Math 564
Dhanvardh Voona [A20543395]
Homework 3

classmate
Date
Page

1) We know that

$$S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}).$$

To show S_{xy} can be also writen as

$$S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i)$$

Proof :-

$$S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

Expand product inside the summation

$$S_{xy} = \underbrace{\sum_{i=1}^{n} (x_i - \bar{x}) y_i}_{1} - \underbrace{\sum_{i=1}^{n} (x_i - \bar{x}) \bar{y}}_{2}$$

We know $\bar{y}$ is constant, so factoring out of summation

Part 2 will be $\bar{y} \underbrace{\sum_{i=1}^{n} (x_i - \bar{x})}_{3}$

lets expand part 3

We know $\sum_{i=1}^{n} x_i = n\bar{x}$

$$\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x} = n\bar{x} - n\bar{x} = 0.$$

$\textcircled{3} \Rightarrow \sum_{i=1}^{n} (x_i - \bar{x}) = 0 \Rightarrow \bar{y} \times 0 = 0$

$\Rightarrow$ Part ② is 0.

$$\therefore \boxed{S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x}) y_i} - \text{eq } ①$$

We know,

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$, and $c_i = \frac{x_i - \bar{x}}{S_{xx}}$

$S_{xy} = \left(\sum_{i=1}^{n} (x_i - \bar{x}) y_i\right.$ from eq ①

$$\hat{\beta}_1 = \sum_{i=1}^{n} \underbrace{\frac{(x_i - \bar{x})}{S_{xx}}}_{c_i} y_i$$

$$\boxed{\hat{\beta}_1 = \sum_{i=1}^{n} c_i y_i}$$

5A

$y = \beta_0 + \beta_1 x + \epsilon$ (True Relationship)

for a collected dataset, (x_i, y_i), the relationship can be expressed as

$$\boxed{y_i = \beta_0 + \beta_1 x_i + \epsilon_i}$$

(a) lets try to prove $E(\hat{\beta}_0) = \beta_0$,

~~In problem ① we pr~~
We know that $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

$$E[\hat{\beta}_0] = \underbrace{E[\bar{y}]}_{\text{Part ①}} - \bar{x}\, \underbrace{E[\hat{\beta}_1]}_{\text{Part ②}} \longrightarrow \text{Eq. ①}$$

Expand Part ① :-
finding expectation of $\bar{y}$

$\cancel{E[y] \; y_i = \beta_0 + \beta_1 x_i + \epsilon_i}$

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{\cancel{n}\beta_0}{\cancel{n}} + \frac{\cancel{n}\beta_1}{\cancel{n}} \times \bar{x} + \frac{1}{n}\sum_{i=1}^{n} \epsilon_i$$

$$\bar{y} = \beta_0 + \beta_1 \bar{X} + \frac{1}{n}\sum_{i=1}^{n} \epsilon_i$$

Lets ~~consider~~ compute Expectation along $\bar{y}$.

$$E[\bar{y}] = E[\beta_0] + E[\beta_1 \bar{x}] + E\left[\frac{1}{n}\sum_{i=1}^{n} \epsilon_i\right]$$

$\downarrow$ β_0 . $\qquad$ $\downarrow$ $\beta_1 \bar{X}$ $\qquad$ $\cancel{E(\epsilon_i)}$

lets consider $E[\epsilon_i] = 0$.

Then $\boxed{E[\bar{y}] = \beta_0 + \beta_1 \bar{x}}$ $\longrightarrow$ Equation ①

Expand part ② :-

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$\Rightarrow$ We proved in problem ① that

$$S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})\, y_i$$

$$\Rightarrow \hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i)}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(\beta_0 + \beta_1 x_i + \epsilon_i)}{S_{xx}}$$

$$= \frac{\sum_{i=1}^{n} (x_i - \bar{x})\beta_0}{S_{xx}} + \frac{\sum_{i=1}^{n} \cancel{(x_i - \bar{x})} \beta_1 x_i}{\cancel{S_{xx}}} + \frac{\sum_{i=1}^{n} (x_i - \bar{x})\epsilon_i}{S_{xx}}$$

$$0 \qquad + \boxed{\beta_1 + E\left[\frac{\sum_{i=1}^{n} (x_i - \bar{x})\epsilon_i}{S_{xx}}\right]}$$

0

Equation ②

Since we considered $E(\epsilon_i) = 0$

$\Rightarrow E[\hat{\beta}_1] = \beta_1$

Now Eq. (1) will be

$$E[\hat{\beta}_0] = E[\bar{y}] - E[\hat{\beta}_1]\bar{x}$$
$$= \beta_0 + \cancel{\beta_1 \bar{x}} - \cancel{\beta_1 \bar{x}}$$

$\boxed{E[\hat{\beta}_0] = \beta_0}$ when $E(\epsilon_i) = 0$

~~$E[\hat{\beta}_0]$~~ $\hat{\beta}_0$ is an unbiased estimator of β_0 when $E(\epsilon_i) = 0$

(b) ~~Var~~ To prove $Var(\hat{\beta}_1) = \dfrac{\sigma^2}{S_{xx}}$.

