Desenvolvimento

Slide 19

Modelo: $y_i = b_0 + b_1 x_i + \epsilon$, onde $\epsilon \sim N(0, \sigma^2)$ e $y_i \sim N(b_0 + b_1 x_i, \sigma^2)$.

Mostre que o estimador $\mathrm{MSE} = \frac{\mathrm{SQE}}{n-2}$ é não viciado para σ^2.

$$E[\mathrm{MSE}] = E\left[\frac{\mathrm{SQE}}{n-2}\right] = \frac{1}{n-2}E[\mathrm{SQE}] = \frac{1}{n-2}E\left[\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\right]$$
$$= \frac{1}{n-2}E\left[\sum_{i=1}^{n}\left(y_i - \hat{b}_0 - \hat{b}_1 x_i\right)^2\right]$$

Como $\hat{b}_0 = \bar{y} - \hat{b}_1\bar{x}$

$$E[\mathrm{MSE}] = \frac{1}{n-2}E\left[\sum_{i=1}^{n}\left(y_i - \bar{y} + \hat{b}_1\bar{x} - \hat{b}_1 x_i\right)^2\right] = \frac{1}{n-2}E\left\{\sum_{i=1}^{n}\left[(y_i - \bar{y}) - \hat{b}_1(x_i - \bar{x})\right]^2\right\}$$
$$= \frac{1}{n-2}E\left\{\sum_{i=1}^{n}\left[(y_i - \bar{y})^2 - 2\hat{b}_1(y_i - \bar{y})(x_i - \bar{x}) + \hat{b}_1^2(x_i - \bar{x})^2\right]\right\}$$
$$= \frac{1}{n-2}E\left[\sum_{i=1}^{n}(y_i - \bar{y})^2 - 2\hat{b}_1\underbrace{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}_{\delta} + \hat{b}_1^2\sum_{i=1}^{n}(x_i - \bar{x})^2\right]$$

Como

$$\hat{b}_1 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \text{ [Mínimos quadrados]}$$

então

$$\delta = \hat{b}_1\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Portanto

$$E[\mathrm{MSE}] = \frac{1}{n-2}E\left[\sum_{i=1}^{n}(y_i - \bar{y})^2 - 2\hat{b}_1^2\sum_{i=1}^{n}(x_i - \bar{x})^2 + \hat{b}_1^2\sum_{i=1}^{n}(x_i - \bar{x})^2\right]$$
$$= \frac{1}{n-2}E\left[\sum_{i=1}^{n}(y_i - \bar{y})^2 - \hat{b}_1^2\underbrace{\sum_{i=1}^{n}(x_i - \bar{x})^2}_{S_{xx}}\right]$$

Obs.:

$$\hat{b}_1 = \frac{S_{xy}}{S_{xx}}$$

Continuando

$$E[\text{MSE}] = \frac{1}{n-2}\left\{\sum_{i=1}^{n} E[(y_i - \bar{y})^2] - S_{xx}\sum_{i=1}^{n} E\left[\hat{b}_1^2\right]\right\} = \frac{1}{n-2}\left\{\sum_{i=1}^{n} E[(y_i - \bar{y})^2] - S_{xx}E\left[\hat{b}_1^2\right]\right\}$$
$$= \frac{1}{n-2}\left\{\sum_{i=1}^{n} E[(y_i - \bar{y})^2] - S_{xx}\left[Var\left(\hat{b}_1\right) + E^2\left(\hat{b}_1\right)\right]\right\}$$

Temos que

$$E^2\left(\hat{b}_1\right) = b_1^2, \qquad Var\left(\hat{b}_1\right) = \frac{\sigma^2}{S_{xx}}$$

Logo

$$E[\text{MSE}] = \frac{1}{n-2}\left\{\sum_{i=1}^{n} \underbrace{E[(y_i - \bar{y})^2]}_{\xi} - S_{xx}\left[\frac{\sigma^2}{S_{xx}} + b_1^2\right]\right\}$$

Em ξ, temos que

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n}(b_0 + b_1 x_i + \epsilon_i) = \frac{1}{n}\sum_{i=1}^{n} b_0 + \frac{1}{n}\sum_{i=1}^{n} b_1 x_i + \frac{1}{n}\sum_{i=1}^{n} \epsilon_i = b_0 + b_1\bar{x} + \bar{\epsilon}$$

Portanto

$$\begin{aligned}\xi &= E[(b_0 + b_1 x_i + \epsilon_i - b_0 - b_1\bar{x} - \bar{\epsilon})^2] = E\{[b_1(x_i - \bar{x}) + (\epsilon_i - \bar{\epsilon})]^2\} \\ &= E[b_1^2(x_i - \bar{x}) + 2b_1(x_i - \bar{x})(\epsilon_i - \bar{\epsilon}) + (\epsilon_i - \bar{\epsilon})^2] \\ &= E[b_1^2(x_i - \bar{x})^2] + E[2b_1(x_i - \bar{x})(\epsilon_i - \bar{\epsilon})] + E[(\epsilon_i - \bar{\epsilon})^2] \\ &= b_1^2(x_i - \bar{x})^2 + \underbrace{2b_1(x_i - \bar{x})E[\epsilon_i - \bar{\epsilon}]}_{0} + E[(\epsilon_i - \bar{\epsilon})^2]\end{aligned}$$

