Lecture (٥): Variance and Covariance of Regression Parameters

١. Variance of the Intercept $\hat{\beta}_0$

$$\hat{\beta}_٠ = \bar{Y} - \hat{\beta}_١\bar{X}$$

$$\Longrightarrow V(\hat{\beta}_٠) = V(\bar{Y} - \hat{\beta}_١\bar{X})$$

$$= V(\bar{Y}) + V(-\bar{X}\hat{\beta}_١) + ٢\text{cov}(\bar{Y}, -\bar{X}\hat{\beta}_١)$$

$$V(\bar{Y}) = V\left(\frac{١}{n}\sum Y_i\right)$$

$$= V\left(\frac{١}{n}Y_١ + \frac{١}{n}Y_٢ + \ldots + \frac{١}{n}Y_n\right)$$

$$= \frac{١}{n^٢}V(Y_١) + \frac{١}{n^٢}V(Y_٢) + \ldots + \frac{١}{n^٢}V(Y_n) \;= \sum \frac{١}{n^٢}V(Y)$$

This is because the variance is homogeneous.

Where $V(Y_١) = V(Y_٢) = \cdots = V(Y_n) = \hat{\sigma}^٢$

$$= \sum \frac{١}{n^٢}V(Y)\,\hat{\sigma}^٢ = \frac{n}{n^٢}\hat{\sigma}^٢$$

$$\Longrightarrow V(\bar{Y}) = \frac{\hat{\sigma}^٢}{n}$$

$$\Longrightarrow V(\hat{\beta}_٠) = V(\bar{Y}) + (\bar{X})^٢ V(\hat{\beta}_١) - ٢\bar{X}\,\text{cov}\,(\bar{Y}, \hat{\beta}_١)$$

$$= \frac{\hat{\sigma}^٢}{n} + (\bar{X})^٢\,\frac{\hat{\sigma}^٢}{S_{xx}} - ٢\bar{X}\,\text{cov}(\bar{Y}, \hat{\beta}_١)$$

Now $\text{cov}(\bar{Y}, \hat{\beta}_١) = ?$

Let $a = a_١\, y_١ + a_٢\, y_٢ + \ldots + a_n\, y_n = \sum a_i\, y_i$

And $c = c_١\, y_١ + c_٢\, y_٢ + \ldots + c_n\, y_n = \sum c_i\, y_i$

We have y_i uncorrected with $y_{i'}$ $i \neq i'$ and $V(Y_١) = V(Y_٢) = \cdots = V(Y_n) = \hat{\sigma}^٢$

$$\text{cov}\,(a, c) = a_١\, c_١ v(y_١) + a_٢\, c_٢ v(y_٢) + \ldots + a_n\, c_n v(y_n) = (a_١\, c_١ + a_٢\, c_٢ + \cdots + a_n\, c_n)\hat{\sigma}^٢$$

$$= \left(\sum a_i\, c_i\right)\hat{\sigma}^٢$$

$$\bar{Y} = \sum \frac{١}{n} Y_i \longrightarrow a_i = \frac{١}{n}$$

$$\hat{\beta}_١ = \sum \frac{X_i - \bar{X}}{S_{xx}} Y_i \longrightarrow c_i \frac{X_i - \bar{X}}{S_{xx}}$$

$$\operatorname{cov}(a, c) = \operatorname{cov}(\bar{Y}, \hat{\beta}_١) = \left(\sum a_i\, c_i\right) \hat{\sigma}^٢$$

$$\therefore \operatorname{cov}(\bar{Y}, \hat{\beta}_١) = \left(\sum \frac{١}{n} \cdot \frac{X_i - \bar{X}}{S_{xx}}\right) \hat{\sigma}^٢$$

$$= \hat{\sigma}^٢ \left[\frac{١}{nS_{xx}} \sum (X_i - \bar{X})\right] = \text{صفر}$$

$$\therefore v(\hat{\beta}_٠) = \left(\frac{١}{n} + \frac{\bar{X}^٢}{S_{xx}}\right) \hat{\sigma}^٢ \quad = \left[\frac{S_{xx} + n\bar{X}^٢}{nS_{xx}}\right] \hat{\sigma}^٢ \quad = \left[\frac{\sum X_i^٢ - n\bar{X}^٢ + n\bar{X}^٢}{nS_{xx}}\right] \hat{\sigma}^٢$$

$$v(\hat{\beta}_٠) = \left[\frac{\sum X_i^٢}{nS_{xx}}\right] \hat{\sigma}^٢$$

٢. Variance of average response $\hat{Y}_0$

$$v(\hat{Y}_٠) = v[\bar{y} + \hat{\beta}_١ (X_٠ - \bar{X})]$$

$$V(\hat{Y}_٠) = V(\bar{y}) + V[\hat{\beta}_١ (X_٠ - \bar{X})] + ٢\operatorname{cov}[\bar{y}, \hat{\beta}_١ (X_٠ - \bar{X})]$$

