Problem 1. $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2)$

(a) Least square Estimators of β_0, β_1 are the following:

$$\begin{cases} \hat{\beta_1} = \dfrac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \dfrac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} \\ \hat{\beta_0} = \bar{y} - \hat{\beta_1}\bar{x} \end{cases}$$

Show that they are unbiased estimators :
in other words, to prove ① $E(\hat{\beta_1}) = \beta_1$ ② $E(\hat{\beta_0}) = \beta_0$

① First prove $E(\hat{\beta_1}) = \beta_1$

Given that $\hat{\beta_1} = \dfrac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \dfrac{S_{xy}}{S_{xx}}$

let $k_i = \dfrac{x_i-\bar{x}}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \dfrac{x_i-\bar{x}}{S_{xx}}$, $i = 1, 2 \ldots n$

so k_i is constant for each level of X at $i = 1, 2, \ldots n$.

k_i has a property that $\sum_{i=1}^{n} k_i = \sum_{i=1}^{n} \dfrac{x_i-\bar{x}}{S_{xx}} = \dfrac{\sum_{i=1}^{n} x_i-\bar{x}}{S_{xx}} = 0$

Rewrite $\hat{\beta_1}$ as a linear combination of y_i observations

$$\hat{\beta_1} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{S_{xx}} = \sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{S_{xx}}\right)(y_i-\bar{y}) = \sum_{i=1}^{n}(k_i)(y_i-\bar{y})$$

$$= \sum_{i=1}^{n} k_i y_i - \underbrace{\bar{y}\sum_{i=1}^{n} k_i}_{0}$$

$$= \sum_{i=1}^{n} k_i y_i$$

because $\sum_{i=1}^{n} k_i = 0$ as proven above

then $E(\hat{\beta_1}) = E(\sum_{i=1}^{n} k_i y_i) = \sum_{i=1}^{n} E(k_i y_i)$ by linear property of expectation.

$$= \sum_{i=1}^{n} k_i E(y_i)$$

we need to find $E(y_i)$

our model is $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $\varepsilon_i \sim N(0, \sigma^2)$

then

$$E(y_i) = E(\beta_0+\beta_1 x_i+\varepsilon_i) = E(\beta_0+\beta_1 x_i)+E(\varepsilon_i) = \beta_0+\beta_1 x_i$$

thus $E(\hat{\beta_1}) = \sum_{i=1}^{n} k_i(\beta_0+\beta_1 x_i) = \underbrace{\beta_0 \sum_{i=1}^{n} k_i}_{=0} + \beta_1 \sum_{i=1}^{n} k_i x_i = \beta_1 \sum_{i=1}^{n} k_i x_i$

work on $\sum_{i=1}^{n} k_i x_i = \sum_{i=1}^{n} \left(\frac{x_i-\bar{x}}{S_{xx}}\right) x_i$

$$= \frac{1}{S_{xx}} \sum_{i=1}^{n} (x_i-\bar{x}) x_i$$

$$= \frac{1}{S_{xx}} \sum_{i=1}^{n} (x_i-\bar{x})(x_i-\bar{x})$$

$\sum (x_i-\bar{x})(x_i-\bar{x})$

$= \sum x_i(x_i-\bar{x}) - \underbrace{\sum \bar{x}(x_i-\bar{x})}_{=0}$

$= \sum x_i(x_i-\bar{x})$

so equivalent

$$= \frac{1}{S_{xx}} \sum_{i=1}^{n} (x_i-\bar{x})^2$$

$$= \frac{1}{S_{xx}} \cdot S_{xx} = 1$$

therefore $E(\hat{\beta_1}) = \beta_1 \sum_{i=1}^{n} k_i x_i = \beta_1$ (unbiased estimator)

② prove $E(\hat{\beta_0}) = \beta_0$

Given that $\hat{\beta_0} = \bar{y} - \hat{\beta_1}\bar{x}$

$E(\hat{\beta_0}) = E(\bar{y}-\hat{\beta_1}\bar{x}) = E(\bar{y}) - \bar{x}E(\hat{\beta_1}) = E(\bar{y}) - \bar{x}\beta_1$

Find $E(\bar{y}) = E\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(y_i) = \frac{1}{n}\sum_{i=1}^{n}(\beta_0+\beta_1 x_i)$

$$= \frac{1}{n}\left(n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i\right)$$

$$= \beta_0 + \beta_1 \bar{x}$$

then $E(\hat{\beta_0}) = \beta_0 + \beta_1\bar{x} - \bar{x}\beta_1 = \beta_0$ (unbiased estimator)

Hence, Least square estimators $\hat{\beta_0}$, $\hat{\beta_1}$ are unbiased estimator of true β_0, β_1.

(b) fitted regression line: $\hat{Y_i} = \hat{\beta_0} + \hat{\beta_1} X_i$

$$\hat{Y_i} = \hat{\beta_0} + \hat{\beta_1}(X_i - \overline{X}) + \hat{\beta_1}\overline{X}$$

$$= \hat{\beta_0} + \hat{\beta_1}\overline{X} + \hat{\beta_1}(X_i - \overline{X})$$

$$= \overline{Y} + \hat{\beta_1}(X_i - \overline{X})$$

because $\hat{\beta_0} = \overline{Y} - \hat{\beta_1}\overline{X}$

then $\overline{Y} = \hat{\beta_0} + \hat{\beta_1}\overline{X}$

this is an alternative form of fitted line:

$$\hat{Y_i} = \overline{Y} + \hat{\beta_1}(X_i - \overline{X})$$

when $X_i = \overline{X}$, $\hat{Y_i} = \overline{Y} + 0 = \overline{Y}$

So this regression line always goes through $(\overline{X}, \overline{Y})$

(c) use MLE to derive estimator for σ^2

Model: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, $\varepsilon_i \sim N(0, \sigma^2)$

Here we assume that ε_i is normally distributed.

