

22/27 challenges solved

Points: 22





D D

Day 7: Pearson Correlation Coefficient I ★

Problem Submissions Leaderboard Editorial ← Tutorial

Covariance

This is a measure of how two random variables change together, or the strength of their correlation.

Consider two random variables, X and Y, each with n values (i.e., x_1 , x_2 , ..., x_n and y_1 , y_2 , ..., y_n). The covariance of X and Y can be found using either of the following equivalent formulas:

$$extstyle{\mathsf{cov}}(X,Y) = rac{1}{n} \sum_{i=1}^n (x_i - ar{x}) \cdot (y_i - ar{y})$$

$$\texttt{cov}(X,Y) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{1}{2} (x_i - x_j) \cdot (y_i - y_j) = \frac{1}{n^2} \sum_i \sum_{j>i} (x_i - x_j) \cdot (y_i - y_j)$$

Here, $ar{x}$ is the mean of X (or μ_X) and $ar{y}$ is the mean of Y (or μ_Y).

Pearson Correlation Coefficient

The Pearson correlation coefficient, $ho_{\pmb{X},\pmb{Y}}$, is given by:

$$ho_{X,Y} = rac{ extstyle extsty$$

Here, σ_X is the standard deviation of X and σ_Y is the standard deviation of Y. You may also see $\rho_{X,Y}$ written as $r_{X,Y}$.