Assignment 2

Rin Meng
Student ID: 51940633

October 23, 2024

1. Given the summary output, it is true that:

   (a)

   $$\text{F-Statistic} = 52.77$$

   $$\text{MSE} = (\text{Residual Standard Error})^2 = 2.792^2 = 7.80$$

   (b) To build the anova table, we need to do calculations as follows:
   Finding the degrees of freedom:

   $$DF_{Reg} = 1$$

   $$DF_{Error} = 7$$

   $$DF_{Total} = 8$$

   $$SS_{Total} = SS_{Reg} + SS_{Error}$$

   Calculate the Mean Squares Reg and Error:

   $$F = \frac{MS_{Reg}}{MS_{Error}} = 52.77$$

   $$MS_{Error} = MSE = 7.80$$

   $$MS_{Reg} = F \cdot MS_{Error} = 52.77 \cdot 7.80 = 411.34$$

   Calculate the Sum Squares Reg and Error:

   $$MS_{Reg} = \frac{SS_{Reg}}{DF_{Reg}} \Leftrightarrow SS_{Reg} = MS_{Reg} \cdot DF_{Reg}$$

$$MS_{Error} = \frac{SS_{Error}}{DF_{Error}} \Leftrightarrow SS_{Error} = MS_{Error} \cdot DF_{Error}$$
$$SS_{Total} = SS_{Reg} + SS_{Error}$$
$$SS_{Reg} = MS_{Reg} \cdot DF_{Reg} = 411.34 \cdot 1 = 411.34$$
$$SS_{Error} = MS_{Error} \cdot DF_{Error} = 7.80 \cdot 7 = 54.60$$
$$SS_{Total} = SS_{Reg} + SS_{Error} = 411.34 + 54.60 = 465.94$$
$$MS_{Total} = \frac{SS_{Total}}{DF_{Total}} = \frac{465.94}{8} = 58.24$$

Source	DF	SS	MS	F
Reg.	1	411.34	411.34	52.77
Error	7	54.60	7.80	
Total	8	465.94	58.24	

(c) The anova function returns:

```
Analysis of Variance Table

Response: R
          Df Sum Sq Mean Sq F value    Pr(>F)
W          1 411.42  411.42  52.767 0.0001679 ***
Residuals  7  54.58    7.80
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Our hand calculations are fairly consistent with the output from the anova function in R, minus a few rounding errors.

(d) Calculating the $\sqrt{F}$ value:

$$\sqrt{F} = \sqrt{52.77} = 7.27$$

The t value for $\hat{\beta}_1$ is 7.264, which is very close to the $\sqrt{F}$ value. This is expected because, for simple linear regression with one predictor, the square of the t-value for the slope is equal to the F-statistic

$$\sqrt{F} = t \Leftrightarrow F = t^2$$

2. Given

(a) - **First task time**: ϵ is exponentially distributed with mean $\frac{1}{\lambda}$, so $E[\epsilon] = \frac{1}{\lambda}$.

- **Second task time**: Proportional to x, with a proportionality constant β. So the time required should be βx.
- **Total time**: sum of times it takes to complete the two tasks impying that the total time is $y = \beta x + \epsilon$.

Then the final linear model would be

$$y = \beta x + \epsilon$$

(b) To derive the maximum likelihood estimator for β and λ, we need to find the pdf of the exponential distribution.

$$f(\epsilon) = \lambda e^{-\lambda \epsilon} \text{ for } x \geq 0$$

$$\Rightarrow f(y) = \lambda e^{-\lambda(y-\beta x)} \text{ for } y \geq \beta x$$

Deriving maximum likelihood estimator for β:

$$\begin{aligned}
L(\beta, \lambda) &= \prod_{i=1}^{n} \lambda e^{-\lambda(y_i - \beta x_i)} \\
\log(\beta, \lambda) = \ell(\beta, \lambda) &= \sum_{i=1}^{n} \log(\lambda e^{-\lambda(y_i - \beta x_i)}) \\
\frac{\partial \ell(\beta, \lambda)}{\partial \beta} &= \sum_{i=1}^{n} \frac{\partial}{\partial \beta} \left(\log(\lambda e^{-\lambda(y_i - \beta x_i)})\right) \\
\frac{\partial \ell(\beta, \lambda)}{\partial \beta} &= \sum_{i=1}^{n} \frac{\partial}{\partial \beta} (\log \lambda - \lambda(y_i - \beta x_i)) \\
\frac{\partial \ell(\beta, \lambda)}{\partial \beta} &= \sum_{i=1}^{n} -x_i \lambda + x_i \lambda \beta \\
\frac{\partial \ell(\beta, \lambda)}{\partial \beta} &= \sum_{i=1}^{n} -x_i \lambda + x_i \lambda \beta = 0 \\
\sum_{i=1}^{n} x_i \lambda &= \sum_{i=1}^{n} x_i \lambda \beta \\
\beta &= 1
\end{aligned}$$

Deriving maximum likelihood estimator for λ:

$$\begin{aligned}
L(\beta, \lambda) &= \prod_{i=1}^{n} \lambda e^{-\lambda(y_i - \beta x_i)} \\
\log L(\beta, \lambda) = \ell(\beta, \lambda) &= \sum_{i=1}^{n} \log \lambda e^{-\lambda(y_i - \beta x_i)} \\
\frac{\partial \ell(\beta, \lambda)}{\partial \lambda} &= \sum_{i=1}^{n} \frac{\partial}{\partial \lambda} \left(\log \lambda e^{-\lambda(y_i - \beta x_i)} \right) \\
\frac{\partial \ell(\beta, \lambda)}{\partial \lambda} &= \sum_{i=1}^{n} \frac{\partial}{\partial \lambda} \left(\log \lambda - \lambda(y_i - \beta x_i) \right) \\
\frac{\partial \ell(\beta, \lambda)}{\partial \lambda} &= \sum_{i=1}^{n} \frac{1}{\lambda} - (y_i - \beta x_i) \\
\frac{\partial \ell(\beta, \lambda)}{\partial \lambda} &= \sum_{i=1}^{n} \frac{1}{\lambda} - (y_i - \beta x_i) = 0 \\
\sum_{i=1}^{n} \frac{1}{\lambda} &= \sum_{i=1}^{n} (y_i - \beta x_i) \\
\frac{n}{\lambda} &= \sum_{i=1}^{n} (y_i - \beta x_i) \\
\lambda &= \frac{n}{\sum_{i=1}^{n} (y_i - \beta x_i)}
\end{aligned}$$

$\therefore$ The maximum likelihood estimator for β is 1 and λ is $\frac{n}{\sum_{i=1}^{n}(y_i - \beta x_i)}$.

End of Assignment 2.