Chanachon Hongsakul

6104640948

1. In Table 1. $X_i$ is total econometrics exam point (total points are 100) and $Y_i$ is GPA of each BE student.

**Table 1**

| Student | GPA $Y_i$ | TEST $X_i$ | $X_iY_i$ | $X_i^2$ | $X_i-\bar{X}=x_i$ | $x_i^2$ |
|---|---|---|---|---|---|---|
| 1 | 2.8 | 63 | 176.4 | 3,963 | -14.625 | 213.8906 |
| 2 | 3.4 | 72 | 244.8 | 5184 | -5.625 | 31.6406 |
| 3 | 3.0 | 78 | 234 | 6084 | 0.375 | 0.1406 |
| 4 | 3.5 | 81 | 283.5 | 6561 | 3.375 | 11.3906 |
| 5 | 3.6 | 87 | 313.2 | 7569 | 9.375 | 87.8906 |
| 6 | 3.0 | 75 | 225 | 5625 | -2.625 | 6.8906 |
| 7 | 2.7 | 75 | 202.5 | 5625 | -2.625 | 6.8906 |
| 8 | 3.7 | 90 | 333 | 8100 | 12.375 | 153.1406 |
| sum | 25.7 | 621 | 2012.4 | 48,717 | | 511.8748 |
| mean | 3.2125 | 77.625 | 251.55 | 6089.625 | | |

1.1 Now consider the two-variable model $Y_i = \beta_1 + \beta_2 X_i + u_i,\ u_i \sim NIID(0,\sigma^2)$

Use OLS to find the estimator of $\beta_1 and \beta_2$. Interpret the regression.

1.1 Find $\hat{\beta}_2$

$$\hat{\beta}_2 = \frac{n\sum X_iY_i - \sum X_i \sum Y_i}{n\sum X_i^2 - (\sum X_i)^2}$$

$$= \frac{8(2012.4)-(621)(25.7)}{8(48,717)-(621)^2}$$

$$= \frac{16099.2 - 15959.7}{389,736 - 385,641}$$

$$= \frac{139.5}{4099}$$

$$\hat{\beta}_2 = 0.0341$$

Find $\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2\bar{X}$

$$= 3.2125 - (0.0341)(77.625)$$

$$= 3.2125 - 2.6393$$

$$\hat{\beta}_1 = 0.5655$$

Therefore, $\hat{Y}_i = \beta_1 + \beta_2 X_i$ equation is $\hat{Y}_i = 0.5655 + 0.0341X_i$

If BE students got 0 in test exam $(X_i)$, their GPA $(Y_i)$ will likely to be 0.5655.

And every Econometric test score increase by 1, on average, GPA will rise by 0.0341.

1.2 Find $\hat{Y}_i$ $and$ $\hat{u}_i$ and show that $\sum_{i=1}^{n} \hat{u}_i \approx 0$

| $i$ | $X_i$ | $\hat{\beta}_1 - \hat{\beta}_2 X_i = \hat{Y}_i$ | $Y_i - \hat{Y}_i = \hat{u}_i$ |
|---|---|---|---|
| 1 | 63 | $0.5655 + 0.0341(63) = 2.7138$ | $2.8 - 2.7138 = 0.0862$ |
| 2 | 72 | $0.5655 + 0.0341(72) = 3.0207$ | $3.4 - 3.0207 = 0.3793$ |
| 3 | 78 | $0.5655 + 0.0341(78) = 3.2253$ | $3.0 - 3.2253 = -0.2253$ |
| 4 | 81 | $0.5655 + 0.0341(81) = 3.3276$ | $3.5 - 3.3276 = 0.1724$ |
| 5 | 87 | $0.5655 + 0.0341(87) = 3.5322$ | $3.6 - 3.5322 = 0.0678$ |
| 6 | 75 | $0.5655 + 0.0341(75) = 3.123$ | $3.0 - 3.123 = -0.123$ |
| 7 | 75 | $0.5655 + 0.0341(75) = 3.123$ | $2.7 - 3.123 = -0.423$ |
| 8 | 90 | $0.5655 + 0.0341(90) = 3.6345$ | $3.7 - 3.6345 = 0.0655$ |

