Problem 3. In the two-variable model

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i, i = 1, \cdots, 11$$

Suppose that $x_1'x_1 = 2, x_2'x_2 = 2, x_1'x_2 = 1, x_1'y = 2, x_2'y = 1, y'y = 7/3$ where x_1, x_2 and y are the column vectors with typical elements x_{1i}, x_{2i} and y_i respectively.
Assume $\epsilon_i \sim i.i.d.N(0, \sigma_\epsilon^2)$. Suppose you would like to make out-of-sample predictions about the left-hand-side(dependent) variable for two hypothetical observations with the following characteristics:

Obs.	x_1	x_2
12	5	-2
13	3	-7

1. Construct 80% prediction intervals for the dependent variable y for observations 12 and 13.

$$\beta = (X'X)^{-1}X'y, \quad X = \begin{bmatrix} x_1 & x_2 \end{bmatrix}, \quad X' = \begin{bmatrix} x_1' \\ x_2' \end{bmatrix}$$

$$\begin{aligned} X'X &= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix} = \begin{bmatrix} x_1'x_1 & x_1'x_2 \\ x_2'x_1 & x_2'x_2 \end{bmatrix} \\ (X'X)^{-1} &= \frac{1}{(x_1'x_1)(x_2'x_2) - (x_1'x_2)(x_2'x_1)} \begin{bmatrix} x_2'x_2 & -x_1'x_2 \\ -x_2'x_1 & x_1'x_1 \end{bmatrix} \\ X'y &= \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} y = \begin{bmatrix} x_1'y \\ x_2'y \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \beta &= \frac{1}{(x_1'x_1)(x_2'x_2) - (x_1'x_2)(x_2'x_1)} \begin{bmatrix} x_2'x_2 & -x_1'x_2 \\ -x_2'x_1 & x_1'x_1 \end{bmatrix} \begin{bmatrix} x_1'y \\ x_2'y \end{bmatrix} \\ &= \frac{1}{(x_1'x_1)(x_2'x_2) - (x_1'x_2)(x_2'x_1)} \begin{bmatrix} (x_2'x_2)(x_1'y) + (-x_1'x_2)(x_2'y) \\ (-x_2'x_1)(x_1'y) + (x_1'x_1)(x_2'y) \end{bmatrix} \\ &= \frac{1}{(2)(2) - (1)(1)} \begin{bmatrix} (2)(2) + (-1)(1) \\ (-1)(2) + (2)(1) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \end{aligned}$$

$$y_i = x_{1i} + \epsilon_i, \quad \hat{y}_i = x_{1i} \quad \because \beta_1 = 1, \beta_2 = 0$$

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{n-2} \sum \hat{\epsilon}_i^2 = \frac{1}{n-2} \sum (y_i - x_{1i})^2 \\ &= \frac{1}{n-2} (\sum y_i^2 - 2 \sum y_i x_{1i} + \sum x_{1i}^2) \\ &= \frac{1}{9} (\frac{7}{3} - 2(2) + 2) = \frac{1}{27} \end{aligned}$$

A 100(1-α)% Prediction Interval for $Y = \beta_0 + \beta_1 x + \epsilon$ when $x = x^*$

$$\hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2} S \sqrt{1 + \frac{1}{n} + \frac{(x^* - \overline{x})^2}{S_{xx}}}$$

where the tabulated $t_{\alpha/2}$ is based on $n-2$ df.

$$\hat{\beta}_0 = 0, \hat{\beta}_1 = 1, t_{0.1}(9) = 1.383, S = \hat{\sigma} = 0.19245, \overline{x} = 0, S_{xx} = x'x$$

at Obs. 12, $x^* = 5$, Prediction Interval $= 5 \pm (1.383)(0.19245)\sqrt{1 + \frac{1}{15} + \frac{5^2}{2}} = 5 \pm 0.9803 = (4.020, 5.980)$

at Obs. 13, $x^* = 3$, Prediction Interval $= 3 \pm (1.383)(0.19245)\sqrt{1 + \frac{1}{15} + \frac{3^2}{2}} = 3 \pm 0.6280 = (2.372, 3.628)$

2. Construct 80% prediction intervals for the expected value of y_{12} and y_{13}.

A 100(1-α)% Confidence Interval for $E(Y) = \beta_0 + \beta_1 x^*$ when $x = x^*$

$$\hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2} S \sqrt{\frac{1}{n} + \frac{(x^* - \overline{x})^2}{S_{xx}}}$$

where the tabulated $t_{\alpha/2}$ is based on $n - 2$ df.

$$\hat{\beta}_0 = 0, \hat{\beta}_1 = 1, t_{0.1}(9) = 1.383, S = \hat{\sigma} = 0.19245, \overline{x} = 0, S_{xx} = x'x$$

at Obs. 12, $x^* = 5$, Prediction Interval $= 5 \pm (1.383)(0.19245)\sqrt{\frac{1}{15} + \frac{5^2}{2}} = 5 \pm 0.9435 = (4.0565, 5.9435)$

at Obs. 13, $x^* = 3$, Prediction Interval $= 3 \pm (1.383)(0.19245)\sqrt{\frac{1}{15} + \frac{3^2}{2}} = 3 \pm 0.5688 = (2.4312, 3.5688)$

3. Do the answers above differ? Why?

$$Var(error) = Var(Y^* - \widehat{Y^*}) = Var(Y^*) + Var(\widehat{Y^*}) - 2Cov(Y^*, \widehat{Y^*})$$
$$Cov(Y^*, \widehat{Y^*}) = 0 \quad \because Y^* \, and \, \widehat{Y^*} \, are \, independent.$$

$$\begin{aligned} Var(error) &= Var(Y^*) + Var(\widehat{Y^*}) = \sigma^2 + Var(\widehat{\beta_0} + \widehat{\beta_1} x^*) \\ &= \sigma^2 + (\frac{1}{n} + \frac{(x^* - \overline{x})^2}{S_{xx}})\sigma^2 \\ &= \sigma^2 \left[1 + \frac{1}{n} + \frac{(x^* - \overline{x})^2}{S_{xx}}\right] \end{aligned}$$

$\therefore$ the variance of particular value of variable Y is bigger than one of the expected value of variable Y.