Homework 8

Tuesday, December 3, 2024 2:53 PM

Exercise 1

a) Using R:

$\bar{x} = 109.341$

$\bar{y} = 19.987$

$S_{xx} = 521.196$

$S_{xy} = -471.542$

b) From notes: $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \cdot \bar{x}$

then

$$\hat{\beta}_1 = \frac{-471.542}{521.196} = \boxed{-0.90473}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \cdot \bar{x} = 19.987 + 0.90473 \cdot 109.34 = \boxed{118.91}$$

c) $$SSE = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i))^2$$

$$= \boxed{11.439}$$

d) $SE(\hat{\beta}_1) = \sqrt{\frac{\sigma_\varepsilon^2}{S_{xx}}} \Rightarrow$ compute $\sqrt{\frac{\hat{\sigma}_\varepsilon^2}{S_{xx}}}$

$$\hat{\sigma}_\varepsilon^2 = \frac{SSE}{n-2} = \frac{11.439}{15-2} = 0.87992$$

then $SE(\hat{\beta}_1) = \sqrt{\frac{0.87992}{521.196}} = \boxed{0.04109}$

$\hat{\beta}_1$ normally distributed around β_1 w/ $Var(\hat{\beta}_1) = SE(\hat{\beta}_1)^2$; for test stat T, $T = \frac{(\hat{\beta}_1 - \beta_1)}{SE(\hat{\beta}_1)} \sim t_{n-2}$

Want to see if $T \in R \mid H_0$ true for some R; since $T \sim t_{n-2}$ we construct $R = (-\infty, -t_{\alpha/2, n-2}] \cup [t_{\alpha/2, n-2}, \infty)$

we construct $K = (-\infty, -t_{\alpha/2, n-2}] \cup [t_{\alpha/2, n-2}, \infty)$

$$= (-\infty, -2.16] \cup [2.16, \infty).$$

Have that $T = \dfrac{(\hat{\beta}_1 - \beta_1)}{SE(\hat{\beta}_1)}$; since H_0 is that $\beta_1 = 0$, compute T

to be $\dfrac{-0.90473}{0.04109} \approx -22.018 \in R$, so can reject H_0 in favor of H_A.

Exercise 2

a) By assumption, $Y_i = \underbrace{\beta_0 + \beta_1 x_i}_{\text{given}} + \underbrace{\varepsilon_i}_{\text{random}}$, with $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$

So then we have that $\boxed{Y_i \sim N(b_0 + b_1 x_i, \sigma_\varepsilon^2)}$

Thus using the formula for PDF of normal dist., we have

$$f_{Y_i}(y_i \mid b_0, b_1, x_i, \sigma_\varepsilon^2) = \boxed{\frac{1}{\sqrt{2\sigma_\varepsilon^2 \pi}} e^{-\frac{[y_i - (b_0 + b_1 x_i)]^2}{2\sigma_\varepsilon^2}}}$$

b) Since ε_i is only source of randomness & are independent, then Y_i are indep, so

$$f(y_1, y_2, \dots, y_n \mid b_0, b_1, \sigma_\varepsilon^2, x_1, \dots, x_n) = \prod_{i=1}^{n} f_{Y_i}(y_i \mid b_0, b_1, x_i, \sigma_\varepsilon^2)$$

$$= \prod_{i=1}^{n} \left[\frac{1}{\sqrt{2\sigma_\varepsilon^2 \pi}} e^{-\frac{[y_i - (b_0 + b_1 x_i)]^2}{2\sigma_\varepsilon^2}} \right] = \left(\frac{1}{2\sigma_\varepsilon^2 \pi} \right)^{n/2} e^{-\sum_{i=1}^{n} \frac{[y_i - (b_0 + b_1 x_i)]^2}{2\sigma_\varepsilon^2}}$$

c) To find MLE of β_0 & β_1, consider likelihood fxn. accross all possible realizations b_0 and b_1 of β_0, β_1 respectively.

