Homework 03

Carolina Rutili de Lima - ID: 80975004h

March 8, 2022

1- First of all, let's consider:

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

$$S_{yy} = \sum_{i=1}^{n} = (y_i - \bar{y})^2$$

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

$$Var(y_{i)} = \sigma^2$$

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

a) Let's find out $Var(\beta_1)$ first:

$$Var(\beta_1) = Var(\frac{S_{xy}}{S_{xx}})$$

$$Var(\beta_1) = Var(\frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{S_{xx}})$$

$$Var(\beta_1) = \frac{Var(y_i)}{(S_{xx})^2}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$Var(\beta_1) = \frac{\sigma^2}{S_{xx}}$$

$$cov(\beta_1) = \frac{\sigma^2}{S_{xx}}$$

$$cov(\hat{\beta}_0, \hat{\beta}_1) = cov(\bar{y} - \hat{\beta}_1\bar{x}, \hat{\beta}_1)$$

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

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

$$= -\frac{\bar{x}}{S_{xx}}\sigma^2$$

b)

$$cov(\bar{y}, \hat{\beta}_1) = cov(\frac{\sum_{i=1}^{n} y_i}{n}, \frac{\sum(x_i - \bar{X})y_i}{S_{xx}})$$

$$= \frac{1}{nS_{xx}} cov(\sum_{i=1}^{n} y_i, \sum (x_i - \bar{X}) y_i)$$

$$= \frac{1}{nS_{xx}} cov(y_1 + ... + y_n, (x_1 - \bar{X}) y_1) + ... + (x_n - \bar{X}) y_n)$$

$$= \frac{1}{nS_{xx}} [cov(y_1, (x_1 - \bar{X}) y_1) + ... + cov(y_n, (x_n - \bar{X}) y_n)]$$

$$= \frac{\sum_{i=1}^{n} cov(y_i - \bar{X} y_i)}{nS_{xx}}$$

$$= \frac{\sum_{i=1}^{n} (x_i - \bar{X}) cov(y_i, y_i)}{nS_{xx}}$$

$$= \frac{(n\bar{X} - x\bar{X})\sigma^2}{nS_{xx}} = 0$$

2 -

$$SS_r = \sum_{i=1}^{n} (\hat{y_i} - \bar{y})^2$$

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

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

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

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

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

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

3 - Considering $H = X(X^T X)^{-1} X^T$

a) **H** is idempotent: it means that any power of itself is itself.

$$H^2 = (X(X^T X)^{-1} X^T)(X(X^T X)^{-1} X^T)$$

$$H^2 = (X(X^T X)^{-1} X^T)(X(X^T X)^{-1} X^T)$$

And,

$$I = (X^T X)^{-1} (X^T X)$$

Thus,

$$H^2 = X(X^T X)^{-1} X^T$$

$$H^2 = H$$

Now considering:

$$(H - I)(I - H) = I - H$$

$$I - H - H + H^2 = I - H$$

$$I - 2H + H^2 = I - H$$

And,

$$H^2 = H$$

Thus,

$$I - 2H + H = I - H$$

$$I - H = I - H$$

b)

$$\hat{Y} = X(X^T X)^{-1} X^T Y = HY$$

$$Var(\hat{Y}) = Var(HY)$$

$$Var(\hat{Y}) = HVar(Y)H^T$$

$$Var(\hat{Y}) = \sigma^2 H H^T$$

Considering that:

$$H^T = H$$

So,

$$Var(\hat{Y}) = \sigma^2 H H$$

Also considering that:

$$H^2 = H$$

Therefore,

$$Var(\hat{Y}) = \sigma^2 H$$

4 -

$$R^2 \equiv 1 - \frac{\mathrm{SS_{res}}}{\mathrm{SS_{tot}}}$$

For R^2 be larger than 1 or smaller than 0, $SS_{res} \leqslant SS_{tot}$, SS_{res} and SS_{tot} be both positive or both negative.

$$SS_{res} = \sum_i (y_i - f_i)^2$$

$$SS_{tot} = \sum_i (y_i - \bar{y})^2$$

SS_{tot} is the gaps between the y observed values and their mean.
At the end we can get two different scenarios:

1 $y = \bar{y}$, the model minimises the residual. Thus, $SS_{res} = SS_{tot}$;

2 The regression results in a different model, since the regression is producing a model that minimises residuals, the residuals of this model must be smaller than y, so $SS_{res} < SS_{tot}$.

Also, SS_{tot} and SS_{res} are given by the sum of squares, thus they are always positive.