ECON 351* -- NOTE 5

Computational Properties and Goodness-of-Fit of the OLS Sample Regression Equation

Outline of Note 5

- State and prove the **five computational properties** of the **OLS SRE**

$$\begin{aligned} Y_i &= \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i \\ &= \hat{Y}_i + \hat{u}_i \end{aligned} \qquad (i = 1, ..., N) \qquad (1)$$

- Derive and interpret the **OLS decomposition equation**, which looks like this:

$$\sum_{i=1}^{N} y_i^2 = \sum_{i=1}^{N} \hat{y}_i^2 + \sum_{i=1}^{N} \hat{u}_i^2 \qquad (5.1)$$

or

$$TSS = ESS + RSS \qquad (5.2)$$

- Define and interpret the **goodness-of-fit measure** called $\mathbf{R^2}$ **(R-squared)**, which is defined as

$$R^2 \equiv \frac{\sum_i \hat{y}_i^2}{\sum_i y_i^2} = 1 - \frac{\sum_i \hat{u}_i^2}{\sum_i y_i^2}$$

or

$$R^2 \equiv \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

Starting Point

The **OLS sample regression equation (OLS-SRE)** is

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i = \hat{Y}_i + \hat{u}_i \qquad (i = 1, ..., N) \qquad (1)$$

where

$\hat{\beta}_0 =$ the OLS estimate of the ***intercept*** **coefficient β_0**;

$\hat{\beta}_1 =$ the OLS estimate of the ***slope*** **coefficient β_1**;

$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ = the i-th estimated (or predicted) value of $E(Y_i \mid X_i) = \beta_0 + \beta_1 X_i$, and is called the **OLS sample regression function** (or **OLS-SRF**);

$\hat{u}_i = Y_i - \hat{Y}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$ = the i-th **OLS residual**.

The OLS sample regression equation (1) exhibits ***five*** **computational properties**. These computational properties are necessary for developing goodness-of-fit measures such as the coefficient of determination, R^2.

Recall that the **OLS normal equations** for the simple (two-variable) linear regression model are:

$$N\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{N} X_i = \sum_{i=1}^{N} Y_i \qquad \textbf{(N1)}$$

$$\hat{\beta}_0 \sum_{i=1}^{N} X_i + \hat{\beta}_1 \sum_{i=1}^{N} X_i^2 = \sum_{i=1}^{N} X_i Y_i \qquad \textbf{(N2)}$$

1. Computational Properties of the OLS SRE

<u>PROPERTY 1</u>: The *OLS sample regression equation* passes through the *point of sample means* $(\bar{\mathbf{Y}}, \bar{\mathbf{X}})$, where

$\bar{Y} = \sum_{i=1}^{N} Y_i / N$ is the sample mean value of Y; and

$\bar{X} = \sum_{i=1}^{N} X_i / N$ is the sample mean value of X.

That is,

$$\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X} \qquad \textbf{... (C1)}$$

- **<u>Proof of (C1):</u>** Follows from the first OLS normal equation (N1)

$$\Sigma_i Y_i = N\hat{\beta}_0 + \hat{\beta}_1 \Sigma_i X_i . \qquad \text{(N1)}$$

Dividing both sides of equation (N1) by N yields

$$\frac{\sum_i Y_i}{N} = \hat{\beta}_0 + \hat{\beta}_1 \frac{\sum_i X_i}{N}$$

or, using the definitions of $\bar{Y}$ and $\bar{X}$,

$$\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X} . \qquad \textbf{... (C1)}$$

<u>PROPERTY 2</u>: The **sample mean of the *estimated* Y_i's** (the $\hat{Y}_i$**'s**) ***equals*** the **sample mean of the *observed* Y_i's**; or the **sum of the *estimated* Y_i's** (the $\hat{Y}_i$**'s**) ***equals*** the **sum of the *observed* Y_i's**.

$$\bar{\hat{Y}} = \bar{Y} \qquad \text{where } \bar{\hat{Y}} \equiv \Sigma_i \hat{Y}_i / N \text{ and } \bar{Y} = \Sigma_i Y_i / N \qquad \text{... (C2)}$$

or

$$\sum_{i=1}^{N} \hat{Y}_i = \sum_{i=1}^{N} Y_i$$ sum of *estimated* Y_i's, (the $\hat{Y}_i$'s) = sum of *observed* Y_i's.

