv \frac{\text{Cov}(X_i, u_i)}{\sqrt{\text{Var}(X_i)\text{Var}(u_i)}} = \frac{\text{Cov}(X_i, u_i)}{\text{sd}(X_i)\,\text{sd}(u_i)}.$$

From this definition of $\rho(X_i, u_i)$, it is obvious that $\text{Cov}(X_i, u_i) = 0$ implies that $\rho(X_i, u_i) = 0$, i.e.,

$$\text{Cov}(X_i, u_i) = 0 \Rightarrow \quad \rho(X_i, u_i) = 0.$$

- **<u>Implication 3 of A2</u>.** Assumption A2 implies that the **conditional mean of the population $Y_i$ values** corresponding to a given value $X_i$ of the regressor X ***equals*** the **population regression function (PRF),** $\mathbf{f(X_i) = \beta_0 + \beta_1 X_i}$:

$$\mathbf{E(u|X) = 0 \quad \Rightarrow \quad E(Y|X) = f(X) = \beta_0 + \beta_1 X} \qquad \textbf{(A2-3)}$$

$$\mathbf{E(u_i|X_i) = 0 \quad \Rightarrow \quad E(Y_i|X_i) = f(X_i) = \beta_0 + \beta_1 X_i \quad \forall\ i.} \qquad \textbf{(A2-3)}$$

***<u>Proof</u>:*** Take the conditional expectation of the PRE (A1) for some given $X_i$:

$$Y_i = \beta_0 + \beta_1 X_{1i} + u_i \qquad (i = 1, \ldots, N) \qquad \textbf{(A1)}$$

$$E(Y_i|X_i) = E(\beta_0 + \beta_1 X_i|X_i) + E(u_i|X_i)$$

$$= E(\beta_0 + \beta_1 X_i|X_i) \qquad \text{since } E(u_i|X_i) = 0 \text{ by assumption A2}$$

$$= \beta_0 + \beta_1 X_i \qquad \text{since } E(\beta_0 + \beta_1 X_i|X_i) = \beta_0 + \beta_1 X_i.$$

- **<u>Meaning of the Zero Conditional Mean Error Assumption A2</u>**

Each value $X_i$ of X identifies a segment or subset of the relevant population. For each of these population segments or subsets, assumption A2 says that the mean of the random error u is zero. In other words, for each population segment, positive and negative values of u "cancel out" so that the average value of the random errors $u_i$ for each population value $X_i$ of X equals zero.

Assumption A2 rules out both ***linear*** **dependence and *nonlinear* dependence between X and u;** that is, **it requires that X and u be *statistically independent***.

- The absence of linear dependence between X and u means that **X and u are *uncorrelated***, or equivalently that **X and u have *zero covariance***.

- But linear independence between X and u is not sufficient to guarantee the satisfaction of assumption A2. It is possible for X and u to be both uncorrelated, or linearly independent, and nonlinearly related.

- Assumption A2 therefore also requires that there be **no *nonlinear* relationship between X and u**.

- <u>**Violations of the Zero Conditional Mean Error Assumption A2**</u>

- The random error term u represents all the ***unknown, unobservable*** **and** ***unmeasured*** **variables** other than the regressor X that determine the population values of the dependent variable Y.

- Anything that causes the random error u to be correlated with the regressor X will violate assumption A2:

$$\text{Cov}(X, u) \neq 0 \;\; or \;\; \rho(X, u) \neq 0 \quad \Rightarrow \quad E(u \mid X) \neq 0 .$$

If X and u are correlated, then $E(u \mid X)$ must depend on X and so cannot be zero.

Note that the converse is not true:

$$\text{Cov}(X, u) = 0 \;\; or \;\; \rho(X, u) = 0 \quad \text{does } \textit{not} \text{ imply that} \quad E(u \mid X) = 0 .$$

<u>*Reason*</u>: $\text{Cov}(X, u)$ measures only ***linear*** **dependence** between u and X. But any ***nonlinear*** **dependence** between u and X will also cause $E(u \mid X)$ to depend on X, and hence to differ from zero. So $\text{Cov}(X, u) = 0$ is not enough to insure that assumption A2 is satisfied.

- **Common causes of correlation or dependence between X and u --** i.e., common causes of violations of assumption A2.

1. **Incorrect specification of the functional form of the relationship between Y and X.**

   *Examples:* Using Y as the dependent variable when the true model has ln(Y) as the dependent variable. Or using X as the independent variable when the true model has ln(X) as the independent variable.

