Formulas and Calculations Simple Linear Correlation and Regression

Rossi A. Hassad, PhD, MPH

Pearson' Correlation Coefficient (r)

NOTE: For the assignment, please organize your calculations in column format rather than horizontally, as shown here.

N = 3,as there are three pairs of values (x and y).

Example (Pearson's Correlation Coefficient)

X	Y
1	2
2	5
3	6

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{N}\right)\left(\sum Y^2 - \frac{(\sum Y)^2}{N}\right)}}$$

$\Sigma XY = (1)(2) + (2)(5) + (3)(6) = 30$

$\Sigma X = 1 + 2 + 3 = 6$

$\Sigma X^2 = 1^2 + 2^2 + 3^2 = 14$

$\Sigma Y = 2 + 5 + 6 = 13$

$\Sigma Y^2 = 2^2 + 5^2 + 6^2 = 65$

$N=3$

$\Sigma XY - \Sigma X \Sigma Y / N = 30 - (6)(13)/3 = 4$

$\Sigma X^2 - (\Sigma X)^2/N = 14 - 6^2/3 = 2$

$\Sigma Y^2 - (\Sigma Y)^2/N = 65 - 13^2/3 = 8.667$

$r = 4 / \sqrt{(2)(8.6667)} = 4/4.16333$

$= .9608$

Simple Linear Regression Model
Y = bx + a

X	Y	X- Mean X	$(X- \text{Mean } X)^2$	Y- Mean Y	(X-Mean X) (Y-Mean Y)
1	2	-1	1	-2.33	2.33
2	5	0	0	.67	0
3	6	1	1	1.67	1.67
Mean = 2	**Mean = 4.33**		**$\sum$ (sum) = 2**		**$\sum$(sum) = 4**

$$b = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{\Sigma(x - \bar{x})^2}$$

b = 4/2 = 2

$$a = \bar{y} - b\bar{x}$$

a = mean of Y – b(mean of X)

a = 4.33 - 2(2)

a = 4.33 - 4 = .33

The regression equation: Y = 2x + .33