1.

Given that $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i = 1 \dots n$, where $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$. Show that.

① $Cov(\hat{\beta}_0, \hat{\beta}_1) = -\bar{x}\sigma_0^2 / S_{xx}$.

② $Cov(\bar{y}, \hat{\beta}_1) = 0$

Step 1. $Var(\hat{\beta}_1) = Var\left[\dfrac{\sum(x_i-\bar{x})y_i}{\sum(x_i-\bar{x})x_i}\right] = \dfrac{\sum Var(y_i)(x_i-\bar{x})^2}{[\sum(x_i-\bar{x})^2]^2} = \dfrac{\sigma_0^2}{S_{xx}}$

Step 2. $Var(\hat{\beta}_0) = Var(\bar{y} - \bar{x}\hat{\beta}_1) = (\bar{x})^2 \cdot Var(\hat{\beta}_1) = \dfrac{\bar{x}^2 \cdot \sigma_0^2}{S_{xx}}$

Step 3. $Cov(\hat{\beta}_0, \hat{\beta}_1) = Cov(\bar{y} - \bar{x}\hat{\beta}_1, \hat{\beta}_1) = Cov(\bar{y}, \hat{\beta}_1) - \bar{x}\,Cov(\hat{\beta}_1, \hat{\beta}_1) = Cov(\bar{y}, \hat{\beta}_1) - \boxed{\bar{x} \cdot \dfrac{\sigma_0^2}{S_{xx}}}$.

If 證明式②成立，則式①成立，Prove that $Cov(\bar{y}, \hat{\beta}_1) = 0$.

Step 1. 已知. $\hat{\beta}_1 = \dfrac{\sum(x_i-\bar{x})y_i}{S_{xx}}$，且 $\bar{y} = \frac{1}{n}\sum y_i$，且 $y_i \sim N(\beta_0+\beta_1 x_i, \sigma_0^2)$. 求 $E[(\bar{y}-E(\bar{y}))(\hat{\beta}_1 - E(\hat{\beta}_1))]$

Step 2. $\bar{y} - E(\bar{y})$: $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $\bar{y} = \beta_0 + \beta_1\bar{x} + \bar{\varepsilon}$, $\bar{y} - E(\bar{y}) = \bar{\varepsilon} - E(\bar{\varepsilon}) = \bar{\varepsilon} - 0 = \bar{\varepsilon} \Rightarrow \bar{y} - E(\bar{y}) = \bar{\varepsilon}$.

Step 3. $\hat{\beta}_1 - E(\hat{\beta}_1)$: $\hat{\beta}_1 = \dfrac{\sum(x_i-\bar{x})y_i}{S_{xx}} = \sum a_i y_i$, $a_i = \dfrac{x_i - \bar{x}}{S_{xx}}$

又因 $E(\hat{\beta}_1) = \beta_1 \Rightarrow \hat{\beta}_1 - \beta_1 = (\sum a_i\beta_0 + \sum a_i\beta_1 x_i + \sum a_i\varepsilon_i) - \beta_1$

$\Rightarrow \bar{\varepsilon}(\hat{\beta}_1 - \beta_1) = \boxed{\sum a_i\beta_0}\,\bar{\varepsilon} + \boxed{\sum a_i\beta_1 x_i}\,\bar{\varepsilon} - \beta_1\bar{\varepsilon} + \boxed{\bar{\varepsilon}\sum a_i\varepsilon_i}$

Step 4. $E((\bar{y}-E(\bar{y}))(\hat{\beta}_1 - E(\hat{\beta}_1))) = E(\bar{\varepsilon}(\hat{\beta}_1 - \beta_1)) = E(x \cdot \bar{\varepsilon} + \bar{\varepsilon} \cdot \sum a_i\varepsilon_i)$

$= \underbrace{E(x \cdot \bar{\varepsilon})}_{=0} + E(\bar{\varepsilon}\sum a_i\varepsilon_i)$