From b), have likelihood fxn.

$$\mathcal{L}(b_0, b_1; (x_i, y_i)_{i=1}^n) = \left(\frac{1}{2\sigma_\varepsilon^2 \pi}\right)^{n/2} e^{-\sum_{i=1}^n \frac{[y_i - (b_0 + b_1 x_i)]^2}{2\sigma_\varepsilon^2}}$$ so

$$\ln(\mathcal{L}(b_0, b_1; (x_i, y_i)_{i=1}^n)) = \underbrace{-\frac{n}{2}\ln(2\sigma_\varepsilon^2\pi)}_{\text{constant}} - \sum_{i=1}^n \frac{[y_i - (b_0 + b_1 x_i)]^2}{2\sigma_\varepsilon^2},$$

So want to maximize $-\sum_{i=1}^n \frac{[y_i - (b_0 + b_1 x_i)]^2}{2\sigma_\varepsilon^2} = -\frac{1}{2\sigma_\varepsilon^2} \overbrace{\sum_{i=1}^n [y_i - (b_0 + b_1 x_i)]^2}^{\text{Formula for } n \cdot MSE},$

So we are maximizing $\frac{-n}{2\sigma_\varepsilon^2} \cdot MSE$. Since

$MSE \geq 0$, $\frac{-n}{2\sigma_\varepsilon^2} < 0$, then the values for b_0 & b_1 that maximize the desired quantity <u>minimize</u> MSE; these are $\hat{\beta}_0$ and $\hat{\beta}_1$, as given.

Thus we have that

$\hat{\beta}_0^{MLE} = \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ and

$\hat{\beta}_1^{MLE} = \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$

<u>Exercise 3</u>

a) know that:

$\frac{S_{xx}}{n} = \frac{1}{100}\sum_{i=1}^{100} x_i^2 - \bar{x}^2$ $\quad \frac{S_{xy}}{n} = \frac{1}{100}\sum_{i=1}^{100} x_i y_i - \bar{x}\bar{y}$ and that

$\bar{x} = \frac{1}{100}\sum_{i=1}^{100} x_i = 3$, $\bar{y} = \frac{1}{100}\sum_{i=1}^{100} y_i = 4$ so have

$\frac{S_{xx}}{n} = \frac{1300}{100} - 3^2 = 4$, $\frac{S_{xy}}{n} = \frac{1500}{100}$ $(3)(4) = 3$ and then since $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$ under the assumed conditions, $\hat{\beta}_1 = \frac{3}{4} = 0.75$. Note that $S_{xx} = 400$, $S_{xy} = 300$.

b) Showing $\hat{\sigma}_\varepsilon = 10$:

$$SSE = n \cdot MSE = \sum_{i=1}^{n} \left[y_i - (b_0 + b_1 x_i)\right]^2 = \sum_{i=1}^{n} \left[y_i^2 - 2y_i(b_0 + b_1 x_i) + (b_0 + b_1 x_i)^2\right] =$$

$$\sum_{i=1}^{n} y_i^2 - 2b_0 \sum_{i=1}^{n} y_i - 2b_1 \sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} (b_0 + b_1 x_i)^2 =$$

$$\sum_{i=1}^{n} y_i^2 - 2b_0 \sum_{i=1}^{n} y_i - 2b_1 \sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} b_0^2 + b_0 b_1 \sum_{i=1}^{n} x_i^2 + b_1^2 \sum_{i=1}^{n} x_i^2.$$

Using $b_1 = \hat{\beta}_1 = 0.75$, $b_0 = \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 1.75$ from (a), plug into the above to obtain

$$SSE = 11625 - 2(1.75)(400) - 2(0.75)(1500) + (100)(1.75)^2 + (1.75)(0.75)(1300) + (0.75)^2(1300)$$
$$= 10{,}718.75$$

Then $\hat{\sigma}_\varepsilon = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{10{,}718.75}{100-2}} = 10.458 \approx 10$, as desired.

