Tutorial 5

1. Suppose we have $n = 10$ observations (X_i, Y_i) and fit the data with model

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

with $\varepsilon_i, i = 1, ..., 10$ are IID $N(0, \sigma^2)$. We have the following calculations.

$$\bar{X} = 0.5669, \quad \bar{Y} = 0.9624, \quad \sum_{i=1}^{n} Y_i^2 = 10.2695,$$

$$\sum_{i=1}^{n} X_i^2 = 4.0169, \quad \sum_{i=1}^{n} X_i Y_i = 6.2841.$$

   (a) Write down the estimated model

   (b) Test $H_0 : \beta_1 = 1$ with $\alpha = 0.05$

   (c) For a new $X = 1$, find the 95% CI for its expected response

   (d) For a new $X = 0.5$, find the 95% prediction interval for its possible response

2. For the least square estimator b_0, b_1 of simple linear regression model, find $Cov(b_0, b_1)$

3. Suppose $A : m \times n$ is a constant matrix and $Y : n \times 1$, is a random vector. Then

$$\mathbf{Var}(AY) = A\mathbf{Var}(Y)A'$$

Please give your proof for $m = 2, n = 3$.

4. For multiple linear regression, the normal equations are

$$\sum_{i=1}^{n} e_i = 0$$
$$\sum_{i=1}^{n} e_i X_{i1} = 0$$
$$\vdots$$
$$\sum_{i=1}^{n} e_i X_{ip} = 0$$

Prove that

$$\sum_{i=1}^{n} \hat{Y}_i e_i = 0$$

5. Consider the simple regression model $Y = \beta_0 + \beta_1 X + \varepsilon$. We have n observations (X_i, Y_i) with sample correlation coefficient r_{XY}. Standardize

$$\tilde{X}_i = \frac{X_i - \bar{X}}{\sqrt{\sum_{j=1}^n (X_j - \bar{X})^2/(n-1)}}, \qquad \tilde{Y}_i = \frac{Y_i - \bar{Y}}{\sqrt{\sum_{j=1}^n (Y_j - \bar{Y})^2/(n-1)}}$$

What is the estimated model for $\tilde{Y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 \tilde{X}_i + \varepsilon_i$?