Question 3

(a)

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \quad \text{for } i = 1, \ldots, n \text{ from question}$$

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

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

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

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

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

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

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

(b) (i) $\mathbb{E}\left[\hat{\beta}_1 \mid x_1, \dots, x_n\right]$

$$= \mathbb{E}\left[\hat{\beta}_1 \mid X\right], \quad X = (x_1, \dots, x_n)$$

$$= \mathbb{E}\left[\frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \,\Big|\, X\right], \quad i = 1, \dots, n \text{ from (9)}$$

$$= \mathbb{E}\left[\frac{\sum_{i=1}^{n}(x_i - \bar{x})(\beta_0 + \beta_1 x_i + u_i)}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \,\Big|\, X\right] \text{ from (4)}$$

$$= \mathbb{E}\left[\frac{\sum_{i=1}^{n}\left[(x_i - \bar{x})\beta_0 + (x_i - \bar{x})\beta_1 x_i + (x_i - \bar{x})u_i\right]}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \,\Big|\, X\right]$$

$$= \mathbb{E}\left[\frac{\beta_0 \sum_{i=1}^{n}(x_i - \bar{x}) + \beta_1 \sum_{i=1}^{n}(x_i - \bar{x})x_i + \sum_{i=1}^{n}(x_i - \bar{x})u_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \,\Big|\, X\right]$$

Now, $\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x}$

$$= n \sum_{i=1}^{n} \frac{x_i}{n} - n\bar{x}$$

$$= n\bar{x} - n\bar{x}$$

$$= 0$$

So, $$\mathbb{E}\left[\hat{\beta}_1 \mid x_1, \dots, x_n\right] = \mathbb{E}\left[\frac{\beta_1 \sum_{i=1}^{n}(x_i - \bar{x})x_i + \sum_{i=1}^{n}(x_i - \bar{x})u_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \,\Big|\, X\right]$$

$$= \mathbb{E}\left[\beta_1 \frac{\sum_{i=1}^{n} (x_i - \bar{x}) x_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2} + \frac{\sum_{i=1}^{n} (x_i - \bar{x}) u_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \,\middle|\, X \right]$$

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

$\sum_{i=1}^{n} (x_i - \bar{x}) = 0$ from before.

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

$$= 1$$

So, $$\mathbb{E}[\hat{\beta}_1 \mid x_1, \dots, x_n] = \mathbb{E}\left[\beta_1 + \frac{\sum_{i=1}^{n} (x_i - \bar{x}) u_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \,\middle|\, X\right]$$

$$= \mathbb{E}[\beta_1 \mid X] + \mathbb{E}\left[\frac{\sum_{i=1}^{n} (x_i - \bar{x}) u_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \,\middle|\, X\right]$$

$$= \beta_1 + \frac{\sum_{i=1}^{n} (x_i - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \mathbb{E}[u_i \mid X]$$

$$= \beta_1 \quad \text{as } \mathbb{E}[u_i \mid X] = \mathbb{E}[u_i \mid x_1, \dots, x_n] = 0$$

from zero conditional mean assumption.

(ii) $Var[\hat{\beta}|x_1, \dots, x_n] = Var[\hat{\beta}|X]$, $X = (x_1, \dots, x_n)$

$$= Var\left[\beta_1 + \frac{\sum_{i=1}^{n}(x_i-\bar{x})u_i}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \middle| X\right] \quad \text{from part(b)i)}$$

$$= Var\left[\frac{\sum_{i=1}^{n}(x_i-\bar{x})u_i}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \middle| X\right] \quad \text{as } \beta_1 \text{ is a constant}$$

$$= \left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^2 Var[u_i|X] \quad \text{as only } u_i \text{ is a r.v.}$$

Now, $Var[y_i|x_1, \dots, x_n] = Var[y_i|X]$

$$= Var[\beta_0 + \beta_1 x_i + u_i|X] \quad \text{from (4)}$$

$$= Var[u_i|X] \quad \text{as only } u_i \text{ is a r.v.}$$

$$= \sigma^2 \quad \text{from (8)}$$

So, $Var[\hat{\beta}|x_1, \dots, x_n] = \left(\frac{\sum_{i}^{n}(x_i-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^2 \sigma^2$

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

(iii) $\mathrm{Var}[\hat{\beta}_1] = \mathbb{E}\left[\mathrm{Var}[\hat{\beta}_1 | x_1, \ldots, x_n]\right] + \mathrm{Var}\left[\mathbb{E}[\hat{\beta}_1 | x_1, \ldots, x_n]\right]$

$= \mathbb{E}\left[\frac{1}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sigma^2\right] + \mathrm{Var}[\beta_1]$

from part (b) i) and part (b) ii)

$= \mathbb{E}\left[\frac{1}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sigma^2\right]$ as β_1 is a constant

$= \frac{\sigma^2}{\mathbb{E}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]}$ as σ^2 is a constant

Now, $\mathbb{E}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right] = \frac{n-1}{n-1}\mathbb{E}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]$

$= (n-1)\mathbb{E}\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]$

$= (n-1)\mathbb{E}[\hat{\sigma}_x^2]$

So, $\mathrm{Var}[\hat{\beta}_1] = \frac{\sigma^2}{(n-1)\mathbb{E}[\hat{\sigma}_x^2]}$

(c) From part (b), we are aware of two things:

$$\mathbb{E}[\hat{\beta}_1 | x_1, \ldots, x_n] = \beta_1$$

and
$$Var[\hat{\beta}_1 | x_1, \ldots, x_n] = \frac{\sigma^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

Now, the circumstances that provide the best opportunity to precisely estimate $\hat{\beta}_1$ is when $\hat{\beta}_1$ is BLUE (best linear unbias estimator). This occurs when we ~~have~~ our population model is linear in the parameter and is correctly specified, data on $(x_1, \ldots, x_n)$ represents a random sample from the population described by the model, there is variation in x, we have zero conditional mean $\mathbb{E}[u_i | x_1, \ldots, x_n] = 0$ and homoskedasticity $Var[u_i | x_1, \ldots, x_n] = \sigma^2$. These assumptions imply, according to the Gauss-Markov theorem, that $\hat{\beta}_1$ is BLUE. Hence, we know that $\frac{\sigma^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$ is the lowest variance of any linear estimator of β_1, i.e. we can precisely estimate β_1.