Solutions to HW1

Guanxi Li

April 14, 2024

1

$\mathbb{E}(\bar{Y}_n) = 0.4$, $\mathbb{V}(\bar{Y}_n) = \mathbb{V}(\frac{\sum_{i=1}^{n} Y_i}{n}) = 0.24/n$.

According to CLT, $Z_n = \frac{\bar{Y}_n - \mathbb{E}[\bar{Y}_n]}{\sqrt{\mathbb{V}[\bar{Y}_n]}} = \frac{\bar{Y}_n - 0.4}{\sqrt{0.24/n}} \xrightarrow{d} \mathcal{N}[0,1]$, then $Pr(\bar{Y}_n \geq m) = Pr(Z_n \geq \frac{m-0.4}{\sqrt{0.24/n}})$, $Pr(\bar{Y}_n \leq m) = Pr(Z_n \leq \frac{m-0.4}{\sqrt{0.24/n}})$.

1.1 n=100, m=0.43

$Pr(\bar{Y}_n \geq m) = Pr(Z_n \geq \frac{m-0.4}{\sqrt{0.24/n}}) = Pr(\bar{Y}_{100} \geq 0.43) = Pr(Z_{100} \geq \frac{0.43-0.4}{\sqrt{0.24/100}}) = 0.27014569.$

1.2 n=400, m=0.37

$Pr(\bar{Y}_n \leq m) = Pr(Z_n \leq \frac{m-0.4}{\sqrt{0.24/n}}) = Pr(\bar{Y}_{400} \leq 0.37) = Pr(Z_{400} \leq \frac{0.37-0.4}{\sqrt{0.24/400}}) = 0.11033568.$

1.3

$t = \frac{0.41-0.4}{\sqrt{0.24/n}} = 1.96$, $n = 9219.84$, so the minimum n is 9220.

2

Given $\mathbb{E}(\hat{\beta}_1|X_1,\cdots,X_n)=\beta_1$,

$$\begin{aligned}
\mathbb{E}(\hat{\beta}_0|X_1,\cdots,X_n) &= \mathbb{E}(\bar{Y}-\hat{\beta}_1\bar{X}|X_1,\cdots,X_n) \\
&= \mathbb{E}(\bar{Y}|X_1,\cdots,X_n)-\mathbb{E}(\hat{\beta}_1|X_1,\cdots,X_n)\bar{X} \\
&= \mathbb{E}(\beta_0+\beta_1\bar{X}+\bar{u}|X_1,\cdots,X_n)-\beta_1\bar{X} \\
&= \beta_0+\mathbb{E}(\bar{u}|X_1,\cdots,X_n) \\
&= \beta_0+\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(u_i|X_1,\cdots,X_n) \\
&= \beta_0. \blacksquare
\end{aligned}$$

3

3.1

To prove $TSS = ESS + SSR$,

$$\begin{aligned}
\sum_{i=1}^{n} \hat{u}_i(\hat{Y}_i - \bar{Y}) &= \sum_{i=1}^{n} \hat{u}_i \hat{Y}_i - \bar{Y} \sum_{i=1}^{n} \hat{u}_i \\
&= \sum_{i=1}^{n} \hat{u}_i(\hat{\beta}_0 + \hat{\beta}_1 X_i) \\
&= \hat{\beta}_1 \sum_{i=1}^{n} \hat{u}_i X_i \\
&= 0.
\end{aligned}$$

$$\begin{aligned}
TSS = \sum_{i=1}^{n}(Y_i - \bar{Y})^2 &= \sum_{i=1}^{2}(\hat{Y}_i + \hat{u}_i - \bar{Y})^2 \\
&= \sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2 + \sum_{i=1}^{n} \hat{u}_i^2 + 2\sum_{i=1}^{n} \hat{u}_i(\hat{Y}_i - \bar{Y}) \\
&= ESS + SSR + 2\sum_{i=1}^{n} \hat{u}_i(\hat{Y}_i - \bar{Y}) \\
&= ESS + SSR. \blacksquare
\end{aligned}$$

3.2

To prove $\rho_{XY}^2 = R^2$,

$$\begin{aligned}
RHS = \frac{ESS}{TSS} &= \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\
&= \frac{\sum_{i=1}^n (\hat{\beta}_0 + \hat{\beta}_1 X_i - \hat{\beta}_0 - \hat{\beta}_1 \bar{X})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\
&= \frac{\hat{\beta}_1^2 \sum_{i=1}^n (X_i - \bar{X})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\
&= \frac{[\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})]^2 \sum_{i=1}^n (X_i - \bar{X}_i)^2}{[\sum_{i=1}^n (X_i - \bar{X})^2]^2 \sum_{i=1}^n (Y_i - \bar{Y})^2} \\
&= \frac{[\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})]^2}{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2} \\
&= LHS. \blacksquare
\end{aligned}$$

4

4.1

The 95% confidence interval is $43.2 \pm 10.2t^{\alpha} = 22.29$ to 64.11.

4.2

$t^{act} = \frac{61.5-55}{7.4} = 0.8784 < 2.05$, accept H_0 and reject H_1.

4.3

$t^{act} = \frac{61.5-55}{7.4} = 0.8784 < 1.70$, accept H_0 and reject H_1.

5

Table 1: Results

	(1) ahe_all	(2) ahe_hs	(3) ahe_ba
age	0.605***	0.298***	0.925***
	(0.0403)	(0.0427)	(0.0606)
Constant	1.082	6.522***	−4.439**
	(1.167)	(1.270)	(1.800)
N	7,711	4,002	3,709
R^2	0.029	0.012	0.059

* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

5.1

The coefficient is significant at the 10 percent level, 5 percent level and 1 percnt level. $t^{act} = 15.02 > 2.58$, so the null hypothesis can be rejected at the 1% confidence level.

5.2

The 5% confidence interval is $0.605 \pm 0.0403 \times 1.96 = 0.526$ to 0.684.

5.3

The coefficient is significant at the 10 percent level, 5 percent level and 1 percnt level. $t^{act} = 6.97 > 2.58$, so the null hypothesis can be rejected at the 1% confidence level.

The 5% confidence interval is $0.298 \pm 0.0427 \times 1.96 = 0.214$ to 0.381.

5.4

The coefficient is significant at the 10 percent level, 5 percent level and 1 percnt level. $t^{act} = 15.27 > 2.58$, so the null hypothesis can be rejected at the 1% confidence level.

The 5% confidence interval is $0.925 \pm 0.0606 \times 1.96 = 0.806$ to 1.043.

5.5

The difference in $\hat{\beta}_1$ is $0.925 - 0.298 = 0.627$, the standard error of $\hat{\beta}_1$ is $(0.0426^2 + 0.0606^2)^{0.5} = 0.074$. $t^{act} = \frac{0.627}{0.074} = 8.47 > 2.58$, so the null hypothesis can be rejected at the 1% level.

The difference can be explained by the tendency for individuals with higher education to opt for higher-paying and technology-intensive jobs. Therefore, all else being equal, average hourly earnings (AHE) increases more on average as age increases by one unit for those with bachelor degree.