- **<u>Proof of (C2)</u>:**

(1) The estimated values of Y_i are given by

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i .$$

(2) Substitute for $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$ in the above expression for $\hat{Y}_i$:

$$\begin{aligned} \hat{Y}_i &= \bar{Y} - \hat{\beta}_1 \bar{X} + \hat{\beta}_1 X_i \\ &= \bar{Y} + \hat{\beta}_1 \left(X_i - \bar{X} \right) \end{aligned}$$

(3) Now sum both sides over i = 1, ..., N:

$$\begin{aligned} \sum_{i=1}^{N} \hat{Y}_i &= N\bar{Y} + \hat{\beta}_1 \left(\sum_{i=1}^{N} X_i - N\bar{X} \right) \\ &= N\bar{Y} + \hat{\beta}_1 \left(N\bar{X} - N\bar{X} \right), && \text{since } \sum_{i=1}^{N} X_i = N\bar{X} \\ &= N\bar{Y}, && \text{since } \left(N\bar{X} - N\bar{X} \right) = 0. \end{aligned}$$

(4) Finally, dividing by N, we get

$$\frac{\sum_{i=1}^{N} \hat{Y}_i}{N} = \bar{Y} \quad \Rightarrow \quad \bar{\hat{Y}} = \bar{Y} \quad \Rightarrow \quad \sum_{i=1}^{N} \hat{Y}_i = \sum_{i=1}^{N} Y_i .$$

- **Implication of Property C2:** The OLS-SRF $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ can be written in **deviation-from-means form** as

$$\hat{y}_i = \hat{\beta}_1 x_i \qquad \text{where } \hat{y}_i \equiv \hat{Y}_i - \bar{Y} \text{ and } x_i \equiv X_i - \bar{X}.$$

Proof:

(1) From line (2) of the proof of Property (C2) above,

$$\hat{Y}_i = \bar{Y} + \hat{\beta}_1 (X_i - \bar{X}).$$

(2) Subtract $\bar{Y}$ from both sides of the above equation to get

$$(\hat{Y}_i - \bar{Y}) = \hat{\beta}_1 (X_i - \bar{X}),$$

which is simply

$$\hat{y}_i = \hat{\beta}_1 x_i$$

where by definition $\hat{y}_i \equiv \hat{Y}_i - \bar{Y}$ and $x_i \equiv X_i - \bar{X}$. □

<u>PROPERTY 3</u>: The ***sample mean*** **of the OLS residuals** $\hat{u}_i$ **equals** ***zero***, or the ***sum*** **of the OLS residuals** $\hat{u}_i$ **equals** ***zero***.

$$\bar{u} = \sum_{i=1}^{N}\hat{u}_i \Big/ N = \frac{\sum_{i=1}^{N}\hat{u}_i}{N} = 0 \quad \textit{or} \quad \sum_{i=1}^{N}\hat{u}_i = 0. \qquad \text{... (C3)}$$

- **<u>Proof of (C3):</u>** Involves demonstrating that $\Sigma_i \hat{u}_i = 0$.

  (1) From the first normal equation (N1), we have

$$-2\sum_{i=1}^{N}\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right) = 0$$

$$\sum_{i=1}^{N}\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right) = 0.$$

  (2) But $\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right) = \hat{u}_i$ by definition, so that equation (N1) implies that

$$\sum_{i=1}^{N}\hat{u}_i = 0 \quad \text{and hence that} \quad \bar{u} = \sum_{i=1}^{N}\hat{u}_i / N = \frac{\sum_i \hat{u}_i}{N} = 0. \quad \square$$

<u>NOTE</u>: **Properties 1-3** depend on their being **an** ***intercept coefficient*** **in the population regression function**. The following two properties do not require an intercept in the regression function.