Zero

$$= v(\bar{y}) + (X_٠ - \bar{X})^٢\, V(\hat{\beta}_١) + ٢(X_٠ - \bar{X}) \operatorname{Cov}(\bar{y}, \hat{\beta}_١)$$

$$= \frac{\hat{\sigma}^٢}{n} + (X_٠ - \bar{X})^٢ \times \frac{\hat{\sigma}^٢}{S_{xx}}$$

$$= \hat{\sigma}^٢ \left[\frac{١}{n} + \frac{(X_٠ - \bar{X})^٢}{S_{xx}}\right]$$

٣. Covariance Between Mean and Intercept Parameter $\operatorname{cov}(\bar{y}, \hat{\beta}_0)$

$$\operatorname{Cov}(\bar{y}, \hat{\beta}_٠) = \operatorname{Cov}(\bar{y}, \bar{y} - \bar{X}\hat{\beta}_١)$$

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

$$= \operatorname{cov}(\bar{y}, \bar{y}) - \bar{x} \operatorname{cov}(\bar{y}, \hat{\beta}_١)$$

$$= v(\bar{y}) - \text{صفر} = v(\bar{y}) = \frac{\hat{\sigma}^٢}{n}$$

$$\operatorname{cov}(\bar{y}, \hat{\beta}_٠) = v(\bar{y}) = \frac{\hat{\sigma}^٢}{n}$$

٤. Covariance Between Regression Parameter and Intercept Parameter $cov(\hat{\beta}_1 , \hat{\beta}_0)$

$$Cov(\hat{\beta}_{١}, \hat{\beta}_{.}) = cov(\hat{\beta}_{١}, \bar{y} - \bar{X}\hat{\beta}_{١})$$

$$= Cov(\bar{y}, \hat{\beta}_{١}) - \bar{X}Cov(\hat{\beta}_{١}, \hat{\beta}_{١})$$

$$= \text{صفر} - \bar{X}\, V(\hat{\beta}_{١}) = -\bar{X}\frac{\hat{\sigma}^{٢}}{S_{xx}}$$

Example (١): The following data represents the quantity supplied (Y) of a certain commodity and the price of one unit of it (X). Find the following:

a) Calculate the equation of the regression line Y/X
b) Find the expected value of the variable Y if you know that X=١١
c) Verify each of the following:
   ١. $\sum e_i = ٠$
   ٢. The sum of the weighted residuals with corresponding X_i values equals zero
   ٣. The sum of the weighted residuals with corresponding Y_i values equals the sum of the squares of the errors.
   ٤. The sum of the weighted residuals with corresponding $\hat{Y}^{\wedge}$ values equals zero
   ٥. $\bar{\hat{Y}} = \bar{Y}$
   ٦. The standard deviation of the regression Y/X

Solution:

X	Y	$X_i Y_i$	$X^2{}_i$	$Y^2{}_i$	$\hat{Y}_i$	$e_i = Y_i - \hat{Y}_i$	e_iX_i
٢	٣	٦	٤	٩	١,٩٧٩٦	١,٠٢٠٤٠	٢,٠٤٠٨
٤	١	٤	١٦	١	٣,٦٦٩٤	-٢,٦٦٩٤	-١٠,٦٧٧٦
٦	٧	٤٢	٣٦	٤٩	٥,٣٥٩٩	١,٦٤٠٨	٩,٨٤٤٨
٨	٥	٤٠	٦٤	٢٥	٧,٠٤٩	-٢,٠٤٩	-١٦,٣٩٢
١٠	٩	٩٠	١٠٠	٨١	٨,٧٣٨٨	٠,٢٦١٢	٢,٦١٢
٧	٨	٥٦	٤٩	٦٤	٦,٢٠٤١	١,٧٩٥٩	١٢,٥٧١
٣٧	٣٣	٢٣٨	٢٦٩	٢٢٩	٣٣,٠٠٠١	٠.٠٠٠١ ≅ ٠	−٠.٠٠٠٧ ≅ ٠

