Tutorial 5

1. Consider a multiple linear regression (MLR) model with one response (Y) and two predictors (X_1 and X_2), $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$, where $\epsilon \sim N(0, \sigma^2)$. Suppose that there are n observations, i.e. we want to analyze the following table of numbers,

Y	X_1	X_2
y_1	x_{11}	x_{12}
y_2	x_{21}	x_{22}
$\vdots$	$\vdots$	$\vdots$
y_n	x_{n1}	x_{n2}

   (i) Write out the matrix form of the above MLR model, including write out the matrices, $\mathbf{Y}$, $\mathbf{X}$ and ϵ, and the corresponding assumptions.

   (ii) Derive the distribution of $\mathbf{c}'\widehat{\beta}$, where $\mathbf{c} = (c_0, c_1, c_2)'$ and $\widehat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ is the least squares estimator.

   If we are further given that $n = 25$

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 2.779 & -0.0112 & -0.106 \\ -0.0112 & 0.146 \times 10^{-3} & 0.175 \times 10^{-3} \\ -0.106 & 0.175 \times 10^{-3} & 0.479 \times 10^{-2} \end{pmatrix}, \quad \mathbf{X}'\mathbf{Y} = \begin{pmatrix} 235.6 \\ 11821.432 \\ 4831.86 \end{pmatrix}$$

   and $s^2 = 0.4377$.

   (iv) Calculate the least squares estimator $\widehat{\beta}$.

   (v) Test for the hypothesis that $\beta_1 + \beta_2 = 0$ at the significant level of 0.05. (Some quantiles of t-distribution: $t_{25}^{0.05/2} = 2.060$, $t_{24}^{0.05/2} = 2.064$ and $t_{22}^{0.05/2} = 2.074$)

2. Many people believe that the length of one's life is linearly related to the length of lifeline on one's left hand. Dr. L.E. Mather and Dr. M.E. Wilson conducted a experiment to study this belief in 1974. In their experiment, two variables were involved,

   Y = Age of person at death (to nearest year) and
   X = Length of lifeline on left hand in centimeters (to nearest 0.15cm),

   and 50 pairs of observations were collected. The first four observations are listed as below.

Case	Y = Age (Year)	X = Length (cm)
1	19	9.75
2	40	9.00
3	42	9.60
4	42	9.75
⋮	⋮	⋮
50	94	9.00

   The summary of the data is as follows.

$$\sum y = 3333, \quad \sum y^2 = 231933, \quad \sum xy = 30549.75,$$

$$\sum x = 459.9, \ \sum x^2 = 4308.57.$$

   Consider a simple linear regression (SLR) model to fit the above data.

   (i) Write out the fitted SLR model, including the estimation of β_0, β_1 and σ^2.

   (ii) Use t test to check whether or not the length of lifeline can affect the lifetime at a size of 5%.

   (iii) Construct the ANOVA table, and check whether the fitted SLR model is significant.

   (v) The length of lifeline on my left hand is 9.32 cm. Tell me the mean lifetime of persons with this length of lifeline. Given me an interval with 95% confidence.