과제1: Pearson Correlation Coefficient 함수

Pearson Correlation Coefficient 함수

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

가 아래의 수식과 동일한 표현이라는 것을 보이시오.

$$r_{XY} = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

2011년 6월 22일

$$r_{xy} = \frac{n\sum (x_i-\bar{x})(y_i-\bar{y})}{n\sqrt{\sum (x_i-\bar{x})^2}\sqrt{\sum (y_i-\bar{y})^2}}$$

$$= \frac{\sum (x_i y_i) - \sum (x_i\bar{y} + \bar{x}y_i - \bar{x}\bar{y})}{\sqrt{\sum (x_i-\bar{x})^2 \sum (y_i-\bar{y})^2}} \quad \cdots ① \quad \cdots ②$$

$$① \times n = \left(\sum (xy) - \sum (x_i\bar{y}) - \sum (\bar{x}y_i) + \sum (\bar{x}\bar{y})\right) \times n$$

$$= n\sum (xy) - n\times \bar{y}\sum x_i - n\cdot\bar{x}\sum y_i + n^2\bar{x}\bar{y}$$

$$= n\sum (xy) - \sum y \sum x - \sum x \sum y + \sum x \sum y$$

$$= n\sum (xy) - \sum x \sum y$$

$$② \times n = n\sqrt{\left\{\sum (x_i-\bar{x})^2\right\} \times \left\{\sum (y_i-\bar{y})^2\right\}}$$

$$= n\sqrt{\left[\sum (x_i^2 - 2\bar{x}x_i + \bar{x}^2)\right] \times \left[\sum (y_i^2 - 2y_i\bar{y} + \bar{y}^2)\right]}$$

$$= n\sqrt{\left[\sum x^2 - 2\bar{x}\sum x + \bar{x}^2 \cdot n\right] \times \left[\sum y^2 - 2\bar{y}\sum y + n\cdot\bar{y}^2\right]}$$

$$= \sqrt{\left(n\Sigma x^2 - 2(\Sigma x)^2 + (\Sigma x)^2\right)\left[n\Sigma y^2 - 2(\Sigma y)^2 + (\Sigma y)^2\right]}$$

$$= \sqrt{\left[n\Sigma x^2 - (\Sigma x)^2\right]\left[n\Sigma y^2 - (\Sigma y)^2\right]}$$

$$\frac{① \times n}{② \times n} = \frac{①}{②}$$

$$\therefore \quad \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma (x_i - \bar{x})^2}\sqrt{\Sigma (y_i - \bar{y})^2}}$$

$$= \frac{n\Sigma(xy) - \Sigma x \Sigma y}{\sqrt{\left[n\Sigma x^2 - (\Sigma x)^2\right]\left[n\Sigma y^2 - (\Sigma y)^2\right]}}$$