Least Squares Linear Regression Equations

Dave Rosenman

August 20, 2017

If you have taken n pairs of measurements $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$, the mean value of x is by definition:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

and the mean value of y is

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

The slope of the best fit line, m is given by:

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

The y-intercept, c, is given by:

$$c = \bar{y} - m\bar{x}$$

The standard error in the slope, Δm, is:

$$\Delta m = \sqrt{\frac{1}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\frac{\sum_{i=1}^{n}(y_i - mx_i - c)^2}{n-2}}$$

The standard error in the y intercept, Δc is:

$$\Delta c = \sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)\frac{\sum_{i=1}^{n}(y_i - mx_i - c)^2}{n-2}}$$

If the best fit is required to pass through the origin, $(0,0)$, then $c = 0$, and

$$m = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

and the standard error of the slope, Δm is

$$\Delta m = \sqrt{\frac{1}{\sum_{i=1}^{n} x_i^2}\frac{(y_i - mx_i)^2}{n-1}}$$