Sheil Sarda
ESE402 Assignment 4
Prof. Hassani
Nov. 17, 2021

Problem 1

PROBLEM 1 A

Given: n data points : $\{(x_i, y_i)\}_{i=1}^{n} \sim y_i = x_i\beta_1 + \beta_0 + \epsilon$

$\epsilon \sim dist(0, \sigma^2)$

RSS estimates $\hat{\beta}_0, \hat{\beta}_1$: $\hat{\mu}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$

$\mu_0 = \beta_0 + \beta_1 x_0$

To find: variance of $\hat{\mu}_0$

We know that β_0 & β_1 have the following form: $\hat{\beta}_0 = \bar{y} - \beta_1 \bar{x}$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\Rightarrow \hat{\mu}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$$
$$= \bar{y} - \hat{\beta}_1 \bar{x} + \hat{\beta}_1 x_0$$
$$= \bar{y} + \hat{\beta}_1 (x_0 - \bar{x})$$

$$\Rightarrow Var(\hat{\mu}_0) = Var(\bar{y} + \hat{\beta}_1 (x_0 - \bar{x}))$$
$$= Var(\bar{y}) + (x_0 - \bar{x})^2 Var(\hat{\beta}_1) \quad - (1)$$

($\because$ Applying Linearity of variance because $\bar{y}$, ~~$x_0, \bar{x}$ are all constants so independent~~ from $\hat{\beta}_1$ is uncorrelated & x_0 & $\bar{x}$ are provided #s so they are constant)

According Theorem 14.B in Rice's text book:

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \quad - (2)$$

We also know $Var(\bar{y}) = \dfrac{Var(\cancel{\epsilon})}{n} = \dfrac{\sigma^2}{n} \quad - (3)$

Substituting (2) & (3) into (1):

$$Var(\hat{\mu}_0) = \frac{\sigma^2}{n} + (x_0 - \bar{x})^2 \frac{\sigma^2}{\sum_{i=1}^{n} (x_i - \bar{x})}$$
$$= \sigma^2 \left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})} \right)$$

PROBLEM 1B

We know : $Var(\hat{\mu}_0) = \dfrac{\sigma^2}{n} + \dfrac{\sigma^2(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})}$

we can assign variables :

$$\frac{\sigma^2}{n} = \alpha_0 \quad , \quad \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})} = \alpha_1 \quad , \quad (x_0 - \bar{x}) = z$$

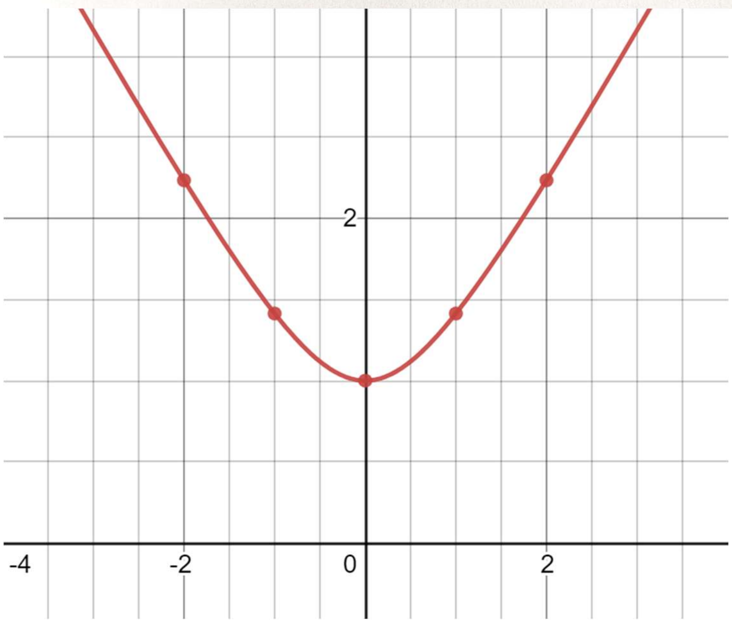
$\Rightarrow Var(\hat{\mu}_0) = \alpha_0 + \alpha_1 z^2$

$\Rightarrow$ std. dev$(\hat{\mu}_0) = \sqrt{\alpha_0 + \alpha_1 z^2}$

Graph of $f(z) = \sqrt{\alpha_0 + \alpha_1 z^2}$ where $\alpha_0, \alpha_1 \geq 0$

For $\alpha_0 = \alpha_1 = 1$, we get the following graph :

(Digital)