Then $SE(\hat{\beta}_1) = \sqrt{\frac{\sigma_\varepsilon^2}{S_{xx}}} = \sqrt{\frac{10^2}{400}} = 0.5$

c) Construct a 2-sided rejection rgn. R:

Since $\frac{\sqrt{S_{xx}}\,(\hat{\beta}_1 - \beta_1)}{\hat{\sigma}_\varepsilon} \sim t_{n-2}$, and $\beta_1 = 0$ under H_0, have

$$\frac{\sqrt{S_{xx}}\,(\hat{\beta}_1 - \beta_1)}{\hat{\sigma}_\varepsilon} = \frac{\hat{\beta}_1 \sqrt{S_{xx}}}{\hat{\sigma}_\varepsilon} = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} \sim t_{n-2} \text{ with } n = 100.$$

$$\hat{\sigma}_\varepsilon \quad - \quad \frac{\hat{\beta}_1}{\hat{\sigma}_\varepsilon} = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} \sim t_{n-2} \quad \text{with } n = 100.$$

Then for a 5% sig. lvl test, $R = (-\infty, -t_{2.5\%,98}] \cup [t_{2.5\%,98}, \infty)$

$= (-\infty, -1.98] \cup [1.98, \infty)$.

Since $\hat{\beta}_1 = 1.1526$ from (a), $\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} = \frac{0.75}{0.5} = 1.5 \notin R$,

so we $\boxed{\text{cannot}}$ reject H_0 in favor of H_A.

d) $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$

$\frac{S_{xy}}{n} = \frac{1}{400}\sum_{i=1}^{400} x_i y_i - \bar{x}'\bar{y}'$, $\frac{S_{xx}}{n} = \frac{1}{400}\sum_{i=1}^{400} x_i^2 - \bar{x}'^2$

Note $\sum_{i=1}^{400} x_i y_i = 4\sum_{i=1}^{100} x_i y_i$, $\sum_{i=1}^{400} x_i^2 = 4\sum_{i=1}^{100} x_i^2$, $\bar{x}' = \bar{x}$, $\bar{y}' = \bar{y}$

with $\bar{x}'$, $\bar{y}'$ denoting $\frac{1}{400}\sum_{i=1}^{400} x_i$ and $\frac{1}{400}\sum_{i=1}^{400} y_i$.

Thus $\underbrace{S_{xy} = 1200, S_{xx} = 1600}_{\text{change}}$, $\underbrace{\hat{\beta}_1 = 0.75}_{\text{same}}$.

e) From (b) SSE =

$$\sum_{i=1}^{n} y_i^2 - 2b_0\sum_{i=1}^{n} y_i - 2b_1\sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} b_0^2 + b_0 b_1 \sum_{i=1}^{n} x_i^2 + b_1^2 \sum_{i=1}^{n} x_i^2.$$

Using $b_1 = \hat{\beta}_1 = 0.75$, $b_0 = \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = 1.75$, note that the summation in each term of the above eqn. is exactly equal to 4x the value in part (b); thus, the new SSE is 4x the old and is therefore 42,875

$$\hat{\sigma}_\varepsilon = \sqrt{\frac{SSE}{\ }} = \sqrt{\frac{42,875}{\ }} = 10.37 \approx 10$$

$$\hat{\sigma}_\varepsilon = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{42,875}{400-2}} = 10.37 \approx 10$$

Then $SE(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}_\varepsilon^2}{1600}} = \sqrt{\frac{10^2}{1600}} = 0.25$

Construct a 2-sided rejection rgn. R:

Since $\frac{\sqrt{S_{xx}}\,(\hat{\beta}_1 - \beta_1)}{\hat{\sigma}_\varepsilon} \sim t_{n-2}$, and $\beta_1 = 0$ under H_0, have

$$\frac{\sqrt{S_{xx}}\,(\hat{\beta}_1 - \beta_1)}{\hat{\sigma}_\varepsilon} = \frac{\hat{\beta}_1 \sqrt{S_{xx}}}{\hat{\sigma}_\varepsilon} = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} \sim t_{n-2} \text{ with } n = 400.$$

Then for a 5% sig. lvl test, $R = (-\infty, -t_{2.5\%, 398}] \cup [t_{2.5\%, 398}, \infty)$

$= (-\infty, -1.97] \cup [1.97, \infty)$.

Since $\hat{\beta}_1 = 0.75$ from (a), $\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} = \frac{0.75}{0.25} = 3 \in R$,

so we $\boxed{\text{can}}$ reject H_0 in favor of H_A.