Pearson Correlation Coefficient 공식 $= \frac{Cov(X,Y)}{\sigma x \sigma y}$ 이다.

1. $Cov(X,Y) = \sum_{i=1}^{n} (x_i - \bar{X})(Y_i - \bar{Y}) \times \frac{1}{n}$

$= \sum_{i=1}^{n} (X_i Y_i - \bar{X} Y_i - \bar{Y} X_i + \bar{X}\bar{Y}) \times \frac{1}{n}$

$= (\sum_{i=1}^{n} X_i Y_i - \bar{X} \sum_{i=1}^{n} Y_i - \bar{Y} \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} \bar{X}\bar{Y}) \times \frac{1}{n}$

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ 이므로,

$= (\sum_{i=1}^{n} X_i Y_i - \frac{1}{n}\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i - \frac{1}{n}\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i$

$+ \frac{1}{n}\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i) \times \frac{1}{n}$

$= (n\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i) \times \frac{1}{n^2}$

2. $\sigma x = \sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2} \times \sqrt{\frac{1}{n}}$

$= \sqrt{\sum_{i=1}^{n} X_i^2 + \sum_{i=1}^{n} \bar{X}^2 - \sum_{i=1}^{n} (2\bar{X}X_i)} \times \sqrt{\frac{1}{n}}$

$= \sqrt{\sum_{i=1}^{n} X_i^2 + \frac{1}{n}(\sum_{i=1}^{n} X_i)^2 - \frac{2}{n}(\sum_{i=1}^{n} X_i)^2} \times \sqrt{\frac{1}{n}}$

$= \sqrt{n\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2} \times \frac{1}{n}$

3. $\sigma y = \sqrt{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} \times \sqrt{\frac{1}{n}}$, 2와 같은방법으로 정리하면

$= \sqrt{n\sum_{i=1}^{n} Y_i^2 - (\sum_{i=1}^{n} Y_i)^2} \times \frac{1}{n}$

따라서 $r_{XY} = \frac{\frac{\sum_i^n (X_i - \bar{X})(Y_i - \bar{Y})}{n}}{\sqrt{\sum_i^n (X_i - \bar{X})^2}\sqrt{\sum_i^n (Y_i - \bar{Y}}} = \frac{\frac{n\sum xy - \sum x \sum y}{n^2}}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$

$= \frac{n \, n(\sum xy) - \sum x \sum y \quad n^2}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$ 임을 확인할 수 있다