١. $\hat{Y}_i = \hat{\beta}_{.} + \hat{\beta}_{١}X_i$

$$\hat{\beta}_{١} = \frac{S_{xy}}{S_{xx}} \Longrightarrow S_{xy} = \sum X_i\, Y_i - \frac{(\sum X_i)(\sum Y_i)}{n}$$

$$= ٢٣٨ - \frac{(٣٧)(٣٣)}{٦} = ٣٤.٥$$

$$S_{xx} = \sum X_i^{٢} - \frac{(\sum X_i)^{٢}}{n}$$

$$= ٢٦٩ - \frac{(٣٧)^{٢}}{٦} = ٤٠.٨٤٤٩$$

$$\therefore \hat{\beta}_{١} = \frac{٣٤.٥}{٤٠.٨٤٤٩} = ٠.٨٤٤٩$$

$$\hat{\beta}_{.} = \bar{Y} - \hat{\beta}_{١}\bar{X} \Rightarrow \bar{Y} = \frac{\sum Y_i}{n} = \frac{٣٣}{٦} = ٥.٥$$

$$\bar{X} = \frac{\sum X_i}{n} = \frac{٣٧}{٦} = ٦.١٦٦٧$$

$$\hat{\beta}_{.} = ٥.٥ - ٠.٨٤٤٩(٦.١٦٦٧) = ٠.٢٨٩٨$$

$$\hat{Y}_{١} = ٠.٢٨٩٨ + ٠.٨٤٤٩(X_i)$$

$$\hat{Y}_i = ٠.٢٨٩٨ + ٠.٨٤٤٩(١١) = ٩.٥٨٣٧$$

$$\sum e_i = ٠$$

$$\sum e_i\, x_i = ٠$$

$e_i y_i$	e_i^2	$e_i \hat{y}$
٣,٠٦١٢	١,٠٤١٢	٢,٠١٩٩٨
-٢,٦٦٩٤	٧,١٢٥٧	-٩,٧٩٥٠٩
١١,٤٨٥٦	٢,٦٩٢٢	٨,٧٩٣٣٧
-١٠,٢٤٥	٤,١٩٨٤	-١٤,٤٤٣٤
٢,٣٥٠٨	٠,٠٦٨٢	٢,٢٨٢٥٧
١٤,٣٦٧٢	٣,٢٢٥٢	١١,١٤١٩
١٨,٣٥	١٨,٣٥	$-٠.٠٠٠٦ \cong ٠$

$$SSe = S_{yy} - \hat{\beta}_{١} S_{xy}$$

$$S_{yy} = \sum Y_i^{٢} - \frac{(\sum Y_i)^{٢}}{n}$$

$$= ٢٢٩ - \frac{(٣٣)^{٢}}{٦} = ٤٧.٥$$

$$\therefore SSe = ٤٧.٥ - ٠.٨٤٤٩(٣٤.٥) = ١٨.٣٥$$

Or from the tables $\sum e_i\, y_i = SSe = \sum e^{٢}{}_i$

$\sum e_i\, \hat{Y}_i = ٠$ *clear from the table*

$$\left.\begin{array}{l} \bar{Y} = \frac{\sum Y_i}{n} = \frac{٣٣}{٦} = ٥.٥ \\[2ex] \bar{\hat{Y}} = \frac{\sum \hat{Y}_i}{n} = \frac{٣٣.٠٠٠١}{٦} = ٥.٥ \end{array}\right\} \bar{\hat{Y}} = \bar{Y}$$

$v(\hat{Y}_i) = \frac{SSe}{n-٢} = \frac{١٨.٣٥}{٦-٢} = ٤.٥٨$ then $S(\hat{Y}_i) = \sqrt{٤.٥٨} = ٢.١٤$

Example (٢): if you know that $\sum(Y_i-\overline{Y})(X_i-\overline{X}) = ١٤١ \qquad \sum(X_i-\overline{X})^٢$

$$\sum Y^٢_i = ٣٥ \;,\; \overline{X} = ٢١ \;,\; \overline{Y} = ٣.٢$$

Find the equation of the regression line for this information and then verify that the point $X_i = ٥$ lie on the regression line

Note: If the equation properties are required to be applied, we must take at least four places after the comma.