2. **Omission of relevant variables that are correlated with X.**

3. **Measurement errors in X.**

4. **Joint determination of X and Y.**

**Assumption A3: The Assumption of *Constant Error Variances***
**The Assumption of *Homoskedastic Errors***
**The Assumption of *Homoskedasticity***

**The *conditional* variances of the random error terms $u_i$ are identical for all observations (i.e., for all population values $X_i$ of X), and equal the same finite positive constant $\sigma^2$ for all i:**

$$\mathbf{Var(u \mid X) = E(u^2 \mid X) = \sigma^2 > 0} \qquad \textbf{(A3)}$$

***or***

$$\mathbf{Var(u_i \mid X_i) = E(u_i^2 \mid X_i) = \sigma^2 > 0 \quad \forall\, i \;\; (for\ all\ X_i\ values)} \qquad \textbf{(A3)}$$

**where $\sigma^2$ is a *finite positive (unknown) constant*.**

- The **first equality in (A3)** follows from the definition of the conditional variance of $u_i$ and assumption (A2):

$$\begin{aligned} \mathrm{Var}(u_i \mid X_i) &\equiv E\{[u_i - E(u_i \mid X_i)]^2 \mid X_i\} && \text{by definition} \\ &= E\{[u_i - 0]^2 \mid X_i\} && \text{because } E(u_i \mid X_i) = 0 \text{ by assumption A2} \\ &= E(u_i^2 \mid X_i). \end{aligned}$$

- **Implication 1 of A3:** Assumption A3 implies that the ***unconditional* variance of the random error u** is also equal to $\sigma^2$:

$$\mathrm{Var}(u_i) = E[(u_i - E(u_i))^2] = E(u_i^2) = \sigma^2 .$$

where $\mathrm{Var}(u_i) = E(u_i^2)$ because $E(u_i) = 0$ by A2-1.

By assumptions A2 and A3, $E(u_i^2 \mid X) = \sigma^2$.
By the law or iterated expectations, $E[E(u_i^2 \mid X_i)] = E(u_i^2)$.
Thus,

$$\mathrm{Var}(u_i) = E(u_i^2) = E[E(u_i^2 \mid X_i)] = E[\sigma^2] = \sigma^2 .$$

- **Implication 2 of A3:** Assumption A3 implies that the conditional variance of the population values $Y_i$ of the regressand Y corresponding to any given value $X_i$ of the regressor X equals the constant conditional error variance $\sigma^2$:

$$\operatorname{Var}\left(u_i \middle| X_i\right) = \sigma^2 \quad \forall\ i \qquad \Rightarrow \qquad \operatorname{Var}\left(Y_i \middle| X_i\right) = \sigma^2 \quad \forall\ i \qquad \textbf{(A3-2)}$$

***Proof:*** Start with the definition of the conditional variance of $Y_i$ for some given $X_i$:

$$\begin{aligned}
\operatorname{Var}\left(Y_i \middle| X_i\right) &\equiv E\left\{\left[Y_i - E(Y_i|X_i)\right]^2 \middle| X_i\right\} && \text{by definition} \\
&= E\left\{\left[Y_i - \beta_0 - \beta_1 X_i\right]^2 \middle| X_i\right\} && \text{since } E(Y_i|X_i) = \beta_0 + \beta_1 X_i \text{ by A2} \\
&= E\left(u_i^2 \middle| X_i\right) && \text{since } u_i = Y_i - \beta_0 - \beta_1 X_i \text{ by A1} \\
&= \sigma^2 && \text{since } E\left(u_i^2 \middle| X_i\right) = \sigma^2 \text{ by assumption A3.}
\end{aligned}$$

- **Meaning of the Homoskedasticity Assumption A3**

- For each population value $X_i$ of X, there is a ***conditional*** **distribution of random errors**, and a corresponding ***conditional*** **distribution of population Y values**.

- Assumption A3 says that the ***variance*** **of the random errors for X = $X_i$** is the ***same*** as the ***variance*** **of the random errors for any other regressor value X = $X_s$** (for all $X_s \neq X_i$). That is, the ***variances*** of the ***conditional*** **random error distributions** corresponding to each population value of X **are all *equal* to the *same* finite positive constant** $\sigma^2$.

$$\operatorname{Var}\left(u_i \middle| X_i\right) = \operatorname{Var}\left(u_s \middle| X_s\right) = \sigma^2 > 0 \quad \text{for all} \quad X_s \neq X_i.$$

- Implication A3-2 says that the ***variance*** **of the population Y values for X = $X_i$** is the ***same*** as the ***variance*** **of the population Y values for any other regressor value X = $X_s$** (for all $X_s \neq X_i$). The ***conditional distributions*** **of the population Y values** around the PRF have the ***same constant*** **variance** $\sigma^2$ for all population values of X.

