Homework-3 A20545853 MATH564.

(1) we know that,

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

To show that,

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

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

where, $c_i = \frac{x_i - \bar{x}}{S_{xx}}$, $i = 1, \ldots, n$.

Now let's solve:

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

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

$$= \sum_{i=1}^{n} x_i y_i - \bar{x} \sum_{i=1}^{n} y_i - \bar{y} \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \bar{x}\bar{y}$$

($\because$ we know, $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$, $\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$)

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

$$= \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y} - \cancel{n\bar{x}\bar{y}} \cancel{+n\bar{x}\bar{y}}.$$

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

Now we can write, $n\bar{y} = \sum_{i=1}^{n} y_i$

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

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

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

Hence, shown.

we know that,

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

$$= \sum_{i=1}^{n} c_i y_i \quad \left(\text{where, } c_i = \frac{x_i - \bar{x}}{S_{xx}}\right) \quad i = 1 \ldots n$$

Hence shown//

(2) From the class, we know that

$\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased estimators of β_0

and β_1.

This means that

$E(\hat{\beta}_0) = \beta_0$ and $E(\hat{\beta}_1) = \hat{\beta}_1$

In that proof, we have assumed $E(\varepsilon_i) = E(\varepsilon_i, x_i) = 0$

what if $E(\varepsilon_i) = E(\varepsilon / x_i) = \alpha$.

Here α is a non-zero constant.

Find out, how would this assumption affect the expectations of $\hat{\beta}_0$ and $\hat{\beta}_1$?

from problem (1), we know.

$$\hat{\beta}_1 = \sum_{i=1}^{n} c_i y_i \quad \left(\text{where, } c_i = \frac{x_i - \bar{x}}{S_{xx}} \quad i = 1 \cdots n\right)$$

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

Now, lets solve for.

$$E(\hat{\beta}_1) = E\left(\sum_{i=1}^{n} c_i y_i\right) = \sum_{i=1}^{n} c_i E(y_i)$$

we know the true relation,

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

$$= \sum_{i=1}^{n} c_i E(\beta_0 + \beta_1 x_i + \varepsilon_i)$$

$$= \sum_{i=1}^{n} c_i E(\beta_0 + \beta_1 x_i) + \sum_{i=1}^{n} c_i E(\varepsilon_i)$$

from earlier, $E(\varepsilon_i) = \alpha$. non-zero constant.

$$= \beta_0 \sum_{i=1}^{n} c_i + \beta_1 \sum_{i=1}^{n} c_i x_i + \alpha \sum_{i=1}^{n} c_i .$$

We leave it as an exercise to show that.

$$\sum_{i=1}^{n} c_i = 0 \quad \& \quad \sum_{i=1}^{n} c_i x_i = 0 .$$

As we have done in the class.

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

$$= \frac{1}{S_{xx}} (n\bar{x} - n\bar{x}) = 0$$

similarly

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

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

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

So, $E(\hat{\beta}_1) = \beta_0 \sum_{i=1}^{n} c_i + \beta_1 \sum_{i=1}^{n} c_i x_i + \alpha \sum_{i=1}^{n} c_i$

$$= \beta_0 (0) + \beta_1 (1) + \alpha (0)$$

Therefore, $\boxed{E(\hat{\beta}_1) = \beta_1}$

So, the expectation of $\hat{\beta}_1$ remains β_1 even if $E(\varepsilon_i) = \alpha$,

Therefore $\hat{\beta}_1$ is still an unbiased estimator of β_1.

Now, $E(\hat{\beta}_0) = E(\bar{y} - \hat{\beta}_1 \bar{x})$

$$= E\left(\frac{\sum_{i=1}^{n} y_i}{n}\right) - \bar{x} E(\hat{\beta}_1)$$

$$= \frac{1}{n} \sum_{i=1}^{n} E(y_i) - \bar{x}\beta_1$$

as we know, $E(\hat{\beta}_1) = \beta_1$ from above

$$= \frac{1}{n} \sum_{i=1}^{n} E(\beta_0 + \beta_1 x_i + \varepsilon_i) - \bar{x}\beta_1$$

$$= \frac{1}{n}\left[\beta_0 \sum_{i=1}^{n} 1 + \beta_1 \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} E(\varepsilon_i)\right] - \bar{x}\beta_1$$

$$= \frac{1}{n}\left[n\beta_0 + \beta_1 n\bar{x} + n\alpha\right] - \bar{x}\beta_1$$

$= \beta_0 + \beta_1 \bar{x} + \alpha - \beta_1 \bar{x}$

$$\boxed{E(\hat{\beta}_0) = \beta_0 + \alpha}$$

Now Here, $\hat{\beta}_0$ is <u>biased</u> by α.

