Simple regression using Least Squares Method

Lucas V. Araujo <lucas.vieira.ar@disroot.org>

December 1, 2019

Part I
Linear Regression

Given N known points, by the Least Squares Method, the linear function that best describes their growth is given by

$$\begin{cases} a\sum_{i=1}^{n} 1 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \\ a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} {x_i}^2 = \sum_{i=1}^{n} x_i y_i \end{cases}$$

What can be written as

$$\begin{cases} an + b\sum x = \sum y & (r_1) \\ a\sum x + b\sum x^2 = \sum xy & (r_2) \end{cases}$$

Starting by r_1, we isolate A:

$$a = \frac{\sum y - b\sum x}{n} \tag{1}$$

Now we can replace A on r_2, to find B:

$$(\frac{\sum y - b\sum x}{n})\sum x + b\sum x^2 = \sum xy$$

Simplifying, and isolating B, we will get

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} \tag{2}$$

Equation (2) allows to get the value of B directly.
Now, using r_1, we isolate B:

$$b = \frac{\sum y - an}{\sum x} \tag{3}$$

Then, we replace B on r_2 to find A:

$$a\sum x + \left(\frac{\sum y - an}{\sum x}\right)\sum x^2 = \sum xy$$

Finally, we simplify this, isolating A, finding a equation to directly get A:

$$a = \frac{\sum xy - \sum y \sum x^2}{\sum x + n \sum x^2} \tag{4}$$

Now, by using equations (2) and (4) we can find the coefficients to a linear function y=ax+b, thats describes (in the best way possible) the growth of the given points.

Part II

Logarithmic Regression

Same as linear regression, but instead of using x, we will use $\ln(x)$. And the function will now be $y = a\ln(x) + b$.

Part III

Exponential Regression

Same as linear regression, but instead of using y, we will use ln(y). And when we find the coefficient B, it actually will be the ln(b), so we just need to calculate ℓ^b. The function will be $y = b * \ell^{ax}$.

Part IV

Potence regression

Again, the same as linear regression, but instead of y we, use ln(y) and instead of x, we use ln(x). The value found to B is actually the ln(b), so we get the real B calculating ℓ^b. The function will be bx^a.