HW 5

Exercise 1

$Y = X\beta + \epsilon \qquad E[\epsilon] = 0,\ var(\epsilon) = \sigma^2 I_n$

know $E[\hat{\beta}] = \beta$

How does $b = AY$ which is unbiased compare to $\hat{\beta}$?

show: $q'(var(b) - var(\hat{\beta}))q \geq 0 \iff q' var(b) q \geq q' var(\hat{\beta}) q$

EX $q = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \rightarrow var(b_1) \geq var(\hat{\beta}_1)$ Ex. $q = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \rightarrow var(b_j) \geq var(\hat{\beta}_j)$

$AX = I$, why?

$E[b] = E[AY]$

$E[b] = AE[Y]$

$\beta = AX\beta \Rightarrow AX = I$

$var(b) = var(AY)$

$= A\,var(Y) A'$

$= A \begin{bmatrix} var(y_i) & cov(y_i, y_j) \\ & \ddots \\ & var(y_n) \end{bmatrix} A' = A \begin{bmatrix} \sigma^2 & 0 \\ & \ddots \\ 0 & \sigma^2 \end{bmatrix} A' = A(\sigma^2 I) A'$

know: $var(\hat{\beta}) = \sigma^2 (X'X)^{-1}$

$= \sigma^2 AA'$

$q'[\sigma^2[AA' - (X'X)^{-1}]]q = \sigma^2 q' [AA' - (X'X)^{-1}] q$

$= \sigma^2 q' [AA' - \underbrace{AX}_{I_n} (X'X)^{-1} \underbrace{X'A'}_{I_n}] q$

$= \sigma^2 q' [A(I - X(X'X)^{-1}X')A'] q$

$= \sigma^2 w' [I - H] w$ where $w = A'q$

Note: $I - H$ is idempotent

proof: $(I-H)(I-H) = I - H + H + H^2$

$= I - H$

Note: $I-H$ is symmetric

Proof: $(I-H)' = (I'-H') = I-H$

$$= \sigma^2 w'[I-H]^2 w \quad \text{because } I-H \text{ is idempotent}$$
$$= \sigma^2 w'(I-H)'(I-H)w \quad \text{because } I-H \text{ is symmetric}$$
$$= \sigma^2 v'v \quad \text{where } v = w(I-H)$$
$$= \sigma^2 \sum_i v_i^2 \geq 0$$

Exercise 2

$$\hat\beta = \begin{bmatrix} \widehat{\text{intercept}} \\ \hat\beta_{(0)} \end{bmatrix} \in \mathbb{R}^{k+1} = \overbrace{\begin{bmatrix} n & 1'X_{(0)} \\ X_{(0)}'1_{n\times1} & X_{(0)}'X_{(0)} \end{bmatrix}^{-1}}^{A^{-1}} \begin{bmatrix} 1'Y \\ X_{(0)}'Y \end{bmatrix}$$

$X = \begin{bmatrix} 1_{n\times1} & X_{(0)} \end{bmatrix}$
where $X_{(0)} \in \mathbb{R}^{n\times k}$

$\hat\beta_{(0)}$: coefficient estimates $\in \mathbb{R}^k$
without intercept

$$\hat\beta_{(0)} = (X_{(0)}^{*\prime} X_{(0)}^{*})^{-1} X_{(0)}^{*\prime} Y^*$$
$$X^*_{(0)} = (I - \tfrac{1}{n} 11')X_{(0)}$$

$$X_{(0)}^* = (I - \tfrac{1}{n}11')X_{(0)} = X_{(0)} - \overline{X_{(0)}} \qquad (\overline{X}_{(0)})_{ij} = \tfrac{1}{n}\Sigma(X_{(0)})_{ij}$$

$$Y_{(0)}^* = (I - \tfrac{1}{n}11')Y = Y - \begin{bmatrix} \bar Y \\ \bar Y \\ \vdots \\ \bar Y \end{bmatrix} \qquad \bar y = \tfrac{1}{n}\Sigma y_i$$