$$\sum_{i=1}^{8} \hat{u}_i = 0.0862 + 0.3793 + (-0.2253) + 0.1724 + 0.0678 + (-0.123) + (-0.423) + 0.0655$$

$$= -0.0001 \approx 0$$

1.3 Find $var(\hat{u}_i), var(\hat{\beta}_1), and\ var(\hat{\beta}_2)$

since $var(u_i|X_1, X_2, \dots, X_n) = \sigma^2$ (Homoscedastic Variance) but $\sigma^2$ is unknown, so we need to find $\hat{\sigma}_u^2$ instead.

$$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-2} = \frac{(0.0862)^2 + (0.3793)^2 + (-0.2253)^2 + (0.1724)^2 + (0.0678)^2 + (-0.123)^2 + (-0.423)^2 + (0.0655)^2}{8-2}$$

$$= \frac{0.0074 + 0.1439 + 0.0508 + 0.0297 + 0.0046 + 0.0151 + 0.1789 + 0.0043}{6}$$

$$= \frac{0.4347}{6}$$

$$var(\hat{u}_i) = \hat{\sigma}^2 = 0.0725$$

$$var(\hat{\beta}_2) = \frac{\hat{\sigma}^2}{\sum x_i^2} = \frac{0.0725}{511.8748}$$

$$= 0.0001$$

$$var(\hat{\beta}_1) = \frac{\sum X_i^2}{n \sum x_i^2} \cdot \sigma^2 = \frac{48{,}717}{8(511.8748)} \cdot 0.0725$$

$$= \frac{3531.9825}{4094.9984} = 0.8625$$

2. Data is listed in the table

| $i$ | $X_i$ | $Y_i$ | $X_iY_i$ | $X_i^2$ | $X_i - \bar{X} = x_i$ | $x_i^2$ |
|---|---|---|---|---|---|---|
| 1 | 10 | 0 | 0 | 100 | −10 | 100 |
| 2 | 12 | 2 | 24 | 144 | −8 | 64 |
| 3 | 14 | 5 | 70 | 196 | −6 | 36 |
| 4 | 16 | 6 | 96 | 256 | −4 | 16 |
| 5 | 18 | 7 | 126 | 324 | −2 | 4 |
| 6 | 22 | 10 | 220 | 484 | 2 | 4 |
| 7 | 24 | 10 | 240 | 576 | 4 | 16 |
| 8 | 26 | 15 | 390 | 676 | 6 | 36 |
| 9 | 28 | 16 | 448 | 784 | 8 | 64 |
| 10 | 30 | 20 | 600 | 900 | 10 | 100 |
| sum | 200 | 91 | 2214 | 4,440 | 0 | 440 |
| mean | 20 | 9.1 | 221.4 | 444 | 0 | 44 |

2.1 From the simple regression model $Y_i = \beta_1 + \beta_2 X_i + u_i, u_i \sim NIID(0, \sigma^2)$

Find estimators of $\beta_1$ and $\beta_2$ from the OLS method and interpret the meaning.

Find $\hat{\beta}_2$

$$\hat{\beta}_2 = \frac{n\sum X_iY_i - \sum X_i \sum Y_i}{n\sum X_i^2 - (\sum X_i)^2}$$

