T6

1. The relationship between the concentration of ozone in New York and other environmental factors is under investigation. The concentration Y, the solar radiation X_1 and the temperature in New York X_2 are collected for 30 consecutive days.

   Consider a multiple linear regression model. Let $\mathbf{Y}$ be the vector of the responses and $\mathbf{X}$ be the design matrix. The following matrices are given for the regression analysis.

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 30 & 5733 & 2085 \\ 5733 & 1463179 & 411076 \\ 2085 & 411076 & 147243 \end{pmatrix} \quad \mathbf{X}'\mathbf{Y} = \begin{pmatrix} 83.84 \\ 17097.38 \\ 5953.47 \end{pmatrix}$$

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 2.2157 & 6.176 \times 10^{-4} & -3.310 \times 10^{-2} \\ 6.176 \times 10^{-4} & 3.341 \times 10^{-6} & -1.807 \times 10^{-5} \\ -3.310 \times 10^{-2} & -1.807 \times 10^{-5} & 5.259 \times 10^{-4} \end{pmatrix}$$

   $\mathbf{Y}'\mathbf{Y} = 251.0472$

   (i) Find the least squares estimators, $\widehat{\beta}_0, \widehat{\beta}_1$ and $\widehat{\beta}_2$, for the regression coefficients.

   (ii) Is the overall regression significant at a 5% level of significance ?

   Consider a reduced regression model without the predictor X_1, $y_i = \beta_0 + \beta_1 x_{2i} + \varepsilon_i$. Denote the design matrix in the reduced model by

$$(\mathbf{X}_r'\mathbf{X}_r)^{-1} = \begin{pmatrix} 2.1015 & -0.029756 \\ -0.029756 & 0.000428 \end{pmatrix}$$

   (iii) Find the least squares estimators $\widehat{\beta}_0, \widehat{\beta}_1$ in the reduced model.

   (iv) Find the standard error for each estimator. Is the overall regression significant at a 5% level of significance ?

   (v) Use an F test to check the significance of the reduced model at a 5% level of significance ?

2. Consider a multiple regression model with two predictors $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$, $\varepsilon_i \sim N(0, \sigma^2)$ and its reduced model $y_i = \beta_0 + \beta_1 x_{i1} + \varepsilon_i$,. Let SSE_ω be the residual sum of square of the reduced model and SSE_Ω the residual sum of square of the full model. Show that $SSE_\Omega \leq SSE_\omega$.