<u>PROPERTY 4</u>: The **OLS residuals** $\hat{\mathbf{u}}_{\mathbf{i}}$ **are** ***uncorrelated with*** **the** ***sample*** **values of X, the X$_i$**; i.e.,

$$\sum_{i=1}^{N} X_i \hat{u}_i = 0 . \qquad \text{... (C4)}$$

- **<u>Proof of (C4)</u>:** Is based on the second OLS normal equation (N2):

$$\sum_{i=1}^{N} X_i Y_i = \hat{\beta}_0 \sum_{i=1}^{N} X_i + \hat{\beta}_1 \sum_{i=1}^{N} X_i^2 . \qquad \text{(N2)}$$

(1) Since $\hat{u}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$, we can pre-multiply by X_i to obtain

$$X_i \hat{u}_i = X_i Y_i - \hat{\beta}_0 X_i - \hat{\beta}_1 X_i^2 .$$

(2) Summing over i = 1,...,N, we get

$$\sum_{i=1}^{N} X_i \hat{u}_i = \sum_{i=1}^{N} X_i Y_i - \hat{\beta}_0 \sum_{i=1}^{N} X_i - \hat{\beta}_1 \sum_{i=1}^{N} X_i^2$$

$$= 0 \qquad \text{by normal equation (N2).}$$

<u>PROPERTY 5</u>: The **OLS residuals** $\hat{u}_i$ **are** ***uncorrelated with*** **the** ***estimated*** **or** ***predicted*** **values of** Y_i**, the** $\hat{Y}_i$; i.e.,

$$\sum_{i=1}^{N} \hat{Y}_i \hat{u}_i = 0 . \qquad \text{... (C5)}$$

- **<u>Proof of (C5)</u>:** Makes use of properties (C3) and (C4) above.

  (1) Since $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$, we can post-multiply by $\hat{u}_i$ to obtain

  $$\hat{Y}_i \hat{u}_i = \hat{\beta}_0 \hat{u}_i + \hat{\beta}_1 X_i \hat{u}_i .$$

  (2) Summing over i = 1, ..., N, we get

  $$\sum_{i=1}^{N} \hat{Y}_i \hat{u}_i = \hat{\beta}_0 \sum_{i=1}^{N} \hat{u}_i + \hat{\beta}_1 \sum_{i=1}^{N} X_i \hat{u}_i .$$

  (3) But $\sum_{i=1}^{N} \hat{u}_i = 0$ by (C3) and $\sum_{i=1}^{N} X_i \hat{u}_i = 0$ by (C4), so that

  $$\sum_{i=1}^{N} \hat{Y}_i \hat{u}_i = 0 .$$

2. Goodness-of-Fit of the OLS-SRE: Objective

The previous section derived the **computational properties of the OLS sample regression equation (OLS-SRE)**.

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i \qquad (i = 1, \ldots, N) \qquad \textbf{(1)}$$

where

$\hat{\beta}_0$ = the OLS estimator of the intercept coefficient β_0,

$\hat{\beta}_1$ = the OLS estimator of the slope coefficient β_1,

$\hat{u}_i$ = the OLS residual for sample observation i.

Our objective now is to derive a measure of how well the OLS-SRE fits the sample data.

- The measure of goodness-of-fit we use is called the ***coefficient of determination***, which is conventionally **denoted as R^2**.

- The R^2 provides a measure of how well the OLS-SRE explains, or accounts for, the observed sample variation of the regressand Y, where

$$\textbf{sample variation of Y} \equiv \sum_{i=1}^{N} (Y_i - \bar{Y})^2 = \sum_{i=1}^{N} y_i^2 .$$

- The derivation of the R^2 for an OLS-SRE is based on the **OLS decomposition equation** for the sample variation of Y.

