

MSE Unbiased estimator

Remind b_1 is linear combination of y_i , it is $b_1 = \sum_{i=1}^n c_i y_i$, where $c_i = (x_i - \bar{x}) / (\sum_{i=1}^n (x_i - \bar{x})^2)$, then

$$\begin{aligned} b_1 &= \sum_{i=1}^n c_i y_i = \sum_{i=1}^n c_i (\beta_0 + \beta_1 x_i + \varepsilon_i) \\ &= \sum_{i=1}^n c_i \beta_0 + \sum_{i=1}^n c_i \beta_1 x_i + \sum_{i=1}^n c_i \varepsilon_i \\ &= \beta_0 \sum_{i=1}^n c_i + \beta_1 \sum_{i=1}^n c_i x_i + \sum_{i=1}^n c_i \varepsilon_i \\ &= \beta_1 + \sum_{i=1}^n c_i \varepsilon_i \end{aligned}$$

above is because $\sum_{i=1}^n c_i = 0$

$$\sum_{i=1}^n \frac{(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{n\bar{x} - n\bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2} = 0$$

and $\sum_{i=1}^n c_i x_i = 1$.

$$\sum_{i=1}^n \frac{(x_i - \bar{x})x_i}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})x_i}{\sum_{i=1}^n (x_i - \bar{x})x_i - \sum_{i=1}^n (x_i - \bar{x})\bar{x}} = \frac{\sum_{i=1}^n (x_i - \bar{x})x_i}{\sum_{i=1}^n (x_i - \bar{x})x_i - (n\bar{x} - n\bar{x})\bar{x}} = 1$$

Then expect value,

$$\begin{aligned}
& \mathbb{E} \left[-2(b_1 - \beta_1) \sum_{i=1}^n (x_i - \bar{x})(\varepsilon_i - \bar{\varepsilon}) \right] \\
&= \mathbb{E} \left[-2(\beta_1 + \sum_{i=1}^n c_i \varepsilon_i - \beta_1) \sum_{i=1}^n (x_i - \bar{x})(\varepsilon_i - \bar{\varepsilon}) \right] \\
&= -2\mathbb{E} \left[\sum_{i=1}^n c_i \varepsilon_i \sum_{i=1}^n (x_i - \bar{x}) \varepsilon_i \right] \\
&= -2\mathbb{E} \left[\frac{(\sum_{i=1}^n (x_i - \bar{x}) \varepsilon_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] \\
&= \frac{-2}{\sum_{i=1}^n (x_i - \bar{x})^2} \mathbb{E} \left[\left((x_1 - \bar{x}) \varepsilon_1 + \sum_{i=2}^n (x_i - \bar{x}) \varepsilon_i \right)^2 \right] \\
&= \frac{-2}{\sum_{i=1}^n (x_i - \bar{x})^2} \mathbb{E} \left[(x_1 - \bar{x})^2 \varepsilon_1^2 + \left(\sum_{i=2}^n (x_i - \bar{x}) \varepsilon_i \right)^2 + 2(x_1 - \bar{x}) \left(\sum_{i=2}^n \varepsilon_i (x_i - \bar{x}) \right) \right] \\
&= \frac{-2}{\sum_{i=1}^n (x_i - \bar{x})^2} \mathbb{E} \left[(x_1 - \bar{x})^2 \varepsilon_1^2 \right] + \mathbb{E} \left[\left(\sum_{i=2}^n (x_i - \bar{x}) \varepsilon_i \right)^2 \right] + \mathbb{E} \left[2(x_1 - \bar{x}) \left(\sum_{i=2}^n \varepsilon_i (x_i - \bar{x}) \right) \right] \\
&= \frac{-2}{\sum_{i=1}^n (x_i - \bar{x})^2} \left((x_1 - \bar{x})^2 \mathbb{E} [\varepsilon_1^2] + \mathbb{E} \left[\left(\sum_{i=2}^n (x_i - \bar{x}) \varepsilon_i \right)^2 \right] + 2(x_1 - \bar{x}) \left(\sum_{i=2}^n (x_i - \bar{x}) \mathbb{E} [\varepsilon_i] \right) \right) \\
&= \frac{-2}{\sum_{i=1}^n (x_i - \bar{x})^2} \left((x_1 - \bar{x})^2 \mathbb{E} [\varepsilon_1^2] + \mathbb{E} \left[\left(\sum_{i=2}^n (x_i - \bar{x}) \varepsilon_i \right)^2 \right] \right) \\
&\dots \\
&= \frac{-2}{\sum_{i=1}^n (x_i - \bar{x})^2} \left((x_1 - \bar{x})^2 \mathbb{E} [\varepsilon_1^2] + (x_2 - \bar{x})^2 \mathbb{E} [\varepsilon_2^2] + \dots + (x_n - \bar{x})^2 \mathbb{E} [\varepsilon_n^2] \right) \\
&= \frac{-2}{\sum_{i=1}^n (x_i - \bar{x})^2} \left((x_1 - \bar{x})^2 \sigma^2 + (x_2 - \bar{x})^2 \sigma^2 + \dots + (x_n - \bar{x})^2 \sigma^2 \right) \\
&= \frac{-2}{\sum_{i=1}^n (x_i - \bar{x})^2} \left((x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2 \right) \sigma^2 \\
&= \frac{-2 \sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \sigma^2 = -2\sigma^2
\end{aligned}$$