Tutorial Questions 1

1. For model

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

assume that $X = 0$ is within the scope of the model. What is the implication for the regression function $Y_i = \beta_0 + \beta_1 X_i$ if $\beta_0 = 0$ so that the model is $Y_i = \beta_1 X_i + \varepsilon_i$? How would the regression function $Y_i = \beta_0 + \beta_1 X_i$ plot on the graph? [hint: what is the interpretation of β_0]

2. Refer to the regression model above, what is the implication for the regression function if $\beta_1 = 0$ so that the model is $Y_i = \beta_0 + \varepsilon_i$? How would the regression function $Y_i = \beta_0 + \beta_1 X_i$ plot on the graph? [hint: what is the interpretation of β_1]

3. (Grade Point average): For graduate students, X is ACT test score, Y is GPA. There are 120 students, and their ACT and GPA are recorded. The data can be found at **(dataTutorial1a.dat)**. Suppose we predict Y based on X by a simple linear regression model above.

   (a) Obtain the LSE of β_0 and β_1

   (b) plot the estimated regression function and the data. Does the estimated regression appear to fit the data well?

   (c) Obtain a point prediction of GPA for a student with ACT $X = 30$

   (d) What is the change of the mean response when ACT increases by one point?

   (e) plot the fitted residuals e_i against X_i

   (f) calculate $\sum_{i=1}^{120} e_i^2$

4. Derive the expression for

$$b_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

from the normal equations.

5. For the model in the first question, prove that the sum of Y_i's is the same as the sum of fitted values

6. For the model in the first question, prove that $\sum_{i=1}^{n} e_i \hat{Y}_i = 0$

7. Show that the least squares regression line fitted to data $(5, Y_{1,1})$, $(5, Y_{1,2})$, $(5, Y_{1,3})$, $(10, Y_{2,1})$, $(10, Y_{2,2})$, $(10, Y_{2,3})$, $(15, Y_{3,1})$, $(15, Y_{3,2})$, $(15, Y_{3,3})$, is the same as a model fitted model to the three points $(5, \bar{Y}_1)$, $(10, \bar{Y}_2)$ and $(15, \bar{Y}_3)$, where $\bar{Y}_1 = (Y_{1,1} + Y_{1,2} + Y_{1,3})/3, \bar{Y}_2 = (Y_{2,1} + Y_{2,2} + Y_{2,3})/3$, and $\bar{Y}_3 = (Y_{3,1} + Y_{3,2} + Y_{3,3})/3$,.

8. In fitting regression model in the first question, it is found that observation (X_i, Y_i) fell directly on the fitted regression line (i.e. $Y_i = \hat{Y}_i$). If this observation is deleted, would the least square regression line fitted to the remaining n-1 cases be changed? [hint: what is the contribution of the observation to the function Q]