Esercizio 3

Siano β_0 e β_1 i coefficienti della retta di regressione, la loro stima è data dal metodo dei minimi quadrati e i suoi stimatori verrano indicati con $\hat{\beta_0}$ e $\hat{\beta_1}$. Si scelgono come stimatori quei particolari valori che minimizzano la somma dei quadrati delle differenze:

$$f(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2$$

Proposizione

La coppia dei valori $\hat{\beta_0}$, $\hat{\beta_1}$ che minimizza la funzione $f(\beta_0, \beta_1)$ e data da:

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

$$\hat{\beta_0} = \frac{\sum_{i=1}^{n} y_i - \hat{\beta_1} \sum_{i=1}^{n} x_i}{n} = \bar{y} - \hat{\beta_1}\hat{x}$$

Dimostrazione

$$f(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2$$

Si tratta di una funzione di 2 variabili per trovare il minimo del gradiente ∇=0, si procede con la derivata parziale della funzione rispetto a β_0:

$$\frac{df}{d\beta_0} = -2 \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))$$

e di conseguenza con la derivata parziale prima rispetto a β_1:

$$\frac{df}{d\beta_1} = -2 \sum_{i=1}^{n} x_i(y_i - (\beta_0 + \beta_1 x_i))$$

Si fa il modo che entrambe le derivate parziali prime cioè il gradiente ∇ sia uguale a 0:

$$\begin{cases} -2\sum_{i=1}^{n} y_i - \beta_0 - \beta_1 x_i = 0 \\ -2\sum_{i=1}^{n} x_i(y_i - (\beta_0 + \beta_1 x_i)) = 0 \end{cases}$$

$$\begin{cases} \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \beta_0 - \sum_{i=1}^{n} \beta_1 x_i = 0 \\ \sum_{i=1}^{n} x_i(y_i - \beta_0 + \beta_1 x_i)) = 0 \end{cases}$$

$$\begin{cases} \sum_{i=1}^{n} y_i - n\beta_0 - \beta_1 \sum_{i=1}^{n} x_i = 0 \\ \sum_{i=1}^{n} x_i(y_i - (\beta_0 + \beta_1 x_i)) = 0 \end{cases}$$

$$\begin{cases} n\beta_0 = \sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i(y_i - (\beta_0 + \beta_1 x_i)) = 0 \end{cases}$$

$$\begin{cases} \beta_0 = \frac{1}{n}\sum_{i=1}^{n} y_i - \beta_1 \frac{1}{n}\sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i(y_i - (\beta_0 + \beta_1 x_i)) = 0 \end{cases}$$

$$\begin{cases} \beta_0 = \bar{y} - \beta_1 \bar{x} \\ \sum_{i=1}^{n} x_i(y_i - (\beta_0 + \beta_1 x_i)) = 0 \end{cases}$$

$$\begin{cases} \beta_0 = \bar{y} - \beta_1 \bar{x} \\ \sum_{i=1}^{n} x_i(y_i - (\bar{y} - \beta_1 \bar{x}) - \beta_1 x_i) = 0 \end{cases}$$

$$\begin{cases} \beta_0 = \bar{y} - \beta_1\bar{x} \\ \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i(\bar{y} - \beta_1\bar{x}) - \beta_1 \sum_{i=1}^n x_i^2 = 0 \end{cases}$$

$$\begin{cases} \beta_0 = \bar{y} - \beta_1\bar{x} \\ \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i\bar{y} + \sum_{i=1}^n x_i\beta_1\bar{x} - \beta_1 \sum_{i=1}^n x_i^2 = 0 \end{cases}$$

$$\begin{cases} \beta_0 = \bar{y} - \beta_1\bar{x} \\ \sum_{i=1}^n x_i\beta_1\bar{x} - \beta_1 \sum_{i=1}^n x_i^2 = -\sum_{i=1}^n x_i y_i + \sum_{i=1}^n x_i\bar{y} \end{cases}$$

$$\begin{cases} \beta_0 = \bar{y} - \beta_1\bar{x} \\ \beta_1(\sum_{i=1}^n x_i\bar{x} - \sum_{i=1}^n x_i^2) = \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i\bar{y} \end{cases}$$

$$\begin{cases} \beta_0 = \bar{y} - \beta_1\bar{x} \\ \beta_1 = \frac{\sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i\bar{y}}{\sum_{i=1}^n x_i\bar{x} - \sum_{i=1}^n x_i^2} \end{cases}$$

$$\begin{cases} \beta_0 = \bar{y} - \beta_1\bar{x} \\ \beta_1 = \frac{\sum_{i=1}^n x_i(y_i - \bar{y})}{\sum_{i=1}^n x_i(x_i - \bar{x})} \end{cases}$$

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} \\ \hat{\beta}_1 = \frac{\sum_{i=1}^n x_i(y_i - \bar{y}) - \sum_{i=1}^n \bar{x}(y_i - \bar{y})}{\sum_{i=1}^n x_i(x_i - \bar{x}) - \sum_{i=1}^n \bar{x}(x_i - \bar{x})} \end{cases}$$

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} \\ \hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \end{cases}$$

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \quad \square$$