$$\begin{bmatrix} \overline{X_{(0)}}_1 & \overline{X_{(0)}}_2 & \cdots & \overline{X_{(0)}}_k \\ \overline{X_{(0)}}_1 & \overline{X_{(0)}}_2 & & \overline{X_{(0)}}_k \\ \vdots & \vdots & & \vdots \\ \overline{X_{(0)}}_1 & \overline{X_{(0)}}_2 & & \overline{X_{(0)}}_k \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} + B_{12}B_{22}^{-1}B_{21} & -B_{12}B_{22}^{-1} \\ -B_{22}^{-1}B_{21} & B_{22}^{-1} \end{bmatrix}$$

$$B_{22} = A_{22} - A_{21}A_{11}^{-1}A_{12}$$
$$B_{12} = A_{11}^{-1}A_{12}$$
$$B_{21} = A_{21}A_{11}^{-1}$$

$$A_{11} = n \in \mathbb{R}^{1\times1}$$
$$A_{12} = 1'X_{(0)} \in \mathbb{R}^{1\times k}$$
$$A_{21} = X_{(0)}'1 \in \mathbb{R}^{k\times1}$$
$$A_{22} = X_{(0)}'X_{(0)}$$

$\longrightarrow$

$$B_{22} = X_{(0)}'X_{(0)} - X_{(0)}'1\tfrac{1}{n}1'X_{(0)}$$
$$B_{12} = \tfrac{1}{n}1'X_{(0)}$$
$$B_{21} = X_{(0)}'1\tfrac{1}{n}$$

$$\begin{bmatrix} \hat\beta_0 \\ \hat\beta_{(0)} \end{bmatrix} = \begin{bmatrix} [A_{11}^{-1} + \beta_{12}\beta_{22}^{-1}\beta_{21}]1'Y - \beta_{12}\beta_{22}^{-1}X_{(0)}'Y \\ -\beta_{22}^{-1}\beta_{21}1'Y + \beta_{22}^{-1}X_{(0)}'Y \end{bmatrix}$$

Note: $(I - \frac{1}{n} 11')(I - \frac{1}{n} 11') = (II - I\frac{1}{n}11' - \frac{1}{n}11'I + \frac{1}{n^2} 11'11')$ Note: $1'1 = n$

$= I - \frac{2}{n}11' + \frac{1}{n^2} 1n1'$

$= I - \frac{2}{n}11' + \frac{n}{n^2} 11'$

$= I - \frac{2}{n}11' + \frac{1}{n} 11'$

$= I - \frac{1}{n}11'$

$\Rightarrow (I - \frac{1}{n}11')$ is idempotent

Note: $(I - \frac{1}{n}11')' = I' - \frac{1}{n}11' = I - \frac{1}{n}11'$

$\Rightarrow (I - \frac{1}{n}11')$ is symmetric

$B_{22} = X_{(0)}'X_{(0)} - X_{(0)}'1\frac{1}{n}1'X_{(0)}$

$= X_{(0)}'(I - 1\frac{1}{n}1')X_{(0)}$

$= X_{(0)}'(I - \frac{1}{n}11')X_{(0)} = X_{(0)}'(I - \frac{1}{n}11')(I - \frac{1}{n}11')X_{(0)}$

$= X_{(0)}'(I - \frac{1}{n}11')'(I - \frac{1}{n}11')X_{(0)}$

$= [(I - \frac{1}{n}11')X_{(0)}]'[(I - \frac{1}{n}11')X_{(0)}]$ Note: $(I - \frac{1}{n}11')X_{(0)} = X_{(0)}^*$

$= X_{(0)}^{*\prime} X_{(0)}^*$

$\Rightarrow B_{22}^{-1} = (X_{(0)}^{*\prime} X_{(0)}^*)^{-1}$

$B_{21} = X_{(0)}'1\frac{1}{n} = X_{(0)}'\frac{1}{n}1$

$\Rightarrow B_{22}^{-1}B_{21} = (X_{(0)}^{*\prime} X_{(0)}^*)^{-1} X_{(0)}'\frac{1}{n}1$

$\hat{\beta}_{(0)} = -B_{22}^{-1}B_{21}1'Y + B_{22}^{-1}X_{(0)}'Y$

$= -(X_{(0)}^{*\prime} X_{(0)}^*)^{-1} X_{(0)}'\frac{1}{n}11'Y + (X_{(0)}^{*\prime} X_{(0)}^*)^{-1} X_{(0)}'Y$

