Two-Dimensional Data Analysis Quiz

Multiple-choice question (Each question has at least one correct answer.) ١.) 1. There are 45 students in a class. The average score and the standard (devastation of the midterm are 60 points and 8 points. The teacher adjusted the scores by multiplying each student's score by 1.5 and then subtracting 10 points. Let the variables x and y represent the original and adjusted scores of these 45 students, respectively. What is the range of the correlation coefficient r between variables xand y?

(1) r < 0(2) r = 0(3) $0 < r \le 0.3$ (4) $0.3 < r \le 0.7$ (5) r=1Ans: (5)

)2. The following scatter plots show the number of children aged 1 to (10(integer count) in five regions with a certain disease. Identify the option where the correlation coefficient between the number of children with the disease and their age is the highest.



Ans: (4)

()3. Suppose judges A, B, C, and D provide scores for 6 contestants as shown in the attached table. The terms r_{AB} , r_{AC} , r_{AD} represent the correlation coefficients between A and B, A and C, and A and D respectively. Which of the following expressions is correct?

Α	4	8	4	6	6	2
В	3	7	3	5	5	1
С	2	4	2	3	3	1
D	4	8	4	6	6	1

(1) $r_{AB} > r_{AC} > r_{AD}$ (4) $r_{AB} = r_{AC} = r_{AD}$ Ans: (2)

(2) $r_{AB} = r_{AC} > r_{AD}$

- (5) $r_{AB} < r_{AC} < r_{AD}$
- (3) $r_{AB} > r_{AC} = r_{AD}$

()4. Points A, B, C, D, and E are located on the coordinate plane. The regression line of these five points is parallel to line L in the diagram. Points A, B, and C are in the upper half-plane of line L, while points D and E are in the lower half-plane of line L. The perpendicular distances from these 5 points to line L are 4, 3, 1, 1, and 2 respectively. Which point among A, B, C, D, and E will lie on the regression line?



5. A research analysis of data on "annual salary" and "years of education" for employees with similar tenure at a company, concludes: "For each additional year of education, the average annual salary increases by 60,000 NTD." From which of the following statistical measures can the above conclusion be directly derived?

(1) Mode of "annual salary" and mode of "years of education"

(2) Range of "annual salary" and range of "years of education"

(3) Mean of "annual salary" and mean of "years of education"

(4) Correlation coefficient between "annual salary" and "years of education"

(5) Slope of the regression line of "annual salary" on "years of education"

Ans: (5)

(1) Point A

Ans: (3)

II. Other question

1. For the following data sets, plot a scatter diagram and describe the correlation.

х	1	2.4	3.6	2.2	4.3	3.3	4.0	0.6
У	6.0	9.0	15.8	7.1	18.6	12.1	15.0	3.7



2. (1) Given 10 data points $x_1 + x_2 + ... + x_{10} = 18.5$, $x_1^2 + x_2^2 + ... + x_{10}^2 = 36$, find S_{xx} Ans: 1.775

(2) Given 10 data points $y_1 + y_2 + ... + y_{10} = 25.7$, $y_1^2 + y_2^2 + ... + y_{10}^2 = 140$, find S_{yy} Ans: 73.951

(3) Given $x_1 + x_2 + ... + x_5 = 15$, $y_1 + y_2 + ... + y_5 = 35$, $x_1y_1 + x_2y_2 + ... + x_5y_5 = 91$, find $S_{yy} = ?$

Ans: 14

(4) Given $x_1 + x_2 + ... + x_6 = 367$, $y_1 + y_2 + ... + y_6 = 270$, $x_1^2 + x_2^2 + ... + x_6^2 = 33845$, $y_1^2 + y_2^2 + ... + y_6^2 = 12976$, $x_1y_1 + x_2y_2 + ... + x_6y_6 = 17135$, n = 6Find the correlation coefficient(r) of x and y

Ans: 0.93

3. A detector at a certain intersection collects data on traffic flow (x) (number of motorcycles passing per minute) and dust particle concentration in the air (y), as shown in the attached table:

x	15	18	24	27	33	27
у	10	16	16	24	28	26

- (1) Find the best-fitting line for y as a function of x. Ans: y = x 4
- (2) Find the predicted value of dust particle concentration in the air when 20 motorcycles pass through. Ans:16

4. For a set of two-dimensional data (x_i, y_i) , the correlation coefficient is 0.8 and the best-fitting line for y as a function of x is y = 2x+5. It knows that the mean of variable x is 6 with a standard deviation of 5.

Find the (1) mean of variable y and (2) the standard deviation of y. Ans:(1) 17, (2)12.5

5. Given four data points (x_i, y_i) as: (3,1), (5,3), (3,2), (1,t). If the best-fitting line for y as a function of x is $y = \frac{1}{4}t + \frac{5}{4}$, then t = ? Ans:2

製作者:國立臺灣師範大學附屬高級中學 蕭煜修