Simple Linear Regression Model.

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad \varepsilon_i \overset{iid}{\sim} N(0,\sigma^2), \quad i=1,\cdots,n.$$

(1) Least Square Objective function

$$Q(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

Least Square Estimates for β_0, β_1:

$$\begin{cases} \frac{\partial Q}{\partial \beta_0} = 0 \\ \frac{\partial Q}{\partial \beta_1} = 0 \end{cases} \Rightarrow \begin{cases} n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i \cdot \beta_0 + \beta_1 \cdot \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \end{cases} \Rightarrow \begin{cases} \hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{S_{xy}}{S_{xx}} \\ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \cdot \bar{x} \end{cases}$$

(2) $\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{S_{xx}}\right)$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{S_{xx}}\right) \cdot y_i = \sum_{i=1}^{n} c_i y_i, \quad \text{where } y_i \overset{ind.}{\sim} N(\beta_0 + \beta_1 x_i, \sigma^2).$$

$$\sum_i c_i = 0, \quad \sum_i c_i (x_i - \bar{x}) = \sum_i c_i x_i = 1, \quad \sum_i c_i^2 = 1/S_{xx}$$

$$E(\hat{\beta}_1) = \sum_i c_i (\beta_0 + \beta_1 x_i) = \beta_1$$

$$Var(\hat{\beta}_1) = \sum_i c_i^2 \cdot Var(y_i) = \frac{\sigma^2}{S_{xx}}, \quad y_1, \cdots, y_n \text{ are mutually independent.}$$

$\because$ $\hat{\beta}_1$ is a linear combination of a sequence of normal variables, $\hat{\beta}_1 = \sum_i c_i y_i$

$\therefore$ it is normally distributed, i.e. $\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{S_{xx}}\right)$

Standardization $\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\sigma^2/S_{xx}}} \sim N(0,1)$ or. $\frac{(\hat{\beta}_1 - \beta_1)^2}{\sigma^2/S_{xx}} \sim \chi^2_{(1)}$

$$\left.\begin{array}{l} Var(\hat{\beta}) = \frac{\sigma^2}{S_{xx}} \\ sd(\hat{\beta}) = \sqrt{\frac{\sigma^2}{S_{xx}}} \end{array}\right\} \text{Standard error } se(\hat{\beta}) = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$ where $\hat{\sigma}^2$ is a good estimator of σ^2.

(3) Decomposition of Sum of Squares

$$\sum_{i=1}^{n}(y_i-\bar{y})^2 = \sum_i (y_i-\hat{y}_i)^2 + \sum_i(\hat{y}_i-\bar{y})^2, \quad \text{where } \hat{y}_i=\hat{\beta}_0+\hat{\beta}_1 x_i \text{ is predicted value.}$$
(SSTO) (SSE) (SSR)

It can be shown that $\sum_i (y_i-\hat{y}_i)(\hat{y}_i-\bar{y})=0$.

a) $DF(SSTO)=(n-1)$ $\quad \frac{SSTO}{\sigma^2}\sim \chi^2(n-1)$ when $\beta_1=0$, $y_1 \cdots y_n \overset{iid}{\sim} N(\beta_0,\sigma^2)$

b) $DF(SSR)=1$

$SSR=\sum_i(\hat{y}_i-\bar{y})^2=\hat{\beta}_1^2\cdot S_{xx}$, only one changing variable.

$$\begin{cases} SSR=\sum_i[\hat{\beta}_0+\hat{\beta}_1 x_i-(\hat{\beta}_0+\hat{\beta}_1\bar{x})]^2=\hat{\beta}_1^2\cdot\sum_{i=1}^{n}(x_i-\bar{x})^2 \\ \therefore SSR=\hat{\beta}_1^2\cdot S_{xx}\end{cases}$$

From (2), $\frac{(\hat{\beta}_1-\beta_1)^2}{\sigma^2/S_{xx}}\sim\chi^2(1) \xrightarrow{\text{when } \beta_1=0} \frac{\hat{\beta}_1^2}{\sigma^2/S_{xx}}\sim\chi^2(1)$, $\frac{SSR}{\sigma^2}\sim\chi^2(1)$

c) $DF(SSE)=DF(SSTO)-DF(SSR)=(n-1)-1=n-2$ as $SSE \perp\!\!\!\perp SSR$.

