Unit 10 · Correlation and Regression

WORKSHEET

Full marks 27

Name _____ Class _____ Roll no. _____ Date _____ Marks _____ / 27

1 Key words [4]

Fill in the blanks with words from the box. Two words in the box are not needed. Each blank carries 0.5 mark.

bivariate	**coefficient**	**covariance**	**curvilinear**	**least squares**	**means**	**negative**	**rank**
scatter diagram	**straight line**						

Two numbers measured on each thing (a station's height and its temperature) make a _____ set of pairs. Plotted as points, they make a _____. When larger x goes with smaller y, the correlation is _____. Karl Pearson's coefficient r measures only how close the points are to a _____. When only an order can be given, Spearman's _____ correlation coefficient is used. The line that makes the sum of the squares of the misses least is the _____ line; its slope is the regression _____. Both regression lines pass through the point of the _____.

2 True or false [3]

Write ✓ for a true statement and ✗ for a false one. Correct the false ones.

i. $r = 0$ means that y does not depend on x in any way. ☐

ii. Writing every height in kilometers instead of meters leaves r unchanged. ☐

iii. To estimate x from a given y, solve the regression line of y on x for x. ☐

iv. The two regression coefficients always have the same sign as r. ☐

v. A high correlation between two quantities shows that one causes the other. ☐

vi. When $r = \pm 1$, the two regression lines are one line. ☐

3 Multiple choice [5]

Choose the correct option for each of the following questions.

i. The regression coefficient of y on x is: [1]

a. $r\dfrac{\sigma_y}{\sigma_x}$ **b.** $r\dfrac{\sigma_x}{\sigma_y}$

c. $\dfrac{\sigma_y}{\sigma_x}$ **d.** $r\sigma_x\sigma_y$

ii. For some data, $r = -0.7$. Which pair can be the two regression coefficients b_{yx} and b_{xy}? [1]

a. 0.7 and 0.7 **b.** −0.98 and 0.5

c. −0.98 and −0.5 **d.** −1.4 and −0.5

iii. Five pairs have $\Sigma x = 15$, $\Sigma y = 40$, $\Sigma xy = 136$, $\Sigma x^2 = 55$ and $\Sigma y^2 = 360$. Then r is: [1]

a. 0.16 **b.** 0.4

c. 0.8 **d.** 1.6

iv. For some data, $r = 0.6$, $\sigma_x = 2$ and $\sigma_y = 4$, and the means are $\bar{x} = 10$, $\bar{y} = 30$. The estimate of y when $x = 15$ is: [1]

a. 31.5 **b.** 36

c. 42 **d.** 15

v. In some data every x is doubled and 5 is added to every y. The regression coefficient b_{yx}: [1]

a. stays the same **b.** doubles

c. becomes half **d.** increases by 5

4 Short answers [7]

Write short answers to the following questions.

a. Explain why the two regression coefficients of one set of data can't both be more than 1 in size. [2]

b. Over a year in a town, the number of umbrellas sold each month and the number of road accidents each month gave $r = 0.85$. Pratik says that selling fewer umbrellas would cut the accidents. Is he right? Give a reason. [2]

c. Two judges ranked six dancers in a competition: [3]

Dancer	A	B	C	D	E	F
Judge 1	1	2	3	4	5	6
Judge 2	3	1	2	6	4	5

Find Spearman's rank correlation coefficient and say what it shows.

Science Playground, Katunje, Bhaktapur
https://science-playground.com/learn/grade-12-math/correlation-and-regression

5 Long answer [5]

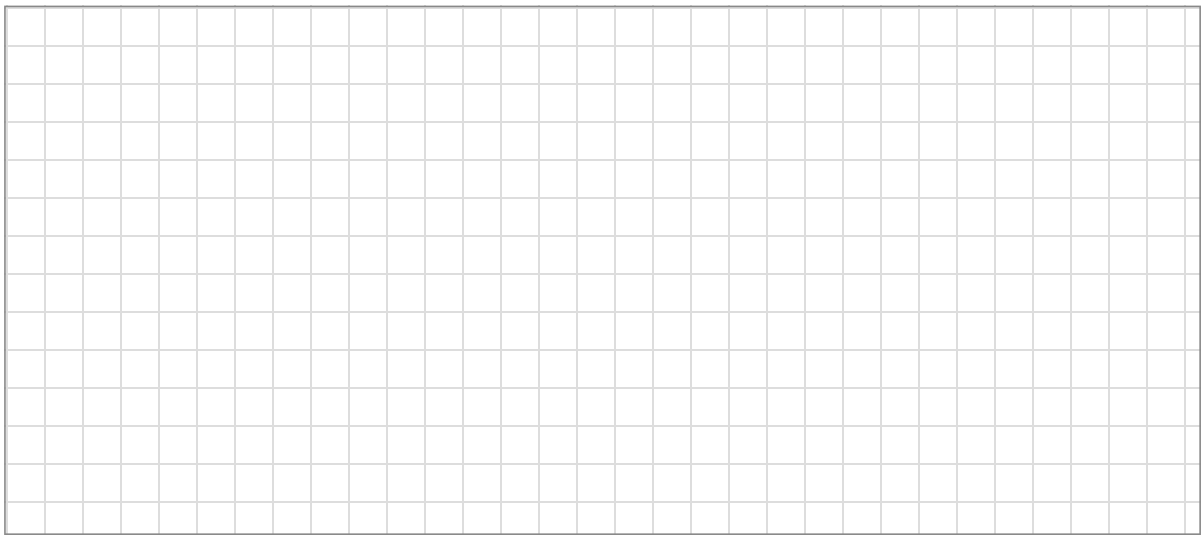
Answer the following question. Show your working.

Five students from a typing class wrote down the hours they practised in a week (x) and their typing speed in words a minute (y):

x	**2**	**4**	**5**	**6**	**8**
y	20	26	31	34	39

a. Find Karl Pearson's correlation coefficient between x and y. [3]

b. Find the regression line of y on x, and estimate the speed of a student who practises 7 hours a week. [2]



Answer: _____

6 Think [3]

The regression line of temperature y (°C) on height x (in hundreds of meters) for Nepal's 15 weather stations is $y = 25.45 - 0.487x$. The highest station is at about 2,400 m. Sunil uses the line to find the average year's temperature at the top of Everest, 8,849 m. What does the line give, and why should he not trust it?

7 Activity: Arm span and height

(not marked)

You need: a measuring tape and six friends. Measure each friend's height and arm span (finger tip to finger tip, arms stretched out) in centimeters.

Friend	1	2	3	4	5	6
Height x (cm)						
Arm span y (cm)						

Draw a scatter diagram on the back of this sheet. Find r, and the regression line of y on x. Is a person's arm span about the same as their height?