So the pdf of Y: $f(Y_i \mid \beta_0, \beta_1, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(Y_i - \beta_0 - \beta_1 X_i)^2}{2\sigma^2}\right]$

$$= \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{Y_i - \beta_0 - \beta_1 X_i}{\sigma}\right)^2\right]$$

The likelihood function becomes:

$$L(\sigma^2 \mid Y_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{Y_i - \beta_0 - \beta_1 X_i}{\sigma}\right)^2\right]$$

$$= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \cdot \prod_{i=1}^{n} \exp\left[-\frac{1}{2}\left(\frac{Y_i - \beta_0 - \beta_1 X_i}{\sigma}\right)^2\right]$$

Take log-likelihood function:

$$\log L(\sigma^2 \mid Y_i) = \log f(\sigma^2 \mid Y_i) = -n\log(\sqrt{2\pi}\sigma) + \log \prod_{i=1}^{n} \exp\left[-\frac{1}{2}\left(\frac{Y_i - \beta_0 - \beta_1 X_i}{\sigma}\right)^2\right]$$

$$= -n\log(\sqrt{2\pi}\sigma) + \sum_{i=1}^{n} -\frac{1}{2}\left(\frac{Y_i - \beta_0 - \beta_1 X_i}{\sigma}\right)^2$$

we need to find $\hat{\sigma}^2$ that maximize this $\log L(\sigma^2)$.

$$\frac{d \log L(\sigma^2)}{d\sigma} = -\frac{n\cdot\sqrt{2\pi}}{\sqrt{2\pi}\sigma} + \left(-\frac{1}{2}\right)\sum_{i=1}^{n} 2\left(\frac{Y_i-\hat{\beta_0}-\hat{\beta_1}X_i}{\sigma}\right)\cdot(Y_i-\hat{\beta_0}-\hat{\beta_1}X_i)\left(-\frac{1}{\sigma^2}\right)$$

$$= -\frac{n}{\sigma} + \sum_{i=1}^{n}\frac{(Y_i-\hat{\beta_0}-\hat{\beta_1}X_i)^2}{\sigma^3}$$

Set $\frac{d \log L(\sigma^2)}{d\sigma} = 0$

$$-\frac{n}{\sigma} + \sum_{i=1}^{n}\frac{(Y_i-\hat{\beta_0}-\hat{\beta_1}X_i)^2}{\sigma^3} = 0$$

$$\frac{1}{\sigma^2}\Sigma(Y_i-\hat{\beta_0}-\hat{\beta_1}X_i)^2 = n$$

then $\hat{\sigma}^2 = \frac{\Sigma(Y_i-\hat{\beta_0}-\hat{\beta_1}X_i)^2}{n} = \frac{\Sigma(Y_i-\hat{Y_i})^2}{n} = \frac{SSE}{n} = MSE$

thus MSE is the MLE estimater of σ^2

Then we need to find $E(MSE)$

Since $MSE = \frac{SSE}{n}$, we can first find $E(SSE)$

$$E(SSE) = E\left(\sum_{i=1}^{n}(Y_i-\hat{Y_i})^2\right)$$

$$= E\left(\sum_{i=1}^{n} e_i^2\right)$$

$$= E\left(\Sigma(e_i-\bar{e_i})^2\right) \quad \text{because } \bar{e_i}=0$$

$$= \Sigma E(e_i-\bar{e_i})^2 = \sum_{i=1}^{n} Var(e_i)$$

Rewrite e_i :

$$e_i = Y_i - \hat{Y_i} = Y_i - (\hat{\beta_0}+\hat{\beta_1}X_i) = Y_i - (\bar{Y}-\hat{\beta_1}\bar{X}+\hat{\beta_1}X_i)$$

$$= (Y_i-\bar{Y}) - (X_i-\bar{X})\hat{\beta_1}$$

then $\sum_{i=1}^{n} Var(e_i) = \sum_{i=1}^{n} Var[(Y_i-\bar{Y})-(X_i-\bar{X})\hat{\beta_1}]$

$$= \sum_{i=1}^{n} Var(Y_i-\bar{Y}) + Var[(X_i-\bar{X})\hat{\beta_1}] - 2\,cov[(Y_i-\bar{Y}),(X_i-\bar{X})\hat{\beta_1}]$$

$$= (n-1)\sigma^2 + \Sigma Var[(X_i-\bar{X})\hat{\beta_1}] - \underbrace{\Sigma 2\,cov[(Y_i-\bar{Y}),(X_i-\bar{X})\hat{\beta_1}]}$$

$$= \Sigma 2\,cov(\hat{\beta_1},\hat{\beta_1})\cdot(X_i-\bar{X})^2$$

$$= \Sigma 2\,Var(\hat{\beta_1})\cdot(X_i-\bar{X})^2$$

$$= (n-1)\sigma^2 + \sum_{i=1}^{n} (x_i-\bar{x})^2 Var(\hat{\beta_i}) - \sum 2(x_i-\bar{x})^2 Var(\hat{\beta_i})$$

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

$$= (n-1)\sigma^2 - \sum (x_i-\bar{x})^2 \cdot \frac{\sigma^2}{\sum (x_i-\bar{x})^2}$$

$$= (n-1)\sigma^2 - \sigma^2 = (n-2)\sigma^2$$

therefore $E(SSE) = (n-2)\sigma^2$

then $E(\frac{SSE}{n-2}) = E(MSE) = \sigma^2$

therefore MSE is an unbiased estimator of σ^2 in any situation.