$$= \frac{(10)(2214) - (200)(91)}{(10)(4440) - 40{,}000}$$

$$= \frac{22{,}140 - 18{,}200}{44{,}400 - 40{,}000}$$

$$= \frac{3{,}940}{4{,}400}$$

$\hat{\beta}_2 = 0.8955$ #

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$$

$$= 9.1 - (0.8955)20$$

$$\hat{\beta}_1 = -8.81$$

From model, $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$

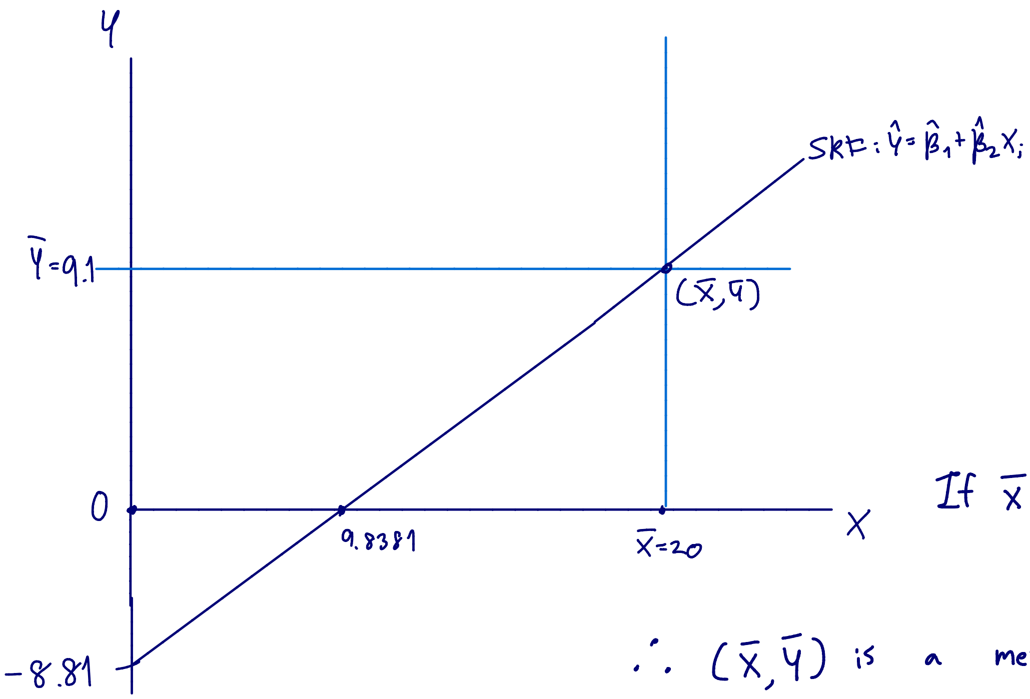
If $X_i = 0$, $\hat{Y}_i = -8.81$

and when $X_i$ rise by 1 unit, $Y_i$ will rise by 0.8955 unit.

2.2 Find the value of $\hat{Y}_i$ and $\hat{u}_i$. Show that $\sum \hat{u}_i \approx 0$

| $i$ | $X_i$ | $\hat{\beta}_1 + \hat{\beta}_2 X_i = \hat{Y}_i$ <br> $-8.81 + (0.8955)(X_i) = \hat{Y}_i$ | $Y_i - \hat{Y}_i = \hat{u}_i$ | $\hat{u}_i^2$ |
|---|---|---|---|---|
| 1 | 10 | 0.145 | $0 - 0.145 = -0.145$ | 0.021 |
| 2 | 12 | 1.936 | $2 - 1.936 = 0.064$ | 0.0041 |
| 3 | 14 | 3.727 | $5 - 3.727 = 1.273$ | 1.6205 |
| 4 | 16 | 5.518 | $6 - 5.518 = 0.482$ | 0.2323 |
| 5 | 18 | 7.309 | $7 - 7.309 = -0.309$ | 0.0955 |
| 6 | 22 | 10.891 | $10 - 10.891 = -0.891$ | 0.7939 |
| 7 | 24 | 12.682 | $10 - 12.682 = -2.682$ | 7.1931 |
| 8 | 26 | 14.473 | $15 - 14.473 = 0.527$ | 0.2777 |
| 9 | 28 | 16.264 | $16 - 16.264 = -0.264$ | 0.0697 |
| 10 | 30 | 18.055 | $20 - 18.055 = 1.945$ | 3.7830 |
| | | | $\sum \hat{u}_i = 0.145$ | $\sum \hat{u}_i^2 = 14.0908$ |

2.3 Plot graph and draw regression line. Does the line pass $(\bar{X}, \bar{Y})$?




If $\bar{X} = 20$, $\hat{Y}_i = -8.81 + 0.8955(20)$

$\hat{Y}_i = 9.1 \quad (\bar{Y})$

$\therefore$ $(\bar{X}, \bar{Y})$ is a member is this SRF family.