$= (X_{(0)}^{*\prime} X_{(0)}^*)^{-1} X_{(0)}'Y - (X_{(0)}^{*\prime} X_{(0)}^*)^{-1} X_{(0)}'\frac{1}{n}11'Y$

$= (X_{(0)}^{*\prime} X_{(0)}^*)^{-1} [X_{(0)}'Y - X_{(0)}'\frac{1}{n}11'Y]$

$= (X_{(0)}^{*\prime} X_{(0)}^*)^{-1} [X_{(0)}'(I - \frac{1}{n}11')Y]$

$= (X_{(0)}^{*\prime} X_{(0)}^*)^{-1} [X_{(0)}'(I - \frac{1}{n}11')(I - \frac{1}{n}11')Y]$ Note: $Y_{(0)}^* = (I - \frac{1}{n}11')Y$

$= (X_{(0)}^{*\prime} X_{(0)}^*)^{-1} [X_{(0)}'(I - \frac{1}{n}11')'Y_{(0)}^*]$

$= (X_{(0)}^{*\prime} X_{(0)}^*)^{-1} [(I - \frac{1}{n}11')X_{(0)}]'Y_{(0)}^*$ Note: $X_{(0)}^* = (I - \frac{1}{n}11')X_{(0)}$

$= (X_{(0)}^{*\prime} X_{(0)}^*)^{-1} X_{(0)}^{*\prime} Y_{(0)}^*$

$$\boxed{\hat{\beta}_{(0)} = (X_{(0)}^{*\prime} X_{(0)}^*)^{-1} X_{(0)}^{*\prime} Y_{(0)}^*}$$

Exercise 3

$Y = X\beta + \epsilon \quad E(\epsilon) = 0 \quad var(\epsilon) = \sigma^2 I$

$$X = \overset{3\times 4}{\begin{bmatrix} 1 & X_{11} & \cdots & X_{1k} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & & X_{nk} \end{bmatrix}} \qquad X' = \overset{4\times 3}{\begin{bmatrix} 1 & \cdots & 1 \\ X_{11} & \cdots & X_{n1} \\ \vdots & & \\ X_{1k} & \cdots & X_{nk} \end{bmatrix}}$$

Note: $\hat{\beta} = (X'X)^{-1} X'Y$

$$X'X = \begin{bmatrix} 1'1 & 1'X_{\cdot 1} & \cdots & 1'_{\cdot k} \\ X_{\cdot 1}'1 & X_{\cdot 1}'X_{\cdot 1} & \cdots & X_{\cdot 1}'X_{\cdot k} \\ \vdots & & & \\ X_{\cdot k}'1 & X_{\cdot k}'X_{\cdot 1} & \cdots & X_{\cdot k}'X_{\cdot k} \end{bmatrix}$$

If the columns of X are orthogonal then the product of any two different columns will be 0. This means X'X must be a diagonal matrix.

$$\Rightarrow \quad X'X = \begin{bmatrix} 1'1 & 0 & \cdots & 0 \\ 0 & X_{\cdot 1}'X_{\cdot 1} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & X_{\cdot k}'X_{\cdot k} \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & \cdots & 1 \\ X_{11} & \cdots & X_{n1} \\ \vdots & & \\ X_{1k} & \cdots & X_{nk} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \Sigma Y_i \\ \Sigma X_{i1} Y_i \\ \vdots \\ \Sigma X_{ik} Y_i \end{bmatrix}$$

$$(X'X)^{-1} = \begin{bmatrix} \frac{1}{n} & 0 & \cdots & 0 \\ 0 & \frac{1}{\Sigma X_{i1}^2} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \frac{1}{\Sigma X_{ik}^2} \end{bmatrix}$$

Because X'X is a diagonal matrix

$$\Rightarrow \quad \hat{\beta} = \begin{bmatrix} \frac{1}{n} & 0 & \cdots & 0 \\ 0 & \frac{1}{\Sigma X_{i1}^2} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \frac{1}{\Sigma X_{ik}^2} \end{bmatrix} \begin{bmatrix} \Sigma Y_i \\ \Sigma X_{i1} Y_i \\ \vdots \\ \Sigma X_{ik} Y_i \end{bmatrix} = \begin{bmatrix} \bar{Y} \\ \Sigma X_{i1}Y_i / \Sigma X_{i1}^2 \\ \vdots \\ \Sigma Y_{ik}Y_i / \Sigma X_{ik}^2 \end{bmatrix} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$$

show the estimates are the same as the estimates from the regression of Y separately on each column.