Veja que

$$\begin{aligned}E[(\epsilon_i - \bar{\epsilon})^2] &= E\left[\epsilon_i^2 - 2\epsilon_1\bar{\epsilon} + \bar{\epsilon}^2\right] = E\left[\epsilon_i^2\right] - 2E[\epsilon_i\bar{\epsilon}] + E[\bar{\epsilon}^2] \\ &= \sigma^2 - 2E\left[\epsilon_i\frac{1}{n}\sum \epsilon_i\right] + E[\bar{\epsilon}^2] = \sigma^2 - 2E\left[\frac{1}{n}\epsilon_i^2\right] + E[\bar{\epsilon}^2] \\ &= \sigma^2 - \frac{2}{n}E\left[\epsilon_i^2\right] + E[\bar{\epsilon}^2] = \sigma^2 - \frac{2}{n}\sigma^2 + E[\bar{\epsilon}^2]\end{aligned}$$

Só nos resta lidar com

$$E[\bar{\epsilon}^2] = E\left[\left(\frac{1}{n}\sum \epsilon_i\right)^2\right] = E\left[\frac{1}{n^2}\sum_i\sum_j \epsilon_i\epsilon_j\right]$$

Como os erros são independentes, temos que

$$E\left[\epsilon_i\epsilon_j\right] = E[\epsilon_i]E\left[\epsilon_j\right] = 0, \qquad i \neq j$$

ao passo que

$$E\left[\epsilon_i^2\right] = \sigma^2$$

Logo

$$E[\bar{\epsilon}^2] = \frac{1}{n^2}\left[\sum_i E\left[\epsilon_i^2\right] + \sum_{i\neq j} E\left[\epsilon_i \epsilon_j\right]\right] = \frac{1}{n^2} n\sigma^2 = \frac{1}{n}\sigma^2$$

Em ξ, teremos

$$\xi = b_1^2(x_i - \bar{x})^2 + E[(\epsilon_i - \bar{\epsilon})^2] = b_1^2(x_i - \bar{x})^2 + \sigma^2 - \frac{1}{n}\sigma^2$$

que, no início

$$\begin{aligned} E[\mathrm{MSE}] &= \frac{1}{n-2}\left\{\sum_{i=1}^{n}\left[b_1^2(x_i - \bar{x})^2 + \sigma^2 - \frac{1}{n}\sigma^2\right] - S_{xx}\left[\frac{\sigma^2}{S_{xx}} + b_1^2\right]\right\} \\ &= \frac{1}{n-2}\{b_1^2 S_{xx} + n\sigma^2 - \sigma^2 - \sigma^2 - b_1^2 S_{xx}\} = \frac{1}{n-2}(n\sigma^2 - 2\sigma^2) \\ &= \frac{1}{n-2}(n-2)\sigma^2 = \sigma^2 \end{aligned}$$

Slide 26

Mostre que

$$t(\hat{\beta}) = \frac{\hat{\beta} - \beta}{S_e}\sqrt{\sum(x_i - \bar{x})^2} \sim t_{n-2}$$

Desenvolvimento:

Temos que, se $Z \sim N(0,1)$ e $Y \sim \mathcal{X}_v^2$, então

$$\frac{Z}{\sqrt{Y/v}} \sim t_v$$

Sabemos que

$$\hat{\beta} \sim N\left(\beta, \frac{\sigma_e^2}{\sum(x_i - \bar{x})^2}\right)$$

Portanto

$$t(\hat{\beta}) = \left(\frac{\hat{\beta} - \beta}{\sigma_e}\sqrt{\sum(x_i - \bar{x})^2}\right)\frac{\sigma_e}{S_e} = \frac{Z}{\frac{S_e}{\sigma_e}}$$

Veja que

$$\left(\frac{S_e}{\sigma_e}\right)^2 = \frac{S_e^2}{\sigma_e^2} = \frac{\sum(y_i - \hat{y}_i)^2}{\sigma_e^2}\frac{1}{n-2}$$

Entretanto

$$\sum(y - \hat{y})^2 = \sum\hat{\varepsilon}_i^2, \qquad \hat{\varepsilon}_i \sim N(0, \sigma_e^2) \text{ e } \frac{\hat{\varepsilon}_i}{\sigma_e} \sim N(0,1)$$

Logo

$$\frac{\sum\varepsilon_i^2}{\sigma_e^2} = \sum\frac{\varepsilon_i^2}{\sigma_e^2} = \sum\left(\frac{\varepsilon_i}{\sigma_e}\right)^2 = \sum Z^2 \sim \mathcal{X}_{n-2}^2$$

Finalmente

$$\frac{S_e^2}{\sigma_e^2} = \frac{Y}{n-2} \Rightarrow \frac{S_e}{\sigma_e} = \sqrt{\frac{Y}{n-2}}$$

tal que

$$t(\hat{\beta}) = \frac{Z}{\sqrt{\frac{Y}{n-2}}} \sim t_{n-2}$$