15 October 2020

Coefficient of determination

$$R^2 = \frac{Reg\ S.S.}{Total\ S.S.} \qquad \text{(coefficient of determination)}$$

$$0 \leq R^2 \leq 1$$

$$\sqrt{R^2} \begin{cases} \text{simple linear reg. - simple correlation coeff. } (= r) \text{ (linear relationship between } y \text{ and } x) \\ \text{multple linear reg. - multiple correlation coeff. (linear relationship between } y \text{ and } \underset{\sim}{x}) \end{cases}$$

Added variable plot / Partial regression plot

Purposes:

1. Look for a graphical representation of $\hat{\beta}_{1,bi}$.
2. Give an explanation of regression coefficient in multiple linear regression.
3. Define partial correlation coefficient.

How to construct Added Variable Plot

1. The y coordinate, $\hat{e}_{Y(x_1)}$, is the residual for the model of y_i on x_{i2}, i.e., $\hat{e}_{Y(x_1),i} = y_i - \bar{y} - \hat{\delta}_1(x_{i2} - \bar{x}_2)$, where $\hat{\delta}_1 = \frac{S_{x_2y}}{S_{x_2x_2}}$ (Regress y on x_2 and the model is $y = \delta_0 + \delta_1 x_2 + e$).

2. The x coordinate, $\hat{e}_1$, is the residual for the model of x_{i1} on x_{i2}, i.e., $\hat{e}_{1,i} = x_{i1} - \bar{x}_1 - \hat{\gamma}_1(x_{i2} - \bar{x}_2)$, where $\hat{\gamma}_1 = \frac{S_{x_1x_2}}{S_{x_2x_2}}$ (Regress x_1 on x_2 and the model is $x_1 = \gamma_0 + \gamma_1 x_2 + e$).

3. It can be proved that the slope of the fitted line of $\hat{e}_{Y(x_1)}$ on $\hat{e}_1$ is equal to $\hat{\beta}_{1,bi}$, i.e., equal to
$$\frac{S_{x_1y}S_{x_2x_2} - S_{x_1x_2}S_{x_2y}}{S_{x_1x_1}S_{x_2x_2} - S^2_{x_1x_2}}\ .$$

Interpretation

1. $\beta_{1,bi}$ can be explained as the increase (or decrease) in y for each unit increase in x_1 after adjusted by x_2 (or, removing the effect of x_2) .

2. The simple correlation coefficient of $\hat{e}_{Y(x_1)}$ on $\hat{e}_1$ is defined as the partial correlation coefficient of y and x_1 after adjusted by x_2. Thus, partial corr. coeff. of y and x_1 on x_2
$$= \frac{r_{1y} - r_{12}r_{2y}}{\sqrt{1 - r^2_{2y}}\sqrt{1 - r^2_{12}}}$$