$\hat{\beta}_0 = (X_{\cdot 1}'X_{\cdot 1})^{-1} X_{\cdot 1}'Y = (1'1)^{-1} 1'Y = \frac{1}{n}\Sigma Y_i = \bar{Y}$

$\hat{\beta}_1 = (X_{\cdot 2}'X_{\cdot 2})^{-1} X_{\cdot 2}'Y = \Sigma X_{i1}Y_i / \Sigma X_{i1}^2$

$\vdots$

$\hat{\beta}_k = (X_{\cdot k}'X_{\cdot k})^{-1} X_{\cdot k}Y = \Sigma Y_{ik}Y_i / \Sigma X_{ik}^2$

Exercise 4

$$R^2 = 1 - \frac{SSE}{SST} \qquad SSE = e'e \qquad Se^2 = \frac{e'e}{n-k-1} \qquad \begin{matrix} n = 10 \\ k = 2 \end{matrix}$$

$$SSE = Se^2 (n-k-1) = Se^2 (7) = (100)(7) = 700$$

$$SST = 5000$$

$$R^2 = 1 - \frac{700}{5000} = 0.86$$

Exercise 5

$Y = X\beta + \epsilon$, $\epsilon \sim N(0, \sigma^2 I_n)$

Note: $\epsilon \sim N(0, \sigma^2 I_n)$
$\Rightarrow Y \sim N(X\beta, \sigma^2 I_n)$

what is the distribution of:

1. $\hat{\beta} \sim N(\beta, \sigma^2 (X'X)^{-1})$
2. $\hat{Y} = HY \sim N(HX\beta, \sigma^2 HH') \rightarrow \sim N(X\beta, \sigma^2 H)$
3. $e = (I-H)Y \sim N((I-H)X\beta, \sigma^2 (I-H)) \rightarrow \sim N(0, \sigma^2 (I-H))$

Note: $var(HY) = H var(Y) H'$
$= \sigma^2 HH' = \sigma^2 H$

Note: $HH' = H$

$cov(\hat{Y}, \hat{\beta}) = cov(X\hat{\beta}, \hat{\beta}) = X var(\hat{\beta}) = \sigma^2 X (X'X)^{-1} \neq 0$ necessarily

$\Rightarrow \hat{Y}, \hat{\beta}$ are dependent

$$\begin{aligned} cov(\hat{Y}, e) = cov(HY, (I-H)Y) &= H\, var(Y)\, (I-H)' \\ &= \sigma^2 H (I-H) \\ &= \sigma^2 (H - H^2) \\ &= 0 \in \mathbb{R}^{n \times n} \end{aligned}$$

Note: $HH = H$

$\Rightarrow \hat{Y}, e$ are independent

$$\begin{aligned} cov(\hat{\beta}, e) = cov(\underbrace{(X'X)^{-1}X'}_{A}Y, (I-H)Y) &= A\, var(Y)\, (I-H)' \\ &= \sigma^2 A (I-H) \\ &= \sigma^2 (X'X)^{-1} X' (I-H) \\ &= \sigma^2 (X'X)^{-1} X' - X'H \\ &= \sigma^2 (X'X)^{-1} (X - H'X)' \\ &= \sigma^2 (X'X)^{-1} (X - HX)' \\ &= 0 \in \mathbb{R}^{(k+1) \times n} \end{aligned}$$

Note: $HX = X$

$\Rightarrow \hat{\beta}, e$ are independent

Exercise 6

$Y = X\beta + \epsilon$, $E[\epsilon] = 0$, $var(\epsilon) = \sigma^2 I \rightarrow \epsilon \sim N(0, \sigma^2 I)$