PROBLEM 1C

We know standard deviation of $\hat{\mu}_0 = \sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2 (x_0 - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})}}$

We also know that the CI can be written as : $\hat{\mu}_0 \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ ①

Alternatively : $\hat{\mu}_0 \pm s_{\hat{\mu}_0} t_{n-2}\left(\frac{\alpha}{2}\right)$ ②

Substituting $\sigma = s_{\hat{\mu}_0} = \sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2 (x_0 - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})}}$,

we have the 95% confidence interval $(\alpha = 0.05)$

Problem 2

PROBLEM 2

Given: $X \sim N(0,1)$; $E \sim N(0,1)$; $Y = X + \beta E$

Prove:

$$r_{XY} = \frac{1}{\sqrt{\beta^2+1}} \quad \text{where} \quad r_{XY} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}$$

$$r_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}$$

$$= \frac{\text{cov}(X, X+\beta E)}{\sigma_X \sigma_{X+\beta E}} = \frac{\text{cov}(X,X) + \text{cov}(X, \beta E)}{\sigma_X \sqrt{\text{Var}(X+\beta E)}}$$

$$= \frac{\text{Cov}(X,X) + \beta \text{cov}(X,E)}{1 \sqrt{\text{Var}(X) + \beta^2 \text{Var}(E)}} = \frac{\text{Var}(X) + 0}{\sqrt{1+\beta^2}}$$

$\Big(\because$ $\text{cov}(X,X) = \text{Var}(X)$

$\text{Var}(X) = \text{Var}(E) = 1$

$\text{cov}(X,E) = 0 \Big)$

$$= \frac{1}{\sqrt{1+\beta^2}}$$

Problem 3

PROBLEM 3

Given: ~~Given~~ $\{x_i \in \mathbb{R}, y_i \in \mathbb{R}\}_{i=1}^{n}$

$y = a + bx$; $x = c + dy$ w/ least squares

Prove: $bd \leq 1$

Hint: $|\text{Cov}(X,Y)|^2 \leq \text{Var}(X)\,\text{Var}(Y)$

$$b = \frac{\text{Cov}(X,Y)}{\text{Var}(X)} \; ; \; d = \frac{\text{Cov}(Y,X)}{\text{Var}(Y)}$$

$$\Rightarrow bd = \frac{(\text{cov}(X,Y))^2}{\text{Var}(X)\,\text{Var}(Y)} \leq \frac{\text{Var}(X)\,\text{Var}(Y)}{\text{Var}(X)\,\text{Var}(Y)} = 1$$

$$\Rightarrow bd \leq 1$$

explanation: $bd = 1$ when $|\text{cov}(X,Y)|^2 = \text{Var}(X)\,\text{Var}(Y) \Rightarrow$ correlation of $+1$

i.e. X & Y are moving in the same direction together.

Problem 4

PROBLEM 4

Given: $Y \in \mathbb{R}^n$ & $X_1, X_2 \in \mathbb{R}^n$

$X_3 = X_1 + X_2$; use multiple regression to predict Y from X_1, X_2 & X_3

Hints: $A_{n,n}$ is invertible $\Leftrightarrow$ $\text{Rank}(A) = n$

$\text{Rank}(AB) \leq \min(\text{Rank}(A), \text{Rank}(B))$

Prove: Why this will not work.

$$Y = \underbrace{\cancel{\beta X}}_{X\beta} + E \quad : \quad Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}, \ \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, \ X = \begin{bmatrix} 1 & X_{1,1} & X_{1,2} & X_{1,3} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n,1} & X_{n,2} & X_{n,3} \end{bmatrix}$$

Looking at X, its column rank is 3 $\because$ the last column is a linear combination of other columns in the matrix. Thus, it does not have full column rank & cannot be invertible.

Least squares estimator for β:

$$Y_i \mid X_i = x_i \sim N\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}, \sigma^2\right) \qquad (\because Y = X\beta + \epsilon)$$

Least squares error: $(y - X\beta)^T (y - X\beta)$

$\Rightarrow$ ~~optimal~~ minimum error is where $X^T X\beta = X^T y \quad \Rightarrow \quad \beta = (X^T X)^{-1} X^T Y$

However, given rank of X = rank of $X^T = 3$, we get:

$\text{rank}(X^T X) = \min\{3, 3\} = 3$

$\Rightarrow$ $X^T X$ is not invertible because it does not have full rank

$\Rightarrow$ we cannot find β because $(X^T X)^{-1}$ is required to find this