EXERCISE 2

$n = 18$

x_{i1} = number of repairs of call i

x_{i2} = type of computer of call i

y_i = number of minutes of call i

a) $Y_i = \beta_1 + \beta_2 x_{i1} + \beta_3 x_{i2} + \varepsilon_i$ $\qquad \varepsilon_i \sim N(0, \sigma^2)$ indep

with $x_{i2} = \begin{cases} 1 & \text{if computer } i \text{ is personal} \\ 0 & \text{otherwise} \end{cases}$

Assumptions:

- linearity w.r.t. $(\beta_1, \beta_2, \beta_3)$
- $\varepsilon_i \sim N(0, \sigma^2)$ indep $i = 1, \dots, 18$
- absence of multicollinearity

b) $\underset{\text{SST}}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = \underset{\text{SSR}}{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2} + \underset{\text{SSE}}{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}$

with SST = total sum of squares

SSR = regression " " "

SSE = error (residual) " " "

since $S_y^2 = \dfrac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n-1} = 1040.889 \Rightarrow SST = 1040.889 \cdot 17 = 17'695.11$

moreover, $R^2 = \dfrac{SSR}{SST} = 0.9808 \Rightarrow SSR = 0.9808 \cdot SST = 0.9808 \cdot 17'695.11 = 17'355.36$

Hence $SSE = SST - SSR = 17'695.11 - 17'355.36 = 339.75$

c) $\begin{cases} H_0: \beta_2 = \beta_3 = 0 \\ H_1: \text{at least one of } (\beta_2, \beta_3) \text{ is } \neq 0 \end{cases}$ equivalent to $\begin{cases} H_0: R^2 = 0 \\ H_1: R^2 \neq 0 \end{cases}$

test statistic

$F = \dfrac{R^2}{1-R^2} \cdot \dfrac{n-p}{p-1} = \dfrac{R^2}{1-R^2} \cdot \dfrac{15}{2} \overset{H_0}{\sim} F_{2,15}$

observed value $f^{obs} = \dfrac{0.9808}{1-0.9808} \cdot \dfrac{15}{2} = 383.125$

The reject region of the test for a level α is $R = (f_{2,15;1-\alpha}; +\infty)$

If $\alpha = 0.001$, $f_{2,15;0.999} = 11.33$

Hence I reject H_0.

d) model B

$Y_i = \beta_1 + \beta_2 x_{i1} + \beta_3 x_{i2} + \beta_4 x_{i3} + \varepsilon_i$ $\qquad \varepsilon_i \sim N(0, \sigma^2)$ indep

with $x_{i2} = \begin{cases} 1 & \text{if computer } i \text{ is personal} \\ 0 & \text{otherwise} \end{cases}$

and $x_{i3} = x_{i1} \cdot x_{i2} = \begin{cases} x_{i1} & \text{if computer } i \text{ is personal} \\ 0 & \text{otherwise} \end{cases}$

e) No, they are nested hence $R_B^2 \geq R_A^2$ by construction

f) $SSE_B = 285.657$

Considering model B, the hyp. are

$\begin{cases} H_0: \beta_4 = 0 \\ H_1: \beta_4 \neq 0 \end{cases}$

model B is the complete model, model A the reduced

test statistic $\qquad F = \dfrac{SSE_{red} - SSE_{full}}{SSE_{full}} \cdot \dfrac{n-p}{p-p_0}$

$F = \dfrac{SSE_A - SSE_B}{SSE_B} \cdot \dfrac{n-p}{p-p_0} = \dfrac{SSE_A - SSE_B}{SSE_B} \cdot \dfrac{14}{1} \overset{H_0}{\sim} F_{1,14}$

with the data $f^{obs} = \dfrac{339.75 - 285.657}{285.657} \cdot 14 = 2.65$

the reject region is $R = (f_{1,14;1-\alpha}; +\infty)$

I do not reject H_0 at usual sig. levels. I prefer model A.