

データ解析基礎レポート課題（第2回）

出題 10月10日（金）、提出期限 10月20日（月）

提出先：情報システム学研究科棟2階事務室脇のポスト

1. 以下の表は、「所属するクラブの数(X)」と「成績(Y)」の関係を6人の中学生について調べたデータである。このデータについて、以下の問いに答えよ。

クラブ数 (X)	5	3	1	3	2	4
成績 (Y)	4.0	3.0	0.5	2.5	1.5	3.5

- X と Y の平均と標準偏差をそれぞれ求めよ。
- X と Y の共分散を求めよ。
- X と Y の相関係数を定義に従って求めよ。
- 次の問題に示された公式を用いて相関係数を求め、cの結果と一致することを確認せよ。

2. Pearson の相関係数を求める公式

$$r_{XY} = \frac{N(\sum X_i Y_i) - (\sum X_i)(\sum Y_i)}{\sqrt{N(\sum X_i^2) - (\sum X_i)^2} \sqrt{N(\sum Y_i^2) - (\sum Y_i)^2}}$$

を、もとの定義

$$r_{XY} = \frac{S_{XY}}{S_X S_Y}$$

から導出せよ。

3. 第2回の講義の内容で理解できなかったこと、また、改善してほしいことなどについて、自由に記述せよ（なければ、特に記載しなくてよい）。