Assuming there is no intercept in the model then matrix X will have the dimensions $n \times K$ rather than $n \times (K+1)$

$$Y \sim N(X\beta, \sigma^2 I) \rightarrow f(y) = (2\pi)^{-n/2} |\sigma^2 I|^{-1/2} e^{-1/2 (Y - X\beta)' \Sigma^{-1} (Y - X\beta)}$$

$$L = f(y) = (2\pi)^{-n/2} ((\sigma^2)^n)^{-1/2} e^{-1/2 (Y - X\beta)(\sigma^2 I)^{-1}(Y - X\beta)}$$
$$= (2\pi\sigma^2)^{-n/2} e^{(Y - X\beta)'(Y - X\beta)}$$

$$\ln(L) = \ell = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} (Y - X\beta)'(Y - X\beta)$$

$$\frac{\partial \ell}{\partial \beta} = 0 \quad \text{From lecture} \Rightarrow \hat{\beta}_{K\times 1} = (X'X)^{-1}X'Y$$

$\hat{\beta}$ for the interceptless model is the same as $\hat{\beta}$ with the intercept however now $\hat{\beta}$ is a $K \times 1$ model unlike the $\hat{\beta}$ for the intercept model where $\hat{\beta}$ is a $(K+1) \times 1$ matrix

$$\frac{\partial \ell}{\partial \sigma^2} = \frac{-\mu}{2\sigma^2} + \frac{1}{2\sigma^4} (Y - X\beta)'(Y - X\beta) = 0$$

$$\hat{\sigma}^2 = \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{n} = \frac{e'e}{n} = \hat{\sigma}^2$$

This is the same as the $\hat{\sigma}^2$ for the intercept model.

$$E[\hat{\beta}] = (X'X)^{-1}X'Y = (X'X)^{-1}X'E[Y] = (X'X)^{-1}X'X\beta = \beta \Rightarrow \hat{\beta} \text{ is unbiased}$$

$$E[\hat{\sigma}] = \frac{1}{n} E[e'e] = \frac{1}{n} E[\underbrace{\epsilon'(I-H)\epsilon}_{\text{scalar}}] = \frac{1}{n} E[tr(\epsilon'(I-H)\epsilon)]$$

$$= \frac{1}{n} tr((I-H)E[\epsilon\epsilon']) = \frac{1}{n} tr[(I-H)(\sigma^2 I + 00')]$$

$$= \frac{\sigma^2}{n} tr(I-H) = \frac{\sigma^2}{n}[tr(I_n) - tr(H)] = \frac{\sigma^2}{n}[tr(I_n) - tr(H)]$$

Note: $tr(I) = n$ when we have $X_{n \times k}$

Note: $tr(H) = tr(X(X'X)^{-1}X')$

$= tr((X'X)^{-1}XX) = tr(I_k) = K$

$= \frac{\sigma^2}{n}(n-K) = \frac{\sigma^2(n-K)}{n} \Rightarrow \hat{\sigma}^2$ is not unbiased

unbiased Estimator: $\frac{n\sigma^2}{n-k}$

$E(\hat{\beta}) = \beta$

$var(\hat{\beta}) = var[(X'X)^{-1}X'Y] = (X'X)^{-1}X'(\sigma^2 I)X(X'X)^{-1} = \sigma^2(X'X)^{-1}$

$\Rightarrow \hat{\beta} \sim N(\beta, \sigma^2(X'X)^{-1})$

We know:

$e'X = [(I-H)Y]X = Y'(I-H)X = Y'(X-X) = 0$

This should still hold because this property is dependent on the dimensions of X

Using what we know from exercise 5

$cov(\hat{Y}, \beta) \neq 0$

$cov(\hat{Y}, e) = 0$

$cov(\hat{\beta}, e) = 0$

$\Rightarrow$ $\hat{Y}$ and $\hat{\beta}$ are not independent, but $\hat{Y}$ and e are independent and $\hat{\beta}$ and e are independent due to normality

Exercise 7

a. Multiple regression with $k=5$ predictors $\quad \epsilon_i \sim N(0, \sigma^2 I_n)$

$$\Rightarrow \hat{\beta} \sim N(\beta, \sigma^2 (X'X)^{-1})$$

Find the distribution and pdf for

$$Q = \begin{bmatrix} \hat\beta_1 - 2\hat\beta_2 + 3\hat\beta_4 \\ \hat\beta_0 + \hat\beta_4 + 3\hat\beta_5 \end{bmatrix} = C_{2\times 6}\,\hat\beta$$

$$= \begin{bmatrix} 0 & 1 & -2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} \hat\beta_1 \\ \hat\beta_2 \\ \hat\beta_3 \\ \hat\beta_4 \\ \hat\beta_5 \\ \hat\beta_6 \end{bmatrix}$$

$$E[Q] = C E[\hat\beta] = C\beta$$

$$var(Q) = var(C\hat\beta) = C\,var(\hat\beta)\,C'$$
$$= \sigma^2 CC'$$

$$Q \sim N(C\beta, \sigma^2 (CC'))$$

b. Let $\beta = \begin{bmatrix} \underline{\beta_1} \\ \underline{\beta_2} \end{bmatrix}$ and $X = \begin{bmatrix} \underline{x_1} & x_2 \end{bmatrix}$

