MA2261 - DLI, Linear Statistical Models, Year 2022-2023

Solutions of exercises for feedback class 4
(Note: the exercise number refers to the workbook)

EXERCISE 3.3 ii) and iii)

ii)
$$RSS = S_{yy} - \frac{S_{xy}^2}{S_{xx}} = 17.367 - \frac{83.723^2}{628.672} = 6.217$$
$$\hat{\sigma}^2 = \frac{RSS}{21-2} = \frac{6.217}{19} = 0.3272$$

To test the hypothesis $b = 0$ we calculate

$$\frac{\hat{b}\sqrt{S_{xx}}}{\hat{\sigma}} = \frac{0.133\sqrt{628.672}}{\sqrt{0.3272}} = 5.83 \sim t_{19}$$

The critical region is $(-\infty, -2.093) \cup (2.093, +\infty)$. The hypothesis $b = 0$ is rejected. Hence the variable income is statistically significant in determining the mean life satisfaction.

iii) Since the hypothesis $b = 0$ is rejected, the 95% confidence interval for b does not contain 0.

EXERCISE 3.4 v)

v) The confidence interval in part ii) does not contain 0. This is expected from the fact the hypothesis $b = 0$ is equivalent to $\rho = 0$; thus since $b = 0$ is rejected, also $\rho = 0$ is rejected, and the 95% confidence interval for ρ does not contain 0.

EXERCISE 3.6

a) $\hat{b} = \frac{S_{xy}}{S_{xx}} = 4, \quad \hat{a} = \bar{y} - \hat{b}\bar{x} = 10.2.$

The regression function is $\hat{y} = 10.2 + 4x$.

b) $RSS = S_{yy} - \dfrac{S_{xy}^2}{S_{xx}} = 17.6$, $\hat{\sigma}^2 = \frac{RSS}{(n-2)} = \frac{17.6}{8} = 2.2, \quad \hat{\sigma} = 1.48.$

The 95% confidence interval for a is

$$\hat{a} \pm t_{0.025,n-2}\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}} = 10.2 \pm 2.306 \times 1.48\sqrt{1/10 + 1/10} = (8.67,\ 11.73).$$

The 95% confidence interval for b is

$$\hat{b} \pm t_{0.025,n-2}\frac{\hat{\sigma}}{\sqrt{S_{xx}}} = 4 \pm 2.306 \times \frac{1.48}{\sqrt{10}} = (2.92,\ 5.08).$$

The 95% confidence interval for σ is

$$\left(\hat{\sigma}\sqrt{\frac{n-2}{\chi^2_{0.025,n-2}}}\ ,\ \hat{\sigma}\sqrt{\frac{n-2}{\chi^2_{0.975,n-2}}}\right) = \left(1.48\sqrt{\frac{8}{17.53}}\ ,\ 1.48\sqrt{\frac{8}{2.18}}\right) = (1,\ 2.84).$$

c) We need to test the hypothesis $b = 0$. We have $\dfrac{\hat{b}\sqrt{S_{xx}}}{\hat{\sigma}} = \dfrac{4\sqrt{10}}{1.48} = 8.55 \sim t_8$

The critical region is $(-\infty, -2.306) \cup (2.306, +\infty)$. Therefore the null hypothesis $b = 0$ is rejected.

e) $\hat{\rho}^2 = \dfrac{S_{yy} - RSS}{S_{yy}} = 0.9$. Hence 90% of the variation in number of broken ampules is explained bt the number of transfers.

EXERCISE 3.7

a) The 95% confidence interval for the expected number of broken ampules when $x_0 = 2$ is given by

$$\hat{a}+\hat{b}x_0 \pm t_{0.025,n-2}\hat{\sigma}\sqrt{\frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}} = 10.2+4\times 2 \pm 2.306\times 1.48\sqrt{\frac{1}{10} + \frac{(2-1)^2}{10}}$$

$= (16.67,\ 19.73).$

b) The 95% prediction interval for the number of broken ampules when two transfers are made is

$$\hat{a}+\hat{b}x_0 \pm t_{0.025,n-2}\hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}} = 10.2+4\times 2 \pm 2.306\times 1.48\sqrt{1 + \frac{1}{10} + \frac{(2-1)^2}{10}}$$

$= (14.453,\ 21.947).$