3. The OLS Decomposition Equation

3.1 Derivation of the OLS Decomposition Equation

1. For each sample observation i, the OLS-SRE is written as

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i$$

or (1)

$$Y_i = \hat{Y}_i + \hat{u}_i \qquad \text{where } \hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i \qquad (i = 1, \ldots, N)$$

2. Subtract the sample mean of the Y_i values, $\bar{Y}$, from both sides of equation (1):

$$Y_i - \bar{Y} = \hat{Y}_i - \bar{Y} + \hat{u}_i$$

or, in deviation-from-means form,

$$y_i = \hat{y}_i + \hat{u}_i \tag{2}$$

where $y_i \equiv Y_i - \bar{Y}$, $\hat{y}_i \equiv \hat{Y}_i - \bar{Y} = \hat{\beta}_1 x_i$, and $x_i \equiv X_i - \bar{X}$.

3. Next, square both sides of equation (2):

$$\begin{aligned} y_i^2 &= (\hat{y}_i + \hat{u}_i)^2 \\ &= \hat{y}_i^2 + \hat{u}_i^2 + 2\hat{y}_i\hat{u}_i \end{aligned} \tag{3}$$

4. Now sum both sides of equation (3) over i = 1,...,N:

$$\sum_{i=1}^{N} y_i^2 = \sum_{i=1}^{N} \hat{y}_i^2 + \sum_{i=1}^{N} \hat{u}_i^2 + 2\sum_{i=1}^{N} \hat{y}_i\hat{u}_i \tag{4}$$

5. But the last term on the right-hand side of equation (4) equals zero:

$$\sum_{i=1}^{N}\hat{y}_i\hat{u}_i = \sum_{i=1}^{N}(\hat{Y}_i - \overline{Y})\hat{u}_i$$
$$= \sum_{i=1}^{N}\hat{Y}_i\hat{u}_i - \overline{Y}\sum_{i=1}^{N}\hat{u}_i$$
$$= 0 \qquad \text{since } \sum_{i=1}^{N}\hat{Y}_i\hat{u}_i = 0 \text{ by (C5) and } \sum_{i=1}^{N}\hat{u}_i = 0 \text{ by (C3).}$$

6. Therefore, setting $\sum_i \hat{y}_i\hat{u}_i = 0$ in equation (4) gives the result that

$$\sum_{i=1}^{N}y_i^2 = \sum_{i=1}^{N}\hat{y}_i^2 + \sum_{i=1}^{N}\hat{u}_i^2 . \qquad \textbf{(5)}$$

❑ ***<u>Result</u>:*** **Equation (5) is the OLS decomposition equation for OLS-SRE (1).**

3.2 Interpretation of the OLS Decomposition Equation

Equation (5) is the OLS decomposition equation for OLS-SRE (1):

$$\sum_{i=1}^{N} \mathbf{y}_i^2 = \sum_{i=1}^{N} \hat{\mathbf{y}}_i^2 + \sum_{i=1}^{N} \hat{\mathbf{u}}_i^2 . \qquad \textbf{(5)}$$

Each of the three terms in equation (5) are defined as follows:

(1) $\sum_{i=1}^{N} y_i^2 \equiv$ TSS $\equiv$ the Total Sum of Squares

= the total sum of squares of the observed sample values of Y about their sample mean $\bar{Y}$

= the total sample variation of the observed Y_i values.

(2) $\sum_{i=1}^{N} \hat{y}_i^2 \equiv$ ESS $\equiv$ the Explained Sum of Squares

= the sum of squares of the estimated or predicted values of Y, the $\hat{Y}_i$, about their sample mean $\bar{Y}$

= the sum of squares explained by the sample regression function, i.e., by the regressor X.

(3) $\sum_{i=1}^{N} \hat{u}_i^2 \equiv$ RSS $\equiv$ the Residual Sum of Squares

= the sum of squares of the OLS residuals $\hat{u}_i$

= the unexplained variation of the observed sample values Y_i of the regressand Y around the sample regression line

Using these definitions, the OLS decomposition equation

$$\sum_{i=1}^{N} y_i^2 = \sum_{i=1}^{N} \hat{y}_i^2 + \sum_{i=1}^{N} \hat{u}_i^2 . \qquad (5.1)$$

can be re-written as

$$\text{TSS} = \text{ESS} + \text{RSS} \qquad (5.2)$$

Equation (5.1) or (5.2) -- the OLS decomposition equation -- decomposes the **sample variation of the regressand Y** into ***two*** **additive components**:

(1) one component, $\text{ESS} \equiv \sum_{i=1}^{N} \hat{y}_i^2$, is attributable to, or explained by, the sample regression function $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$;

(2) a second component, $\text{RSS} \equiv \sum_{i=1}^{N} \hat{u}_i^2$, is attributable to the OLS residuals $\hat{u}_i$ representing unknown random factors that influence the observed Y_i values.

