HW 3

Gehua ZHANG (UNI: gz2280)

---

* 2.55.

Derive: $SSR = b_1^2 \cdot \sum (x_i - \bar{x})^2$

$$SSR = \sum (\hat{Y}_i - \bar{Y})^2$$

$$= \sum \left(\hat{Y}_i - \frac{\sum \hat{Y}_i}{n}\right)^2$$

$$= \sum \left[(b_0 + b_1 x_i) - \frac{\sum (b_0 + b_1 x_i)}{n}\right]^2$$

$$= \sum \left[b_0 + b_1 x_i - \left(b_0 + b_1 \cdot \frac{\sum x_i}{n}\right)\right]^2$$

$$= \sum (b_1 x_i - b_1 \cdot \bar{x})^2$$

$$= b_1^2 \cdot \sum (x_i - \bar{x})^2 \quad \blacksquare$$

* 2.56

a. $X_i = \{1, 4, 10, 11, 14\}$, $\bar{x} = 8$

$\hat{Y}_i = \{8, 17, 35, 36, 47\}$

$$E(MSE) = E\left(\frac{\sum (Y_i - \hat{Y}_i)^2}{n-2}\right) = \sigma^2 = 0.36$$

$$E(MSR) = \sigma^2 + \beta_1^2 \sum (x_i - \bar{x})^2$$

$$= 0.36 + 9(7^2 + 4^2 + 2^2 + 3^2 + 6^2)$$

$$= 1026.36$$

b. $X_i = \{6, 7, 8, 9, 10)$

$E(MSR) = 0.36 + 9(4+1+0+1+4) = 90.36$

E(MSR) decreses From 1026.36 to 90.36

If we have a dataset where X_i is more spread out than a centered sampling of X_i, it is more likely that we have a smaller variance of b_1:

$$\sigma^2\{b_1\} = \frac{\sigma^2}{\sum (X_i - \bar{x})^2}$$

Since $\sigma^2\{b_1\}$ decreases as X_i Spread out, The Confidence Interval of b_1 would be more narrow, thus I think we would have more Confidence to estimate if there's Any Linear Relation exists.

To estimate mean Response, Howere, I think both of them are the same.

$E(\hat{Y}_i) = \beta_0 + \beta_1 \cdot X_i$

* 2.57

a. The Reduced model:

$$Y_i = 5 \cdot X_i + \varepsilon_i$$

degree of Freedom $= n-1$, Since we Restricted b_1, and need to estimate b_0.

b. The Reduced Model

$$Y_i = 5 \cdot X_i + 2 + \varepsilon_i$$

degree of Freedom $= n$.

We have both parameters Restricted.