(5) Given that consider the true relation

$$y = \beta_0 + \beta_1 x + \varepsilon.$$

Data = $(x_1, y_1) \dots (x_n, y_n)$, can be expressed as.

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i.$$

To find,

* what assumptions about the error terms ε_i are necessary to obtain;

(a) <u>$E(\hat{\beta}_0) = \beta_0$.</u>

Let solve $E(\hat{\beta}_0)$.

we know, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

so, $E(\hat{\beta}_0) = E(\bar{y} - \hat{\beta}_1 \bar{x})$

$= E(\bar{y}) - E(\hat{\beta}_1 \bar{x})$

$$= E\left(\frac{\sum_{i=1}^{n} y_i}{n}\right) - \bar{x} E(\hat{\beta}_1)$$

we know that $\hat{\beta}_1$ is unbiased estimator of β_1 even after error term.

~~So,~~ So, $\cancel{\sum_{i}}$

$$= \frac{1}{n} \sum_{i=1}^{n} E(\beta_0 + \beta_1 x_i + \varepsilon_i) - \bar{x}\beta_1.$$

$$= \frac{1}{n}\left[\beta_0 \sum_{i=1}^{n} 1 + \beta_1 \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} E(\varepsilon_i)\right] - \bar{x}\beta_1$$

$$= \beta_0 + \beta_1 \bar{x} + E(\varepsilon_i) - \bar{x}\beta_1$$

$$E(\hat{\beta}_0) = \beta_0 + E(\varepsilon_i)$$

* for the given result to be to, we have to assume that $E(\varepsilon_i) = 0$.

(b) $\operatorname{var}(\hat{\beta}_1) = \dfrac{\sigma^2}{S_{xx}}$

we know, $\hat{\beta}_1 = \dfrac{S_{xy}}{S_{xx}}$.

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

$$\mathrm{var}(\hat{\beta}_1) = \mathrm{var}\left(\frac{\sum_{i=1}^{n}(x_i - \bar{x})\,y_i}{S_{xx}}\right)$$

$$= \frac{1}{S_{xx}^2}\left(\mathrm{var}\left[\sum_{i=1}^{n}(x_i - \bar{x})(y_i)\right)\right)$$

$$= \frac{1}{\cancel{S_{xx}^2}}\ \cancel{\mathrm{var}}$$

Can also be written as

$$\mathrm{var}(\hat{\beta}_1) = \mathrm{var}\left(\sum_{i=1}^{n} c_i y_i\right) \quad \text{where } c_i = \frac{x_i - \bar{x}}{S_{xx}}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j \,\mathrm{cov}(y_i, y_j)$$

Here assume $\mathrm{cov}(y_i, y_j) = 0$.

$$\sum_{i=1}^{n} c_i^2\,\mathrm{var}(y_i)$$

Since, ~~E~~ we have assume the same as above $E(\varepsilon_i) = 0$ so that

$\mathrm{var}(\varepsilon_i) = \sigma^2$.

and

Assume $\mathrm{cov}(y_i, y_j) = 0$, $i \neq j$.

(c) $\hat{\beta}_1 \sim N(\beta_1, \sigma^2/S_{xx})$

we have got the above resulte,

from the above we can say that

$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{S_{xx}}\right)$ ~~where~~ ~~$\varepsilon_i \sim$~~

* for the above result to be right,

we have to assume that, error term

$\varepsilon_i \sim N(0, \sigma^2)$ and $\mathrm{cov}(\varepsilon_i, \varepsilon_j) = 0$,

for $i \neq j$