<u>To show $SSE \perp\!\!\!\perp SSR$</u> (independent), we need to use matrix algebra.

$$Y_i=\beta_0+\beta_1 x_i+\varepsilon_i \Rightarrow \begin{pmatrix}Y_1\\ \vdots\\ Y_n\end{pmatrix}=\begin{pmatrix}1 & x_1\\ \vdots & \vdots\\ 1 & x_n\end{pmatrix}\begin{pmatrix}\beta_0\\ \beta_1\end{pmatrix}+\begin{pmatrix}\varepsilon_1\\ \vdots\\ \varepsilon_n\end{pmatrix} \text{ or } Y=X\beta+\varepsilon$$

Response vector $Y\in R^n$, Parameter $\beta\in R^2$, error vector $\in R^n$, $E(\varepsilon)=0$, $Var(\varepsilon)=\sigma^2 I_n$

The normal equation: $\frac{\partial Q}{\partial\beta}=\frac{\partial}{\partial\beta}[(Y-X\beta)'(Y-X\beta)]=0$

$$\Leftrightarrow X'X\beta=X'Y \quad \therefore \hat{\beta}=(X'X)^{-1}X'Y$$

Residual $\hat{e}=Y-\hat{Y}=Y-X\cdot\hat{\beta}=Y-X\cdot(X'X)^{-1}X'Y=(I-H)Y$

where $H=X\cdot(X'X)^{-1}X'$ is symmetric and idempotent,

i.e. $H'=H$, $H^2=H$.

$$Cov(\hat{\beta}',\hat{e}')=Cov(\underbrace{(X'X)^{-1}X'}_{A}Y,\underbrace{(I-H)}_{B}Y)=A\cdot Var(Y)\cdot B$$
$$=A\cdot(\sigma^2 I_n)\cdot B=\sigma^2\cdot A\cdot B=\sigma^2\cdot(X'X)^{-1}X'\cdot(I-H)=0$$

$$\left.\begin{array}{l}\because \hat{\beta},\ \hat{e} \text{ both normally distributed}\\ \therefore \hat{\beta}\perp\!\!\!\perp\hat{e}\end{array}\right\}\Rightarrow SSR=\hat{\beta}_1^2\cdot S_{xx}\perp\!\!\!\perp SSE=\hat{e}'\cdot\hat{e}$$

(4) Expectation of Sum Squares

a) $SSTO = \sum_{i=1}^{n}(y_i-\bar{y})^2 = (n-1)S_y^2$, where S_y^2 is the sample variance

Under $\beta_1=0$, based on Studard's theorem, $\frac{(n-1)S_y^2}{\sigma^2} = \frac{SSTO}{\sigma^2} \sim \chi^2(n-1)$.

For general (β_0,β_1), $Y_i \overset{ind.}{\sim} N(\beta_0+\beta_1 x_i, \sigma^2)$ are not identical,

Let $Y_i^* = Y_i-(\beta_0+\beta_1 x_i)$, then $Y_i^* \overset{iid}{\sim} N(0,\sigma^2)$, $i=1,\dots,n$

i.e. $\frac{\sum_{i=1}^{n}(y_i^*-\bar{y}^*)^2}{\sigma^2} \sim \chi^2(n-1)$, $\therefore E(y_i^*-\bar{y}^*)^2 = (n-1)\sigma^2$. (*)

