고급소프트웨어실습1

2주차 과제

20161663 허재성

Pearson Correlation Coefficient 함수

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

가 아래 수식과 동일한 표현이라는 것을 보이시오.

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

먼저 분모의 $\sqrt{\frac{\sum_i^n (X_i - \bar{X})^2}{n}}$ 를 바꾸기 위해 괄호의 제곱식을 전개한다.

$$\sqrt{\frac{\sum_i^n (X_i{}^2 - 2\bar{X}X_i + \bar{X}^2)}{n}} = \sqrt{\frac{\sum_i^n X_i{}^2 - 2\bar{X}\sum_i^n X_i + \sum_i^n \bar{X}^2}{n}} =$$

$$\sqrt{\frac{\sum_i^n X_i{}^2}{n} - \bar{X}^2} = \sqrt{\frac{\sum_i^n X_i{}^2}{n} - \left(\frac{\sum_i^n X_i}{n}\right)^2} = \sqrt{\frac{n\sum_i^n X_i{}^2 - \left(\sum_i^n X_i\right)^2}{n^2}}$$ 이다.

같은 방법으로 분모의 $\sqrt{\frac{\sum_i^n (Y_i - \bar{Y})^2}{n}} = \sqrt{\frac{n\sum_i^n Y_i{}^2 - \left(\sum_i^n Y_i\right)^2}{n^2}}$ 이다.

다음으로 분자의 $\frac{\sum_i^n (X_i - \bar{X})(Y_i - \bar{Y})}{n}$ 를 바꾸기 위해 식을 전개한다.

$$\frac{\sum_i^n (X_i Y_i - \bar{Y}X_i - \bar{X}Y_i + \bar{X}\bar{Y})}{n} = \frac{\sum_i^n X_i Y_i}{n} - \bar{X}\bar{Y} = \frac{\sum_i^n X_i Y_i}{n} -$$

$$\frac{\sum_i^n X_i}{n}\frac{\sum_i^n Y_i}{n} = \frac{n\sum_i^n X_i Y_i - \sum_i^n X_i \sum_i^n Y_i}{n^2}$$ 이다.

분모, 분자에 각각의 식을 대입하여 정리하면 결과식을 얻을 수 있다.