$$\operatorname{Var}\left(Y_i \middle| X_i\right) = \operatorname{Var}\left(Y_s \middle| X_s\right) = \sigma^2 > 0 \quad \text{for all} \quad X_s \neq X_i.$$

**<u>Assumption A4</u>: The Assumption of *Zero Error Covariances***
**The Assumption of *Nonautoregressive Errors***
**The Assumption of *Nonautocorrelated Errors***

**Consider the random error terms $u_i$ and $u_s$ ($i \neq s$) corresponding to two different population values $X_i$ and $X_s$ of the regressor X, where $X_i \neq X_s$. This assumption states that $u_i$ and $u_s$ have zero conditional covariance:**

$$\mathbf{Cov}\left(u_i, u_s \mid X_i, X_s\right) = \mathbf{E}\left(u_i u_s \mid X_i, X_s\right) = 0 \quad \forall \; i \neq s \;\; (\text{for all } X_i \neq X_s) \qquad \textbf{(A4)}$$

- The **first equality in (A4)** follows from the definition of the conditional covariance of $u_i$ and $u_s$ and assumption (A2):

$$\begin{aligned} \mathrm{Cov}\left(u_i, u_s \mid X_i, X_s\right) &\equiv E\left\{\left[u_i - E(u_i \mid X_i)\right]\left[u_s - E(u_s \mid X_s)\right] \mid X_i, X_s\right\} && \text{by definition} \\ &= E\left(u_i u_s \mid X_i, X_s\right) && \text{since } E(u_i \mid X_i) = E(u_s \mid X_s) = 0 \text{ by A2.} \end{aligned}$$

- The **second equality in (A4)** states the assumption that all pairs of error terms corresponding to different values of X have zero covariance.

- **<u>Implication of A4</u>:** Assumption A4 implies that the conditional covariance between the population values $Y_i$ of Y when $X = X_i$ and the population values $Y_s$ of Y when $X = X_s$ where $X_i \neq X_s$ is equal to zero:

$$\mathbf{Cov}\left(u_i, u_s \mid X_i, X_s\right) = 0 \quad \Rightarrow \quad \mathbf{Cov}\left(Y_i, Y_s \mid X_i, X_s\right) = 0 \qquad \forall \; X_i \neq X_s.$$

***Proof:***

(1) Begin with the definition of the conditional covariance for $Y_i$ and $Y_s$ for given $X_i$ and $X_s$ values:

$$\begin{aligned} \mathrm{Cov}\left(Y_i, Y_s \mid X_i, X_s\right) &\equiv E\left\{\left[Y_i - E(Y_i \mid X_i)\right]\left[Y_s - E(Y_s \mid X_s)\right] \mid X_i, X_s\right\} \\ &= E\left(u_i u_s \mid X_i, X_s\right) \end{aligned}$$

since

$$Y_i - E(Y_i \mid X_i) = Y_i - \beta_0 - \beta_1 X_i = u_i \qquad \text{by assumptions A1 and A2,}$$

and similarly

$$Y_s - E(Y_s \mid X_s) = Y_s - \beta_0 - \beta_1 X_s = u_s \qquad \text{by assumptions A1 and A2.}$$

(2) Therefore

$$\mathrm{Cov}\left(Y_i, Y_s \mid X_i, X_s\right) = E\left(u_i u_s \mid X_i, X_s\right) = 0 \qquad \text{by assumption A4.}$$

- **Meaning of A4:** Assumption A4 states that the random error terms $u_i$ corresponding to $X = X_i$ have ***zero covariance with***, or are ***uncorrelated with***, the random error terms $u_s$ corresponding to any other regressor value $X = X_s$, where $X_i \neq X_s$. Equivalently, assumption A4 states that the **population values $Y_i$ of Y** corresponding to $X = X_i$ have ***zero covariance with***, or are ***uncorrelated with***, the **population values $Y_s$ of Y** corresponding to $X = X_s$ for any distinct pair of regressor values $X_i \neq X_s$. This means there is ***no systematic linear dependence or association* between $u_i$ and $u_s$, or between $Y_i$ and $Y_s$**, where i and s correspond to different observations (that is, to different regressor values $X_i \neq X_s$).

  - The **assumption of *zero covariance*, or *zero correlation*,** between each pair of distinct observations is ***weaker*** than the **assumption of *independent random sampling A5*** from an underlying population.

  - The assumption of independent random sampling implies that the sample observations are statistically independent. The assumption of statistically independent observations is *sufficient* for the assumption of zero covariance between observations, but is stronger than necessary.