Therefore

$$\hat\beta_1 \sim N(\overset{\mu_1}{\beta_1}, \overset{\Sigma_{11}}{var(\hat\beta_1)})$$
$$\hat\beta_2 \sim N(\underset{\mu_2}{\beta_2}, \underset{\Sigma_{22}}{var(\hat\beta_2)})$$

Theorem 4:

If $Q_1 \sim N_p(\mu_1, \Sigma_{11})$ and $Q_2 \sim N_{n-p}(\mu_2, \Sigma_{22})$

Then $Q_1 | Q_2 \sim N_p(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(Q_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$

Using Theorem 4:

$$\hat\beta_1 | \hat\beta_2 \sim N_p(\beta_1 + cov(\hat\beta_1, \hat\beta_2)[var(\hat\beta_2)]^{-1}(\hat\beta_2 - \beta_2), var(\hat\beta_1) - cov(\hat\beta_1, \hat\beta_2)[var(\hat\beta_2)]^{-1} cov(\hat\beta_2, \hat\beta_1))$$

Also: $\quad \hat\beta_0 \sim N(\beta_0, var(\hat\beta_0)) \quad \hat\beta_{(0)} \sim N(\beta_{(0)}, var(\hat\beta_{(0)}))$

Using Theorem 4:

$$\hat\beta_0 | \hat\beta_{(0)} \sim N_p(\beta_0 + cov(\hat\beta_0, \hat\beta_{(0)})[var(\hat\beta_{(0)})]^{-1}(\hat\beta_{(0)} - \beta_{(0)}), var(\hat\beta_0) - cov(\hat\beta_0, \hat\beta_{(0)})[var(\hat\beta_{(0)})]^{-1} cov(\hat\beta_{(0)}, \hat\beta_0))$$

Exercise 8

a. True Model: $Y = X_1\beta_1 + X_2\beta_2 + \epsilon \qquad var(\epsilon) = \sigma^2$

our Model: $Y = X_1\beta_1 + \epsilon$

$$Se^2 = \frac{e'e}{n-k-1} = \frac{tr(e'e)}{n-k-1}$$

Note: $tr(a) = a$ when a is a scalar

Note: $e = (I-H)Y$

$e'e = [(I-H)Y]'[(I-H)Y]$

$= Y'(I-H)Y$

$$= \frac{tr(Y'(I-H_1)Y)}{n-k-1}$$

$$= \frac{tr((I-H_1)YY')}{n-k-1}$$

Note: $tr(ABC) = tr(CAB) = tr(BCA) \neq tr(BAC)$

Note: using the true model for Y:

$Y = X_1\beta_1 + X_2\beta_2$

$$= \frac{tr((I-H_1)(X_1\beta_1 + X_2\beta_2 + \epsilon)(X_1\beta_1 + X_2\beta_2 + \epsilon)')}{n-k-1}$$

Note: $(I-H_1)(X_1\beta_1 + X_2\beta_2 + \epsilon)(X_1\beta_1 + X_2\beta_2 + \epsilon)'$

$= [(I-H_1)X_1\beta_1 + (I-H_1)X_2\beta_2 + (I-H_1)\epsilon][X_1\beta_1 + X_2\beta_2 + \epsilon]'$

Note: $(I-H_1)X_1 = 0$

$= [(I-H_1)X_2\beta_2 + (I-H_1)\epsilon][X_1\beta_1 + X_2\beta_2 + \epsilon]'$

$= [(I-H_1)X_2\beta_2\beta_1'X_1' + (I-H_1)X_2\beta_2\beta_2'X_2 + (I-H_1)X_2\beta_2\epsilon' + (I-H_1)\epsilon\beta_1'X_1' +$
$(I-H_1)\epsilon\beta_2'X_2' + (I-H_1)\epsilon\epsilon']$

