M339G: March 8th, 2024.

Exam Aftermath.

Problem 1.5.

Consider a simple linear regression fitted on 20 observations. In our usual notation, you are given:

(i) $\boxed{\sum (y_i - \hat{y}_i)^2 = 10}$ ←

(ii) $\sum (\hat{y}_i - \bar{y})^2 = 112$ ←

Find the coefficient of determination.

→: We know:

$$\boxed{TSS = \sum (y_i - \bar{y})^2 = \underbrace{\sum (y_i - \hat{y}_i)^2}_{RSS} + \boxed{\sum (\hat{y}_i - \bar{y})^2}}$$

$$\sum (y_i - \bar{y})^2 = \sum (y_i - \hat{y}_i + \hat{y}_i - \bar{y})^2$$

$$= \sum (y_i - \hat{y}_i)^2 + \boxed{2 \sum (y_i - \hat{y}_i)(\hat{y}_i - \bar{y})} + \sum (\hat{y}_i - \bar{y})^2$$

Task. Show that

$$\sum \underbrace{(y_i - \hat{y}_i)}_{\varepsilon_i} \underline{(\hat{y}_i - \bar{y})} \overset{?}{=} 0$$

$$\Leftrightarrow$$

$$\sum \varepsilon_i \left(\cancel{\beta_0} + \beta_1 \cdot x_i - (\cancel{\beta_0} + \beta_1 \cdot \bar{x}) \right) \overset{?}{=} 0$$

$$\Leftrightarrow$$

$$\sum \varepsilon_i \beta_1 (x_i - \bar{x}) \overset{?}{=} 0$$

$$\Leftrightarrow$$

$$\beta_1 \sum \varepsilon_i (x_i - \bar{x}) \overset{?}{=} 0$$

$$\beta_1 \left(\underbrace{\sum \varepsilon_i x_i}_{=0} - \bar{x} \underbrace{\sum \varepsilon_i}_{=0} \right) \overset{?}{=} 0$$

$$\sum_i (y_i - \hat{y}_i)^2 \longrightarrow \min$$

$$\sum_i (y_i - \beta_0 - \beta_1 x_i)^2 \longrightarrow \min$$

$$\frac{\partial}{\partial \beta_0} \sum_i (y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

$$\sum_i 2(y_i - \beta_0 - \beta_1 x_i)(-1) = 0$$

$$\frac{\partial}{\partial \beta_1} \sum (y_i - \beta_0 \underbrace{- \beta_1 x_i}) ^2 = 0$$

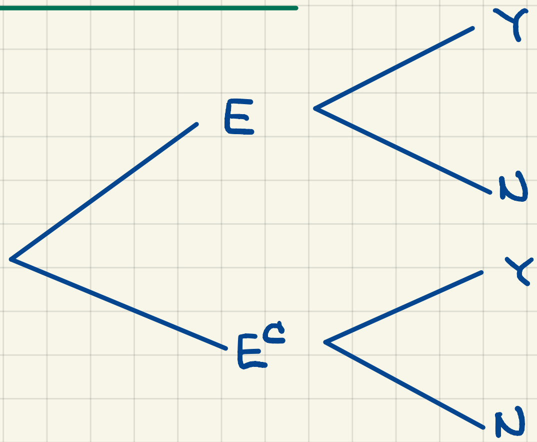
$$\sum_i 2\underbrace{(y_i - \beta_0 - \beta_1 x_i)}(-x_i) = 0$$

$$\sum_i \varepsilon_i x_i = 0$$

By def'n:

$$\boxed{R^2 = \frac{TSS - RSS}{TSS} = \frac{112}{10+112} = \frac{112}{122}}$$

Bayes Theorem.



In m362k:

$$\mathbb{P}[E \mid Y] = ?$$