# 4. Properties of the Sample Data

**<u>Assumption A5</u>: Random Sampling *or* Independent Random Sampling**

The **sample data** consist of **N *randomly selected observations*** on the regressand Y and the regressor X, the two observable variables in the PRE described by A1.

In other words, the sample observations are **randomly selected** from the underlying population; they are a **random subset** of the population **data points**.

These N randomly selected sample observations on Y and X are written as the N pairs

$$\textbf{Sample data} \equiv [(Y_1, X_1), (Y_2, X_2), \ldots (Y_N, X_N)]$$
$$\equiv (Y_i, X_i) \qquad i = 1, \ldots, N.$$

- **<u>Implications of the Random Sampling Assumption A5</u>**

  The **assumption of random sampling** implies that **the sample observations are *statistically independent***.

  **1.** It thus means that the error terms **$u_i$ and $u_s$ are *statistically independent***, and hence have zero covariance, for any two observations i and s.

  $$\text{Random sampling} \Rightarrow \text{Cov}(u_i, u_s \mid X_i, X_s) = \text{Cov}(u_i, u_s) = 0 \qquad \forall \; i \neq s.$$

  **2.** It also means that the dependent variable values **$Y_i$ and $Y_s$ are *statistically independent***, and hence have zero covariance, for any two observations i and s.

  $$\text{Random sampling} \Rightarrow \text{Cov}(Y_i, Y_s \mid X_i, X_s) = \text{Cov}(Y_i, Y_s) = 0 \qquad \forall \; i \neq s.$$

  The assumption of random sampling is therefore sufficient for assumption A4 of zero covariance between observations, but is stronger than is necessary for A4.

- **<u>When is the Random Sampling Assumption A5 Appropriate?</u>**

  The random sampling assumption is usually appropriate for ***cross-sectional*** **regression models**, i.e., for regression models formulated for ***cross section data***.

  - ***Definition:*** A **cross-sectional data set** consists of a sample of observations on individual economic agents or other units taken at a **single *point* in time** or over a **single *period* of time**.

  - A ***distinguishing characteristic*** of any cross-sectional data set is that the individual observations have **no natural ordering**.

  - A ***common, almost universal characteristic*** of cross-sectional data sets is that they usually are constructed by **random sampling** from underlying populations.

  The random sampling assumption is hardly ever appropriate for ***time-series*** **regression models**, i.e., for regression models formulated for ***time series data***.

  - ***Definition:*** A **time-series data set** consists of a sample of observations on one or more variables over several successive periods or intervals of time.

  - A ***distinguishing characteristic*** of any time-series data set is that the observations have a natural ordering -- specifically a **chronological ordering**.

  - A ***common, almost universal characteristic*** of time-series data sets is that the **sample observations exhibit a high degree of *time dependence***, and therefore cannot be assumed to be generated by random sampling.

<u>**Assumption A6**</u>**: The number of sample observations N is greater than the number of unknown parameters K:**

**number of sample observations > number of unknown parameters**

$$N > K. \qquad \textbf{(A6)}$$

- <u>**Meaning of A6**</u>**:** Unless this assumption is satisfied, it is not possible to compute from a given sample of N observations estimates of all the unknown parameters in the model.

<u>**Assumption A7**</u>**: Nonconstant Regressor**

**The sample values $X_i$ of the regressor X in the sample (and hence in the population) are not all the same; i.e., they are not constant:**

$$X_i \neq c \quad \forall\ i = 1, ..., N \qquad \text{where } c \text{ is a constant.} \qquad \textbf{(A7)}$$

- <u>**Technical Form of A7**</u>**:** Assumption A7 requires that the **sample variance of the regressor values $X_i$** (i = 1, ..., N) must be **a *finite positive* number** for any sample size N; i.e.,

  $$\text{sample variance of } X_i \equiv \mathrm{Var}(X_i) = \frac{\sum_i (X_i - \bar{X})^2}{N-1} = s_X^2 > 0,$$

  where $s_X^2 > 0$ is **a *finite positive* number**.

- <u>**Meaning of A7**</u>**:** Assumption A7 requires that the regressor sample values $X_i$ take **at least two different values** in any given sample.

  Unless this assumption is satisfied, it is not possible to compute from the sample data an estimate of the effect on the regressand Y of changes in the value of the regressor X. In other words, to calculate the effect of changes in X on Y, the sample values $X_i$ of the regressor X must vary across observations in any given sample.

**Assumption A8: No Perfect Multicollinearity**

**The sample values of the regressors in a multiple regression model do not exhibit perfect multicollinearity.**

This assumption is relevant only in multiple regression models that contain two or more non-constant regressors.

Its nature is examined later in the context of multiple regression models.