We know $\hat{\beta}_1 = \beta_1 + \dfrac{\sum_{i=1}^{n} (x_i - \bar{x})\epsilon_i}{S_{xx}}$

$$Var(\hat{\beta}_1) = \underbrace{Var(\beta_1)}_{0} + Var\left(\frac{\sum_{i=1}^{n} (x_i - \bar{x})\epsilon_i}{S_{xx}}\right)$$

Variance of constant is zero.

W.K.T $Var(aX) = a^2 Var(X)$

$$\Rightarrow Var(\hat{\beta}_1) = \frac{1}{S_{xx}^2} Var\left(\sum_{i=1}^{n} \underbrace{(x_i - \bar{x})}_{a} \underbrace{\epsilon_i}_{X}\right)$$

$\downarrow$

a^2

$$\Rightarrow Var(\hat{\beta}_1) = \frac{1}{S_{xx}^2} \sum_{i=1}^{n} (x_i - \bar{x})^2 \, Var(\epsilon_i)$$

Lets assume $Var(\epsilon_i) = \sigma^2$.

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}^{\cancel{2}}} \cancel{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\boxed{Var(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}} \text{ when } Var(\epsilon_i) = \sigma^2$$

(c) We know $\hat{\beta}_1 = \beta_1 + \sum_{i=1}^{n} \frac{(x_i - \bar{x})\epsilon_i}{S_{xx}}$

Lets assume $\epsilon_i \sim N(0, \sigma^2)$.

Then according to linear property of Normal distribution, the linear combination $\sum_{i=1}^{n} (x_i - \bar{x})\epsilon_i$ is also normal distribution

for $\hat{\beta}_1$, we found $E[\hat{\beta}_1] = \beta_1$

$$Var[\hat{\beta}_1] = \frac{\sigma^2}{S_{xx}}.$$

$\Rightarrow$ $\therefore$ We have:

$$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{S_{xx}}\right) \text{ when } \epsilon_i \sim N(0, \sigma^2)$$

2A. Lets consider $E[\epsilon_i] = E[\epsilon(x_i)] = \alpha$. $\xrightarrow{}$ non zero constant.

from eq (2) in Question (5) pt (a):-

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n}(x_i - \bar{x})\epsilon_i}{S_{xx}}$$

$$E[\hat{\beta}_1] = \underbrace{E[\beta_1]}_{\text{constant}} + E\left(\frac{\sum_{i=1}^{n}(x_i - \bar{x})\epsilon_i}{S_{xx}}\right)$$

$$= \beta_1 + \cancel{E} \frac{\sum_{i=1}^{n}(x_i - \bar{x})\underbrace{E[\epsilon_i]}_{\alpha}}{S_{xx}}$$

$= \cancel{\beta_1 +}$

$$= \beta_1 + \sum_{i=1}^{n}\frac{(x_i - \bar{x})\alpha}{S_{xx}}$$ $\rightarrow$ non zero constant, lets factor it out.

$$= \beta_1 + \alpha \underbrace{\frac{\sum_{i=1}^{n}(x_i - \bar{x})}{S_{xx}}}_{\rightarrow 0}$$

$\boxed{E[\hat{\beta}_1] = \beta_1}$ $\hat{\beta}_1$ is unbiased even when $E[\epsilon_i] = \alpha \neq 0$.

lets calculate $E[\hat{\beta}_0]$.

We know $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$.

$\bar{y} = \beta_0 + \beta_1 \bar{x} + \bar{\epsilon}$

$$\hat{\beta}_0 = \underbrace{\beta_0 + \beta_1 \bar{x}}_{\text{constant}} + \bar{\epsilon} - \hat{\beta}_1 \bar{x}$$

$$E[\hat{\beta}_0] = \beta_0 + \beta_1 \bar{x} + E(\bar{\epsilon}) - E[\hat{\beta}_1 \bar{x}]$$

We know
$E[\hat{\beta}_1] = \beta_1$

$$\Rightarrow E[\hat{\beta}_0] = \beta_0 + \cancel{\beta_1 \bar{x}} + E[\bar{\epsilon}] - \cancel{\beta_1 \bar{x}}$$

$$= \beta_0 + \underbrace{E[\bar{\epsilon}]}_{\text{Part (2)}}$$

- Lets expand Part (2)

$$E[\bar{\epsilon}] = E\left[\frac{1}{n}\sum_{i=1}^{n} \epsilon\right] = \frac{1}{n}\sum_{i=1}^{n} \underbrace{E[\epsilon_i]}_{\alpha}$$

$$= \frac{1}{\cancel{n}} \times \cancel{n}\,\alpha$$

$$= \alpha.$$

$$\Rightarrow \boxed{E[\hat{\beta}_0] = \beta_0 + \alpha.}$$

The expectation of $\hat{\beta}_0$ is biased estimator, which is biased by α.