2.4 If $X_i = 18$, what is the predicted Y?

If $X_i = 18$, then $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$

$= -8.81 + (0.8955)(18)$

$\hat{Y}_i = 7.309$ #

## 2.5 Find $var(\hat{u}_i)$, $var(\hat{\beta}_1)$, $var(\hat{\beta}_2)$

$$var(\hat{u}_i) = \hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-2} = \frac{14.0908}{10-2} = 1.7614$$

$$var(\hat{\beta}_2) = \frac{\hat{\sigma}^2}{\sum x_i^2} = \frac{1.7614}{440} = 0.004$$

$$var(\hat{\beta}_1) = \frac{\sum X_i^2}{n \sum x_i^2} \cdot \hat{\sigma}^2 = \frac{4{,}440}{(10)(440)} \cdot 1.7614 = 1.7774$$

3. Consider the below regression function: consider the two-variable model

$$Y_i = \beta_1 + \beta_2 X_i + u_i,\ u_i \sim NIID(0, \sigma^2)$$

Find an OLS estimator of $\beta_1$. Then, provide a proof that this is an unbiased estimator.

Please state the assumption(s) of CLRM when used (pages 66-75 in Gujarati).

We need to find $\hat{\beta}_2$ first.

$\hat{\beta}_2$ is unbiased when $E(\hat{\beta}_2) = \beta_2$

$$\hat{\beta}_2 = \sum_{i=1}^{n} k_i Y_i$$

$$= \sum_{i=1}^{n} k_i (\beta_1 + \beta_2 X_i + u_i)$$

$$= \sum_{i=1}^{n} (k_i \beta_1 + \beta_2 k_i X_i + k_i u_i)$$

$$= \beta_1 \sum_{i=1}^{n} k_i + \beta_2 \sum_{i=1}^{n} k_i X_i + \sum_{i=1}^{n} k_i u_i \rightarrow \sum_{i=1}^{n} k_i = 0,\ \sum_{i=1}^{n} k_i X_i = 1$$

$$\hat{\beta}_2 = \beta_2 + \sum_{i=1}^{n} k_i u_i$$

$$E(\hat{\beta}_2) = E(\beta_2) + E\left(\sum_{i=1}^{n} k_i u_i\right) \rightarrow \text{Assumption 3}: E(u_i | X_i) = 0$$

$$E(\hat{\beta}_2) = \beta_2$$

So $\hat{\beta}_2$ is unbiased estimator

And $\hat{\beta}_1$ is unbiased when $E(\hat{\beta}_1) = \hat{\beta}_1$

Since $\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$

Therefore, $E(\hat{\beta}_1) = E(\bar{Y} - \hat{\beta}_2 \bar{X})$

$$= E(\bar{Y}) - \bar{X} E(\hat{\beta}_2)$$

And $E(\bar{Y}) = E\left[\frac{1}{n} \sum Y_i\right] = \frac{1}{n} \sum E(Y_i)$

$$= \frac{1}{n} \sum \beta_1 + \beta_2 X_i$$

$$= \frac{n\beta_1}{n} + \beta_2 \frac{\sum X_i}{n}$$

$$E(\bar{Y}) = \beta_1 + \beta_2 \bar{X}$$

From $E(\hat{\beta}_1) = E(\bar{Y}) - \bar{X} E(\hat{\beta}_2)$

If $E(\bar{Y}) = \beta_1 + \beta_2 \bar{X}$ and $\beta_2 = E(\hat{\beta}_2)$

Then $E(\hat{\beta}_1) = \beta_1 + \beta_2 \bar{X} - \bar{X} \beta_2$

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

$\hat{\beta}_1$ is unbiased estimator.