Exercises: Simple Linear Regression

Valentina Zangirolami - valentina.zangirolami@unimib.it

November 23, 2023

(Referring to the theoretical parts: 1, 2, 3, 6)

1 Mother and Daughter heights data

Let consider a sample of data with $n = 11$ observations (Table 1) with two variables:

- **mother's height** x (independent variable);
- **daughter's height** y (dependent variable).

Table 1: Mother and Daughter heights data: data are expressed in centimeters.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
x	153.7	156.7	173.5	157.0	161.8	140.7	179.8	150.9	154.4	162.3	166.6
y	163.1	159.5	169.4	158.0	164.3	150.0	170.3	158.9	161.5	160.8	160.6

We would like to find out if there exists a relationship between these two variables.

Exercise 1.1
Starting from the data (in Table 1), write the equation of the simple linear regression model. Compute $\bar{x}$, $\bar{y}$, $\sum_{i=1}^{n} x_i y_i$, $\sum_{i=1}^{n} x_i^2$ and, then, find the estimates of the linear model parameters.

Exercise 1.2

Given the results of the previous exercise, compute the fitted values for each i. Make a plot involving the observations of the couple (y, x) and the estimated regression line.

Exercise 1.3

Compute the residuals (e_i) and the unbiased estimate (s^2) of the variance σ^2. Then, find the estimates of the variances of $\hat{\beta}_1$ and $\hat{\beta}_2$. (You can also use the notation of the notes part 6: $\sum_{i=1}^{n} e_i^2 = \ldots$)

Exercise 1.4

Compute the total sum of squares (SST), the residual sum of squares (SSE) and the regression sum of squares (SSR). Then, find the coefficient of determination R^2. Compute the correlation coefficient r_{xy} and its squared. What happens in this case?