Example (٣): If you have the following data (information):

$$n = ١١ \;, \overline{Y} = ٢٠ \;, \overline{X} = ٤ \;, \sum X_i^٢ = ٢١٠ \;, \sum X_i\,Y_i = ٩٢٠ \,, \sum Y_i^٢ = ٤٨٠٠$$

Find the following:

١. Variance and standard deviation of $\beta_١$ and $\beta_٠$
٢. Variance of Y
٣. Covariance cov $(\hat{\beta}_١, \hat{\beta}_٠)$
٤. Variance of mean response $\hat{Y}_٠$ when $X_٠ = ١٢$
٥. Response when $X_٠ = ٢$ and ٧
٦. Plot the equation of the regression line

Solution :

$$\hat{\sigma}^٢ = \frac{S_{yy} - \hat{\beta}_١^٢ S_{xx}}{n-٢}$$

$$\hat{\beta}_١ = \frac{S_{xy}}{S_{xx}}$$

$$S_{xy} = \sum X_i\,Y_i - \frac{\sum X_i \sum Y_i}{n} = ٩٢٠ - \frac{(٤٤)(٢٢٠)}{١١} = ٤٠$$

Where $\sum X_i = n\overline{x} = ١١(٤) = ٤٤$

$$\sum Y_i = n\overline{y} = ١١(٢٠) = ٢٢٠$$

$$S_{xx} = \sum X_i^{\,٢} - \frac{\sum X_i^{\,٢}}{n} = ٢١٠ - \frac{(٤٤)^٢}{١١} = ٣٤$$

$$S_{yy} = \sum Y_i^{\,٢} - \frac{\sum Y_i^{\,٢}}{n} = ٤٨٠٠ - \frac{(٢٢٠)^٢}{١١} = ٤٠٠$$

$$\therefore \hat{\beta}_١ = \frac{٤٠}{٣٤} = ١.١٧٦$$

$$\hat{\beta}_٠ = \overline{Y} - \hat{\beta}_١\overline{X} = ٢٠ - (١.١٧٦)(٤) = ١٥.٢٩٦$$

$$\therefore \hat{Y}_i = ١٥.٢٦٩ + ١.١٧٦(X_i)$$

$$\therefore \hat{\sigma}^٢ = \frac{٤٠٠ - (١.١٧٩)^٢(٣٤)}{١١ - ٢} = ٣٩.٢٢$$

$$S^٢(\hat{\beta}_١) = \frac{٣٩.٢٢}{٣٤} = ١.١٥٣$$

$$\therefore S(\hat{\beta}_١) = \sqrt{١.١٥٣} = ١.٠٧$$

$$S^٢(\hat{\beta}_.) = \hat{\sigma}^٢ \left[\frac{\sum X_i{}^٢}{nS_{xx}}\right] = \hat{\sigma}^٢ \left[\frac{١}{n} - \frac{\overline{X}^٢}{S_{xx}}\right] = ٣٩.٢٢ \left[\frac{١}{١١} + \frac{١٦}{٣٤}\right] = ٢٢.٠٢$$

$$\therefore S(\hat{\beta}_.) = \sqrt{٢٢.٠٢} = ٤.٦٩$$

$$V(\overline{y}) = \frac{\hat{\sigma}^٢}{n} = \frac{٣٩.٢٢}{١١} = ٣.٥٦$$

$$S(\overline{y}) = \sqrt{٣.٥٦} = ١.٨٨٦٧٩ \cong ١.٨٨٧$$

$$\text{cov}\left(\hat{\beta}_., \hat{\beta}_١\right) = -\overline{X}\frac{\hat{\sigma}^٢}{S_{xx}} = -\overline{X}.V(\hat{\beta}_١) = -٤(١.١٥٣) = -٤.٦١٤$$

$$X_. = ١٢$$

$$v(\hat{Y}_.) = \hat{\sigma}^٢ \left[\frac{١}{n} + \frac{(X_. - \overline{X})^٢}{S_{xx}}\right] = ٣٩.٢٢ \left[\frac{١}{١١} + \frac{(١٢ - ٤)^٢}{٣٤}\right]$$

$$= ٣٩.٢٢[٠.٠٩٠ + ١.٨٨٢] = ٧٦.٩٢$$

$$\hat{Y}_i = ١٥.٢٦٩ + ١.١٧٦(٢) = ١٧.٦٤$$

$$\hat{Y}_i = ١٥.٢٦٩ + ١.١٧٦(٧) = ٢٣.٥٢٨$$