Note: $tr((I-H_1)X_2\beta_2\beta_1'X_1') = tr(X_2\beta_2\beta_1'X_1'(1-H_1)') = 0$

Note: $tr(I-H_1)\epsilon\beta_1'X_1') = tr((I-H_1\epsilon\beta_2'X_1') = 0$

using this information, we get:

$$Se^2 = \frac{tr(X_2\beta_2\beta_2'X_2' + (I-H_1)X_2\beta_2\epsilon' + (I-H_1)\epsilon\epsilon')}{n-k-1}$$

Note: $tr(A+B) = tr(A) + tr(B)$

Note: $E[tr(A)] = tr(E[A])$

$$= \frac{1}{n-k-1}[tr(X_2\beta_2\beta_2'X_2') + tr((I-H_1)X_2\beta_2\epsilon') + tr((I-H_1)\epsilon\epsilon')]$$

Note: $E[tr(A)] = tr(E[A])$

$$\Rightarrow E[s_e^2] = \frac{1}{n-k-1}\left[tr(X_2\beta_2\beta_2'X_2') + tr((I-H_1)X_2\beta_2 E[e']) + tr((I-H_1)E[ee'])\right]$$

Note: $E[e'] = 0$

Note: $E[e'e] = \sigma^2 I$

$$= \frac{1}{n-k-1}\left[tr((I-H_1)X_2\beta_2\beta_2'X_2') + tr((I-H_1)\sigma^2 I)\right]$$

$$= \frac{1}{n-k-1}\left[tr((I-H_1)X_2\beta_2\beta_2'X_2') + \sigma^2 tr(I-H_1)\right]$$

$$= \frac{tr((I-H_1)X_2\beta_2\beta_2'X_2')}{n-k-1} + \sigma^2 = \frac{tr(\beta_2'X_2'(I-H_1)X_2\beta_2)}{n-k-1} + \sigma^2$$

$$= \frac{tr(\beta_2'X_2'(I-H_1)'(I-H_1)X_2\beta_2)}{n-k-1} + \sigma^2$$

Let : $u = (I-H_1)X_2\beta_2$

$$= \frac{tr(u'u)}{n-k-1} + \sigma^2 = \frac{u'u}{n-k-1} + \sigma^2 = \frac{\Sigma u_i}{n-k-1} + \sigma^2$$

we know $\Sigma u_i \geq 0$ $\Rightarrow$ $E[s_e^2] = \frac{\Sigma u_i}{n-k-1} + \sigma^2 \geq \sigma^2$

b. $Y \sim N(\mu 1, \sigma^2 V)$

$$L = f(Y) = (2\pi)^{-n/2}|\sigma^2 V|^{-1/2} e^{-1/2 (Y-\mu 1)'(\sigma^2 V)^{-1}(Y-\mu 1)}$$

$$= (2\pi)^{-n/2}(\sigma^2)^{-n/2}|V|^{-1/2} e^{-1/2 (Y-\mu 1)'\frac{1}{\sigma^2}V^{-1}(Y-\mu 1)}$$

$$= (2\pi\sigma^2)^{-n/2}|V|^{-1/2} e^{-1/2 (Y-\mu 1)'\frac{1}{\sigma^2}V^{-1}(Y-\mu 1)}$$

$$\ln(L) = \ell = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2}\ln|V| - \frac{1}{2\sigma^2}(Y-\mu 1)'V^{-1}(Y-\mu 1)$$

$$\frac{\partial \ell}{\partial \mu} = -\frac{1}{2\sigma^2}(Y-\mu 1)'V^{-1}(Y-\mu 1) = 0$$

c. $\mu = \frac{1'V^{-1}Y}{1'V^{-1}1}$ $\sigma^2 = \frac{(Y-\hat{\mu}1)'V^{-1}(Y-\hat{\mu}1)}{n}$ $Y \sim N(\mu 1, \sigma^2 V)$

$$E[\mu] = E\left[\frac{1'V^{-1}Y}{1'V^{-1}1}\right] = \frac{1'V^{-1}}{1'V^{-1}1} E[Y] = \frac{1'V^{-1}\mu 1}{1'V^{-1}1} = \mu \frac{1'V^{-1}1}{1'V^{-1}1} = \mu$$