

Fundamentals of Data Analysis: Assignment #2

Deadline: 10/20/2003 (Monday)

Please post to the mailbox next to the IS management office (2nd floor of IS building)

1. A teacher tries to determine, for his 6 students in his class, if there is a relationship between the number of club to which a student belongs to (X) and his/her grade (Y). The data are given below. Answer the following questions.

Number of Club (X)	5	3	1	3	2	4
Grade (Y)	4.0	3.0	0.5	2.5	1.5	3.5

- Calculate the means and standard deviations of X and Y.
- Calculate the covariance between X and Y.
- Compute the Pearson's correlation coefficient between X and Y, according to its definition.
- Compute the correlation coefficient according to the formula indicated in the next question and check whether the answer agrees with that to the previous question.

2. Induce the following formula for the Pearson's correlation coefficient

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

from its original definition

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

3. Please feel free to write your comments and requests on this lecture (if any).