3.3 An Unbiased Estimator of the Error Variance

The Residual Sum of Squares (RSS) in the OLS decomposition equation can be used to construct an ***unbiased*** **estimator** of the ***unknown*** **error variance** σ^2.

- ***Question:*** Why do we need an estimator of the error variance σ^2?

- ***Answer:*** We need an estimator of the error variance σ^2 so that we can obtain estimators of the variances of the OLS coefficient estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ which as we have seen are given by the formulas

$$\text{Var}(\hat{\beta}_1) = \frac{\sigma^2}{\sum_i x_i^2} = \frac{\sigma^2}{\sum_i (X_i - \bar{X})^2}$$

$$\text{Var}(\hat{\beta}_0) = \frac{\sigma^2 \sum_i X_i^2}{N \sum_i x_i^2} = \frac{\sigma^2 \sum_i X_i^2}{N \sum_i (X_i - \bar{X})^2}.$$

❑ ***<u>Result</u>:*** An ***unbiased*** **estimator** of the **error variance** $\boldsymbol{\sigma^2}$ is given by the formula

$$\hat{\sigma}^2 = \frac{\sum_i \hat{u}_i^2}{(N-2)} = \frac{\text{RSS}}{(N-2)}, \qquad \hat{u}_i = Y_i - \hat{Y}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i \quad (i = 1, ..., N)$$

where 2 is the number of regression coefficients estimated, and N–2 is the degrees of freedom for RSS.

- **Explanation:** $\hat{\sigma}^2$ is an unbiased estimator of the error variance because it can be shown that

$$E(\text{RSS}) = E\left(\Sigma_i \hat{u}_i^2\right) = (N-2)\sigma^2.$$

Therefore

$$E(\hat{\sigma}^2) = E\left(\frac{\sum_i \hat{u}_i^2}{(N-2)}\right) = \frac{E\left(\sum_i \hat{u}_i^2\right)}{(N-2)} = \frac{(N-2)\sigma^2}{(N-2)} = \sigma^2.$$

- ***Summary:*** $\hat{\sigma}^2$ is an ***unbiased*** **estimator** of the error variance σ^2:

$$E(\hat{\sigma}^2) = \sigma^2 \quad \text{because} \quad E(RSS) = E\left(\Sigma_i \hat{u}_i^2\right) = (N-2)\sigma^2.$$

3.4 Unbiased Estimators of the Variances of the OLS Coefficient Estimates

Unbiased estimators of $Var(\hat{\beta}_1)$ and $Var(\hat{\beta}_0)$ are obtained by simply replacing the unknown σ^2 with its unbiased estimator $\hat{\sigma}^2$ in the formulas for $Var(\hat{\beta}_1)$ and $Var(\hat{\beta}_0)$.

- The **unbiased estimator of** $\mathbf{Var(\hat{\beta}_1)}$, the variance of $\hat{\beta}_1$, is therefore

$$\hat{Var}(\hat{\beta}_1) = \frac{\hat{\sigma}^2}{\sum_i x_i^2} = \frac{\hat{\sigma}^2}{\sum_i (X_i - \bar{X})^2}.$$

- The **unbiased estimator of** $\mathbf{Var(\hat{\beta}_0)}$, the variance of $\hat{\beta}_0$, is therefore

$$\hat{Var}(\hat{\beta}_0) = \frac{\hat{\sigma}^2 \sum_i X_i^2}{N \sum_i x_i^2} = \frac{\hat{\sigma}^2 \sum_i X_i^2}{N \sum_i (X_i - \bar{X})^2}.$$

