Proof:

$$V(\hat{\beta}_0) = V(\bar{y} - \hat{\beta}_1 \bar{X})$$
$$= V(\bar{y} + (-\bar{X})\hat{\beta}_1)$$

Inserting the variance of the brackets on the right side, we get:

$$V(\hat{\beta}_0) = V(\bar{y}) + 2\,\mathrm{cov}(\bar{y}, (-\bar{X})\hat{\beta}_1) + \bar{X}^2 V(\hat{\beta}_1)$$

First: $V(\bar{y})$

Proof:

$$\bar{y} = \frac{\sum y_i}{n} = \left(\frac{1}{n} y_1 + \frac{1}{n} y_2 + \ldots + \frac{1}{n} y_n\right)$$
$$= \frac{1}{n^2} V(y_1) + \frac{1}{n^2} V(y_2) + \ldots + \frac{1}{n^2} V(y_n)$$
$$= \frac{n}{n^2} V(y_i)$$
$$= \frac{V(y_i)}{n}$$
$$= \frac{\hat{\sigma}^2}{n}$$

Second: $cov(\bar{y}, (-\bar{X})\hat{\beta}_1)$

Proof:

$$\mathrm{cov}(\bar{y}, (\bar{X})\hat{\beta}_1) = (-\bar{X})\,\mathrm{cov}(\bar{y}, \hat{\beta}_1)$$
$$a = a_1 y_1 + a_2 y_2 + \ldots + a_n y_n$$
$$c = c_1 y_1 + c_2 y_2 + \ldots + c_n y_n$$
$$\mathrm{cov}(a,c) = a_1 c_1 \,\mathrm{cov}(y_1, y_1) + a_2 c_2 \,\mathrm{cov}(y_2, y_2) + \ldots + a_n c_n \,\mathrm{cov}(y_n, y_n)$$
$$= a_1 c_1 V(y_1) + a_2 c_2 V(y_2) + \ldots + a_n c_n V(y_n)$$

We have $V(y_1) = V(y_2) = \ldots = V(y_n) = V(y_i)$

$$\mathrm{cov}(a,c) = (a_1 c_1 + a_2 c_2 + \ldots + a_n c_n) V(y_i)$$
$$= (a_1 c_1 + a_2 c_2 + \ldots + a_n c_n)\hat{\sigma}^2$$
$$= \sum_{i=1}^{n} (a_i c_i)\,\hat{\sigma}^2$$

Suppose: $c = \hat{\beta}_1, a = \bar{y}$

$$\bar{y}=\frac{\sum y_i}{n}=\left(\frac{1}{n}y_1+\frac{1}{n}y_2+\ldots+\frac{1}{n}y_n\right)$$

$$=\left(\frac{1}{n}+\frac{1}{n}+\ldots+\frac{1}{n}\right)y_i$$

$$\therefore a_i=\frac{1}{n}$$

and

$$\hat{\beta}_1=\frac{S_{Xy}}{S_{XX}}=\frac{\sum(x_i-\bar{X})y_i}{\sum(x_i-\bar{X})^2}=\frac{(x_1-\bar{X})y_1}{\sum(x_i-\bar{X})^2}+\frac{(x_2-\bar{X})y_2}{\sum(x_i-\bar{X})^2}+\ldots+\frac{(x_n-\bar{X})y_n}{\sum(x_i-\bar{X})^2}$$

$$\hat{\beta}_1=\left(\frac{(x_1-\bar{X})}{\sum(x_i-\bar{X})^2}+\frac{(x_2-\bar{X})}{\sum(x_i-\bar{X})^2}+\ldots+\frac{(x_n-\bar{X})}{\sum(x_i-\bar{X})^2}\right)y_i$$

So: $c=\frac{(x_i-\bar{X})}{\sum(x_i-\bar{X})^2}$

$$cov(\bar{y},\hat{\beta}_1)=cov(a,c)=\sum\left(\frac{1}{n}\frac{(x_i-\bar{X})}{\sum(x_i-\bar{X})^2}\right)\hat{\sigma}^2$$

$$=\left(\frac{1}{n}\frac{\sum(x_i-\bar{X})}{S_{XX}}\right)\hat{\sigma}^2$$

$$cov(\bar{y},\hat{\beta}_1)=0$$

Were $\sum(x_i-\bar{X})=0$

Third: $\bar{X}^2V(\hat{\beta}_1)$

Proof:

$$\bar{X}^2V(\hat{\beta}_1)=\bar{X}^2\frac{\hat{\sigma}^2}{S_{XX}}$$

$$\therefore V(\hat{\beta}_0)=\frac{\hat{\sigma}^2}{n}+\bar{X}^2\frac{\hat{\sigma}^2}{S_{XX}}V(\hat{\beta}_0)=\hat{\sigma}^2\left(\frac{1}{n}+\frac{\bar{X}^2}{S_{XX}}\right)$$

$$V(\hat{\beta}_0)=\hat{\sigma}^2\left(\frac{S_{XX}+n\bar{X}^2}{nS_{XX}}\right)$$

$$V(\hat{\beta}_0)=\hat{\sigma}^2\left(\frac{\sum X_i^2-n\bar{X}^2+n\bar{X}^2}{nS_{XX}}\right)$$

$$V(\hat{\beta}_0)=\hat{\sigma}^2\frac{\sum X_i^2}{nS_{XX}}$$

11- Estimation of mean response variance

So: $V(\hat{y}_0)$

Proof:

$$\begin{aligned}\hat{y}_0 &= \hat{\beta}_0 + \hat{\beta}_1 X_0 \\ &= (\bar{y} - \hat{\beta}_1 \bar{X}) + \hat{\beta}_1 X_0 \\ &= \bar{y} + \hat{\beta}_1 (X_0 - \bar{X})\end{aligned}$$

