Indian Institute of Technology Guwahati
Statistical Inference (MA682)
Problem Set 07

1. Consider the simple linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$, with usual assumptions on ε. Show that

$$Cov\left(\widehat{\beta}_0,\, \widehat{\beta}_1\right) = -\frac{\overline{x}\sigma^2}{S_{xx}} \quad \text{and} \quad Cov\left(\overline{y},\, \widehat{\beta}_0\right) = \frac{\sigma^2}{n}.$$

2. Consider the simple linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$ with usual assumptions on ε. Show that

$$E\left(MS_R\right) = \sigma^2 + \beta_1^2 S_{xx} \quad \text{and} \quad E\left(MS_{Res}\right) = \sigma^2.$$

3. Suppose that we have fit a simple linear regression model $\widehat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x_1$, but the response is affected by a second variable x_2 such that the true regression is $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$. Is the least square estimator $\widehat{\beta}_1$ in the simple linear regression model unbiased?

4. Consider the simple linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$, where ε's are independent and identicall $N(0,\, \sigma^2)$ random variables. Find the MLEs of β_0, β_1, and σ^2. Is the MLE of σ^2 UE?

5. Suppose that we are fitting a straight line and wish to make standard error of the slope as small as possible. Suppose that the region of interest for x is $-1 \leq x \leq 1$. Where should the observations $x_1,\, x_2,\, \ldots,\, x_n$ be taken?

6. Consider the simple linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$, with usual assumptions on ε. Also assume that β_0 is known. Find the LSE of β_1 and its' variance.

7. Consider the multiple linear regression model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon.$$

   Using the procedure for testing a general linear hypothesis, show how to test

   (a) $H_0 : \beta_1 = \beta_2,\, \beta_3 = \beta_4$.
   (b) $H_0 : \beta_1 - 2\beta_2 = 4\beta_3,\, \beta_1 + 2\beta_2 = 0$.

8. Show that $Var(\widehat{\boldsymbol{y}}) = \sigma^2 H$.

9. Consider the multiple linear regression model $\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. Show that LSE of $\boldsymbol{\beta}$ can be written as $\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta} + R\boldsymbol{\varepsilon}$, where $R = \left(X'X\right)^{-1} X'$.

10. Prove that R^2 is the square of the correlation between $\boldsymbol{y}$ and $\widehat{\boldsymbol{y}}$.

11. Consider the linear regression $y = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p + \epsilon$, with usual assumptions on ϵ. Assuming that X is a full column rank matrix, show that

$$\sum_{i=1}^{n} Var(\widehat{y}_i) = (p+1)\sigma^2.$$