Linear Statistical Models — Quiz 02 — Full Name ______________________________

1. Suppose we have covariate data $x_1, \ldots, x_n$ and response data $y_1, \ldots, y_n$.

(i) (3) Write down the simple linear statistical model for y_i, making sure to define each and every term you specify, including all assumptions. Use b_0 and b_1 as notation for the regression coefficients.

$$Y_i = b_0 + b_1 x_i + \varepsilon_i$$

$$\varepsilon_i \overset{ind}{\sim} N(0, \sigma^2)$$

(ii) (3) Write out the design matrix X and the response vector $\boldsymbol{y}$ for the simple linear model.

$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

(iii) (4) Calculate $X^T X$ and verify that $(X^T X)^{-1} = \frac{1}{-n(\overline{x})^2 + \sum_{i=1}^n x_i^2} \begin{bmatrix} \frac{1}{n}\sum_{i=1}^n x_i^2 & -\overline{x} \\ -\overline{x} & 1 \end{bmatrix}$

(hint: multiply)

$$X^T X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & \sum x_i^2 \end{bmatrix}$$

$$(X^T X)(X^T X)^{-1} = \frac{1}{-n(\bar{x})^2 + \sum x_i^2} \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & \sum x_i^2 \end{bmatrix} \begin{bmatrix} \frac{1}{n}\sum x_i^2 & -\bar{x} \\ -\bar{x} & 1 \end{bmatrix}$$

$$= \frac{1}{-n(\bar{x})^2 + \sum x_i^2} \begin{bmatrix} \sum x_i^2 - n(\bar{x})^2 & -n\bar{x} + n\bar{x} \\ \bar{x}\sum x_i^2 - \bar{x}\sum x_i^2 & -n(\bar{x})^2 + \sum x_i^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

2. Consider the linear model $\boldsymbol{Y} = X\boldsymbol{b} + \boldsymbol{\varepsilon}$, $\varepsilon_i \overset{ind}{\sim} N(0, \sigma^2)$

(i) (2) Derive $E(\boldsymbol{Y})$.

$$E(\underline{Y}) = E(X\underline{b} + \underline{\varepsilon}) = Xb + E(\underline{\epsilon}) = Xb$$

(ii) (3) Derive $E(\widehat{\boldsymbol{B}}) = E((X^T X)^{-1} X^T \boldsymbol{Y})$.

$$E(\underline{\hat{B}}) = E((X^TX)^{-1}X^T\underline{Y}) = (X^TX)^{-1}X^TE(\underline{Y})$$
$$= (X^TX)^{-1}X^TX\underline{b} = \underline{b}$$

(iii) (2) Given that $\mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2 I$, derive $\mathrm{Cov}(\boldsymbol{Y})$.

$$\mathrm{Cov}(\underline{Y}) = \mathrm{Cov}(X\underline{b} + \underline{\varepsilon}) = \mathrm{Cov}(\underline{\varepsilon}) = \sigma^2 \underline{I}$$

(iv) (3) Derive $\mathrm{Cov}(\widehat{\boldsymbol{B}})$.

$$\mathrm{Cov}(\underline{\hat{B}}) = \mathrm{Cov}((X^TX)^{-1}X^T\underline{Y}) = (X^TX)^{-1}X^T\mathrm{Cov}(\underline{Y})X(X^TX)^{-1}$$
$$= (X^TX)^{-1}X^T\sigma^2 I X(X^TX)^{-1}$$
$$= \sigma^2(X^TX)^{-1}X^TX(X^TX)^{-1}$$
$$= \sigma^2(X^TX)^{-1}$$

3. Consider the same model as above, but with the following design matrix: $X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$

(i) (3) Write down the expected values of the four responses and describe the pattern in words.

$E(Y_1) = b_0 + b_1$

$E(Y_2) = b_0 + b_1$

$E(Y_3) = b_0$

$E(Y_4) = b_0$

first two equal $b_0 + b_1$
second two are b_0

(ii) (2) Calculate $(X^T X)$ and $(X^T X)^{-1}$. (check your work!)

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \checkmark$$

$$(X^T X)^{-1} = \frac{1}{8-4} \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}$$

(iii) (2) Calculate the projection matrix for X.

$$X(X^T X)^{-1} X^T = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

(iv) (3) Given response data y_1, y_2, y_3, y_4, calculate the fitted values, and explain how they are related to your answer in part (i).

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} (y_1+y_2)/2 \\ (y_1+y_2)/2 \\ (y_3+y_4)/2 \\ (y_3+y_4)/2 \end{bmatrix}$$

(v) (+1) Find a matrix U with orthonormal columns, such that UU^T is the projection matrix for X.

$$U = \begin{bmatrix} \sqrt{1/2} & 0 \\ \sqrt{1/2} & 0 \\ 0 & \sqrt{1/2} \\ 0 & \sqrt{1/2} \end{bmatrix}$$