Taking the variance of both sides, we get:

$$\begin{aligned}V(\hat{y}_0) &= V(\bar{y} + \hat{\beta}_1 (X_0 - \bar{X})) \\ &= V(\bar{y}) + \operatorname{cov}(\bar{y}, \hat{\beta}_1 (X_0 - \bar{X})) + (X_0 - \bar{X})^2 V(\hat{\beta}_1)\end{aligned}$$

We have: $cov(\bar{y}, \hat{\beta}_1) = 0$

$$V(\hat{y}_0) = V(\bar{y}) + 0 + (X_0 - \bar{X})^2 V(\hat{\beta}_1)$$

We have: $V(\hat{\beta}_1) = \frac{\hat{\sigma}^2}{S_{XX}}$ and $V(\bar{y}) = \frac{\hat{\sigma}^2}{n}$, so:

$$V(\hat{y}_0) = \frac{\hat{\sigma}^2}{n} + (X_0 - \bar{X})^2 \frac{\hat{\sigma}^2}{S_{XX}}$$

12-Covariance $cov(\bar{y}, \hat{\beta}_0)$

$$\operatorname{cov}(\bar{y}, \hat{\beta}_0)$$

Proof:

By equation

$$\operatorname{cov}(A, B - \alpha D) = \operatorname{cov}(A, B) - \alpha \operatorname{cov}(A, D)$$

So:

$$\begin{aligned}\operatorname{cov}(\bar{y}, \hat{\beta}_0) &= \operatorname{cov}(\bar{y}, \bar{y} - \bar{X}\hat{\beta}_1) \\ &= \operatorname{cov}(\bar{y}, \bar{y}) - \bar{X} \operatorname{cov}(\bar{y}, \hat{\beta}_1)\end{aligned}$$

We have $cov(\bar{y}, \hat{\beta}_1) = 0$

$$\operatorname{cov}(\bar{y},\hat{\beta}_0) = \operatorname{cov}(\bar{y},\bar{y}) - 0$$
$$\operatorname{cov}(\bar{y},\hat{\beta}_0) = \operatorname{cov}(\bar{y},\bar{y})$$
$$\therefore \operatorname{cov}(\bar{y},\hat{\beta}_0) = V(\bar{y})$$
$$\therefore \operatorname{cov}(\bar{y},\hat{\beta}_0) = \frac{\hat{\sigma}^2}{n}$$

13- Covariance $cov(\hat{\beta}_0, \hat{\beta}_1)$

$$\operatorname{cov}(\hat{\beta}_0,\hat{\beta}_1)$$

Proof:

$$\operatorname{cov}(\hat{\beta}_0,\hat{\beta}_1) = \operatorname{cov}(\bar{y} - \bar{X}\hat{\beta}_1, \hat{\beta}_1)$$
$$\operatorname{cov}(\hat{\beta}_0,\hat{\beta}_1) = \operatorname{cov}(\bar{y},\hat{\beta}_1) - \bar{X}\operatorname{cov}(\hat{\beta}_1,\hat{\beta}_1)$$
$$\operatorname{cov}(\hat{\beta}_0,\hat{\beta}_1) = 0 - \bar{X}V(\hat{\beta}_1)$$
$$\operatorname{cov}(\hat{\beta}_0,\hat{\beta}_1) = -\bar{X}\frac{\hat{\sigma}^2}{S_{XX}}$$

Example:

The following data represents the independent variable (X) which represents the age variable in years, and the dependent variable or response variable (y) which represents blood pressure, for ten people who were randomly selected. The data are shown in the table below:

y_i^2	X_i^2	$X_i y_i$	X_i	y_i	ت
12544	1225	3920	35	112	1
16384	1600	5120	40	128	2
16900	1444	4940	38	130	3
19044	1936	6072	44	138	4
24964	4487	10586	67	158	5
26244	4096	10368	64	162	6
19600	3481	8260	59	140	7
30625	4761	12075	69	175	8
15625	625	3125	25	125	9
20164	2500	7100	50	142	10
202094	**26157**	**71566**	**491**	**1410**	**المجموع**

Required:

1- Find the estimate of the regression parameters β_1, β_0.

2- Find the equation of the simple regression line.

3- What is the expected value (predictive value) of the blood pressure of an 80-year-old person.

4- Find the estimate of the population variance σ^2.

5- Find the estimate of the variance of each of the β_1, β_0.

Sol:

1- We find the estimate of the regression parameters β_1, β_0 as follows:

$$\hat{\beta}_1 = \frac{S_{Xy}}{S_{XX}} = \frac{\sum (x_i - \bar{X})(y_i - \bar{y})}{\sum (x_i - \bar{X})^2} = \frac{\sum X_i y_i - \frac{(\sum X_i)(\sum y_i)}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}$$

$$\hat{\beta}_1 = \frac{S_{Xy}}{S_{XX}} = \frac{\sum X_i y_i - \frac{(\sum X_i)(\sum y_i)}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}$$

$$\hat{\beta}_1 = \frac{71566 - \frac{(491)(1410)}{10}}{26157 - \frac{(491)^2}{10}} = 1.1396$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{X}$$

$$\hat{\beta}_0 = 141 - (1.1396)(49.1)$$

$$\hat{\beta}_0 = 85.043$$

2- We find the equation of the simple regression line.

$$\hat{y}_i = 85.043 + 1.1396 x_i$$

3- We find the expected value (predictive value) of blood pressure for an 80-year-old person.

$$\hat{y}_i = 85.043 + 1.1396(80)$$

$$\hat{y}_i = 176.2$$

4- We find an estimate of the population variance σ^2.