Step 5: prove that $E(\bar{\varepsilon}\sum a_i\varepsilon_i) = 0$

已知 $\sum a_i = \sum\dfrac{(x_i-\bar{x})}{S_{xx}} = 0$. $\rightarrow \bar{\varepsilon}\sum_i a_i\varepsilon_i = \frac{1}{n}\sum_j \varepsilon_j \sum a_i\varepsilon_i = \frac{1}{n}(\varepsilon_1 + \varepsilon_2 \dots \varepsilon_n)(a_1\varepsilon_1 + a_2\varepsilon_2 \dots a_n\varepsilon_n)$.

$= \frac{1}{n}\left[\sum_{i,j} a_i\varepsilon_i\varepsilon_j\right] = \sum_i a_i\varepsilon_i^2 + \sum_{i \neq j} a_i\varepsilon_i\varepsilon_j = \sum a_i\,Var(\varepsilon_i) = \sigma^2\sum a_i = 0$ #

back to ①, $\because Cov(\bar{y}, \hat{\beta}_1) = 0$ $\therefore Cov(\hat{\beta}_0, \hat{\beta}_1) = -\bar{x}\sigma_0^2 / S_{xx}$. #.

QED.

2. $SS_R = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i^2 - 2\bar{y}\hat{y}_i + \bar{y}^2) = \sum_{i=1}^{n}\hat{y}_i^2 - 2\underbrace{\boxed{\sum_{i=1}^{n}\hat{y}_i}}_{}\bar{y} + \sum_{i=1}^{n}\bar{y}^2 = \sum_{i=1}^{n}\hat{y}_i^2 - \sum_{i=1}^{n}\bar{y}^2 = \sum_{i=1}^{n}\hat{y}_i^2 - n\bar{y}^2$. Q.E.D.

$=\bar{y}$

3. (10%) The matrix, $\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$, derived in multiple regression is usually defined as $\mathbf{H}$. Show that:
   a. $\mathbf{H}$ is idempotent, i.e., $\mathbf{HH} = \mathbf{H}$ and $(\mathbf{I}-\mathbf{H})(\mathbf{I}-\mathbf{H}) = \mathbf{I}-\mathbf{H}$
   b. $V(\hat{\mathbf{y}}) = \sigma^2\mathbf{H}$

a. $H^2 = X(X^TX)^{-1}\overbrace{\boxed{X^TX(X^TX)^{-1}}}^{I}X^T = X(X^TX)^{-1}X^T = H$. and $(I-H)^2 = I^2 - 2H + H^2 = I^2 - 2H + H = I - H$. Q.E.D.

b. $\hat{y} = X\hat{\beta} = X(X^TX)^{-1}X^Ty = X(X^TX)^{-1}X^T(X\beta + \varepsilon) = X[\beta + (X^TX)^{-1}X^T\varepsilon]$.

$Var(\hat{y}) = Var(X\beta + X(X^TX)^{-1}X^T\varepsilon)$

$= Var(X(X^TX)^{-1}X^T\varepsilon) = X(X^TX)^{-1}X^T Var(\varepsilon) = X(X^TX)^{-1}X^T\sigma^2 = H\sigma^2$. Q.E.D.

4. (10%) Investigate and explain why R^2 cannot be larger than 1 or smaller than 0. (Do not copy directly from the source you found, but explain in your own words.)

Since $SS_T = SS_R + SS_{RES}$ and SS_R, SS_{RES} must larger than 0 because they are "sum of squares", thus SS_R must be a part of SS_T and R^2 is $\left(\frac{SS_R}{SS_T}\right)^2$, $\because 0 \le SS_R \le SS_T \therefore 0 \le \frac{SS_R}{SS_T} \le 1 \rightarrow 0 \le \left(\frac{SS_R}{SS_T}\right)^2 \le 1 \rightarrow 0 \le R^2 \le 1$.

5.