4. The Coefficient of Determination, R^2

4.1 Definition of R^2

1. Start with the OLS decomposition equation (5.1) or (5.2):

$$\sum_{i=1}^{N} y_i^2 = \sum_{i=1}^{N} \hat{y}_i^2 + \sum_{i=1}^{N} \hat{u}_i^2 \qquad (5.1)$$

$$TSS = ESS + RSS \qquad (5.2)$$

2. Divide both sides of the OLS decomposition equation (5.1) or (5.2) by TSS = $\sum_{i=1}^{N} y_i^2$:

$$1 = \frac{\sum_i \hat{y}_i^2}{\sum_i y_i^2} + \frac{\sum_i \hat{u}_i^2}{\sum_i y_i^2} \qquad (6.1)$$

or

$$1 = \frac{ESS}{TSS} + \frac{RSS}{TSS} \qquad (6.2)$$

3. The **coefficient of determination R^2** is defined as:

$$R^2 \equiv \frac{\sum_i \hat{y}_i^2}{\sum_i y_i^2} = 1 - \frac{\sum_i \hat{u}_i^2}{\sum_i y_i^2} \qquad \text{from equation (6.1)}$$

or

$$R^2 \equiv \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \qquad \text{from equation (6.2)}$$

4.2 Interpretation of R^2: The Values of R^2

- **<u>What Does R^2 Measure?</u>**

R^2 = the proportion of the total sample variation of the dependent variable Y that is explained by the sample regression function, i.e., by the values of the regressor X.

- **<u>The Values of R^2</u>**

R^2 values lie in the closed unit interval [0, 1]; i.e., $\mathbf{0 \le R^2 \le 1}$.

- **<u>Interpreting the Values of R^2</u>**

- ***<u>Rule 1</u>*****:** The ***closer*** **is the value of R^2 to 1**, the ***better*** **the goodness-of-fit** of the OLS-SRE to the sample data.

  - The upper limiting value $R^2 = 1$ corresponds to a **perfect fit** of the OLS-SRE to the sample data.

$$R^2 = 1 \quad \Rightarrow \quad \frac{ESS}{TSS} = 1 \quad \Rightarrow \quad ESS = TSS \quad \Rightarrow \quad RSS = \sum_i \hat{u}_i^2 = 0.$$

  - But since $\hat{u}_i^2 \ge 0$ for all i, $RSS = \sum_i \hat{u}_i^2 = 0$ if and only if

$$\hat{u}_i = 0 \quad \forall \; i = 1,...,N.$$

  - Therefore, a perfect fit of the OLS-SRE means that

$$\hat{u}_i = 0 \quad \forall \; i = 1,...,N \qquad \textit{or} \qquad Y_i = \hat{Y}_i \quad \forall \; i = 1,...,N.$$

- ***Rule 2*****:** The ***closer*** **is the value of R^2 to 0**, the ***worse*** **the goodness-of-fit** of the OLS-SRE to the sample data.

  - The lower limiting value $R^2 = 0$ corresponds to the **worst possible fit** of the OLS-SRE to the sample data.

$$R^2 = 0 \quad \Rightarrow \quad \frac{ESS}{TSS} = 0 \quad \Rightarrow \quad ESS = 0 \quad \Rightarrow \quad TSS = RSS.$$

  - But ESS = 0 if and only if $\hat{\beta}_1 = 0$:

$$ESS = 0 \quad \Rightarrow \quad \sum_{i=1}^{N} \hat{y}_i^2 = \hat{\beta}_1^2 \sum_{i=1}^{N} x_i^2 = 0 \qquad \Rightarrow \qquad \hat{\beta}_1 = 0.$$

  - Finally, since $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ and $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$, it follows that $R^2 = 0$ means that

$$\hat{Y}_i = \hat{\beta}_0 = \bar{Y} \qquad \forall\ i = 1,...,N.$$

  The reason is that

$$\begin{aligned} \hat{\beta}_1 = 0 \quad & \Rightarrow \quad \hat{Y}_i = \hat{\beta}_0 & \text{since } \hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i \\ & \Rightarrow \quad \hat{Y}_i = \hat{\beta}_0 = \bar{Y} & \text{since } \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}. \end{aligned}$$