Define $SSTO^* = \sum_{i=1}^{n}(y_i^*-\bar{y}^*)^2$, $\bar{y}^* = \frac{1}{n}\sum_{i=1}^{n} y_i^* = \bar{y}-(\beta_0+\beta_1\bar{x})$

$$SSTO^* = \sum_{i=1}^{n}\left[y_i-(\beta_0+\beta_1 x_i)-(\bar{y}-(\beta_0+\beta_1\bar{x}))\right]^2$$

$$= \sum_{i=1}^{n}(y_i-\bar{y}-\beta_1(x_i-\bar{x}))^2$$

$$= \sum_{i=1}^{n}(y_i-\bar{y})^2 - 2\cdot\sum_{i=1}^{n}\beta_1\cdot(x_i-\bar{x})(y_i-\bar{y}) + \beta_1^2\cdot\sum_{i=1}^{n}(x_i-\bar{x})$$

$$\therefore SSTO^* = SSTO - 2\beta_1\cdot\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) + \beta_1^2\cdot S_{xx}$$

Take expectation on both sides. [$y_i \sim N(\beta_0+\beta_1 x_i, \sigma^2)$, $\bar{y} \sim N(\beta_0+\beta_1\bar{x}, \frac{\sigma^2}{n})$)

$$E[SSTO^*] = E[SSTO] - 2\beta_1\cdot\sum_{i=1}^{n}(x_i-\bar{x})\Big[\overset{\to E y_i}{(\beta_0+\beta_1 x_i)} - \overset{\to E\bar{y}}{(\beta_0+\beta_1\bar{x})}\Big] + \beta_1^2 S_{xx}$$

$$= E[SSTO] - 2\beta_1\cdot\sum_{i=1}^{n}(x_i-\bar{x})\cdot\beta_1\cdot(x_i-\bar{x}) + \beta_1^2\cdot S_{xx}$$

$$E[SSTO^*] = E[SSTO] - 2\beta_1^2\cdot\sum_{i=1}^{n}(x_i-\bar{x})^2 + \beta_1^2 S_{xx}$$

$$\therefore E[SSTO] = E[SSTO^*] + \beta_1^2\cdot S_{xx} \overset{(*)}{=} (n-1)\sigma^2 + \beta_1^2\cdot S_{xx}.$$

When $\beta_1=0$, then $E[SSTO] = (n-1)\sigma^2$

b). $SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = \hat{\beta}_1^2 \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2$. $\hat{\beta}_1 \sim N(\beta_1, \frac{\sigma^2}{S_{xx}})$

when $\beta_1 = 0$. $\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\sigma^2/S_{xx}}} = \frac{\hat{\beta}_1 \cdot \sqrt{S_{xx}}}{\sqrt{\sigma^2}} \sim N(0,1)$, $\Leftrightarrow$ $\frac{\hat{\beta}_1^2 \cdot S_{xx}}{\sigma^2} \sim \chi^2(1)$

For a general β_1, $\frac{(\hat{\beta}_1 - \beta_1)^2 \cdot S_{xx}}{\sigma^2} \sim \chi^2(1)$

$\therefore E[SSR] = \sigma^2 + \beta_1^2 \cdot S_{xx}$ (homework problem).

c). $SSE = SSTO - SSR$

For any β_1 value, $E[SSE] = E[SSTO] - E[SSR]$

$$\begin{cases} E[SSTO] = (n-1)\sigma^2 + \beta_1^2 \cdot S_{xx} \\ E[SSR] = \sigma^2 + \beta_1^2 \cdot S_{xx} \end{cases}$$

$\therefore E[SSE] = (n-2)\sigma^2$, (independant of β_1)

$\therefore$ Mean Square Error $MSE = \frac{SSE}{DF(SSE)} = \frac{SSE}{n-2}$

is unbiased for σ^2 unconditionally, i.e.

$$E[MSE] = E\left[\frac{SSE}{n-2}\right] = \sigma^2$$

Mean Square of Regression

$$E[MSR] = E\left[\frac{SSR}{1}\right] = \sigma^2 + \beta_1^2 \cdot S_{xx}$$

If $\beta_1 \neq 0$, then $E[MSR] > E[MSE]$.

If $\beta_1 = 0$, then $E[MSR] = E[MSE]$.