S.9 Interpreting the Regression Equation

The regression equation shows us the relationship between the two variables being graphed. Its components give us additional information.

gradient: rate of change

y-intercept: initial value.

Sometimes the mathematical interpretation of one of these components has no valid interpretation in the real-life context of the problem.

Example 1:

A study was done to investigate the relationship between the age x in years of a young person and the time y in minutes in which the child can run one kilometer. Data from children between the ages of 7 and 18 were collected. The equation of the regression line was found to be $y = 20 - \frac{1}{2}x$.

Interpret the gradient and the y-intercept.

time decrease

$\frac{1}{2}$ min. each yr.

time when 0 yrs.

Example 2:

A biologist wants to study the relationship between the number of trees x per hectare and the number of birds y per hectare. She calculates the equation of the regression line to be $y = 8 + 5.4x$.

Interpret the ~~gradient~~ and the y-intercept.

birds if 0 trees

for ea. tree +5.4 birds

In order to receive full credit when drawing a regression line on a scatter plot, the line must be drawn through both the mean point of the data $(\bar{x}, \bar{y})$ and the y-intercept of the regression equation.

Example 3:

A shopkeeper wanted to investigate whether or not there was a correlation between the prices of food in 1992 and 2002. He chose 8 everyday items. Their prices are given in the table below.

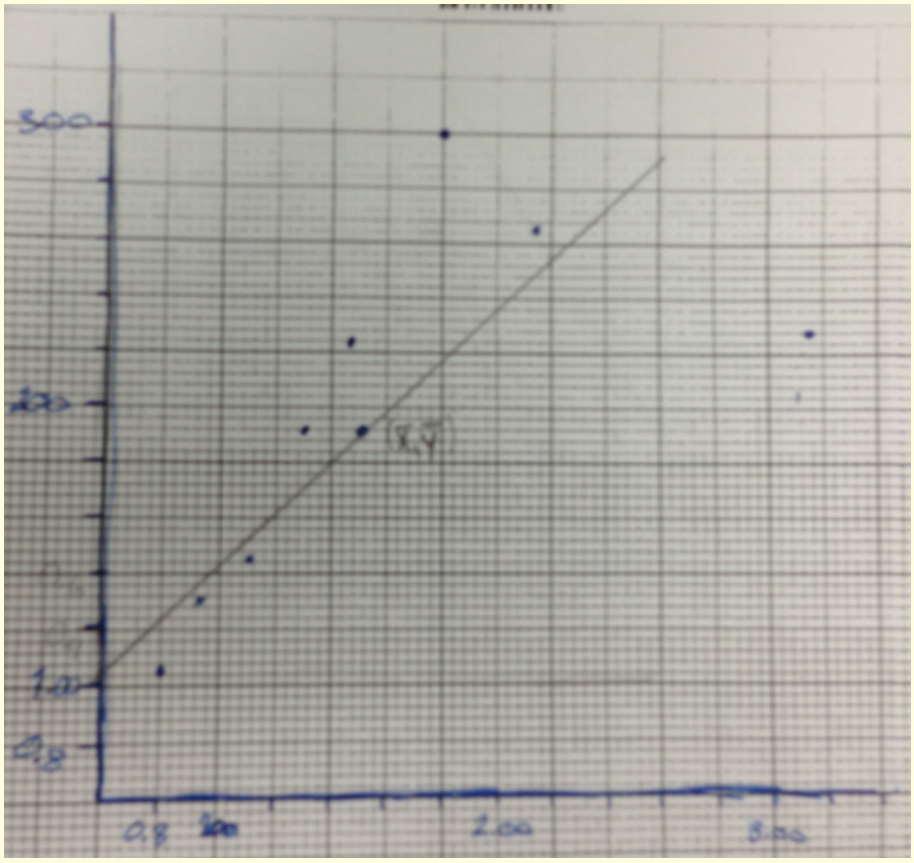
	sugar	milk	eggs	rolls	tea bags	coffee	potatoes	flour
1992 price	$ 1.44	$ 0.80	$ 2.16	$ 1.80	$ 0.92	$ 3.16	$ 1.32	$ 1.12
2002 price	$ 2.20	$ 1.04	$ 2.64	$ 3.00	$ 1.32	$ 2.28	$ 1.92	$ 1.44

a) Calculate the mean and standard deviation of the prices

i) in 1992 $\bar{x} = 1.59$, $\sigma = 0.727$

ii) in 2002 $\bar{x} = 1.98$, $\sigma = 0.635$

b) Plot the points on graph paper. Use the horizontal axis (x) for 1992 and the vertical axis (y) for 2002. Use a scale of 1 cm for $0.20.


500
1.00
0.8
2.00
3.00

c) Write down the equation of the regression line for y on x.

$y = 0.588x + 1.05$

d) Accurately draw your regression line on your graph.

e) Interpret the meaning of the
   i) gradient
   ii) y-intercept

f) What would you expect to pay in 2002 for an item costing \$2.60 in 1992?

\$2.58

g) Write down Pearson's product-moment correlation coefficient.

0.672

h) Which item would you omit to increase the correlation coefficient? Why?

Coffee

IB Practice A

A social science teacher has collected data on the number of days x per year a particular student plays sports and the number of hours y of homework that the same student does per week. She came up with the equation of the regression line $y = 40 - 0.3x$.

Interpret the gradient and the y-intercept.

↓ ↳ 0 days, # hrs of HW

HW hrs. decrease by 0.3/days

IB Practice B

A doctor is researching the relationship between the number of packs of cigarettes x a person smokes per day and the number of days per year y the person is sick. The doctor comes up with the equation of the regression line $y = 7 + 2.4x$.

Interpret the gradient and the y-intercept.

IB Practice C

The table below shows the scores for 12 golfers for their first two rounds in a local golf tournament.

Round 1 (x)	71	79	66	73	69	76	68	75	82	67	69	74
Round 2 (y)	73	81	68	75	70	79	69	77	83	68	72	76

a) i) Write down the mean score in Round 1. 72.4 [1]

ii) Write down the standard deviation in Round 1. 4.77 [1]

iii) Find the number of these golfers that had a score of more than one standard deviation above the mean in Round 1. [3]

2

b) Write down the correlation coefficient, r. [2]

$r = 0.990$

c) Write down the equation of the regression line for y on x. [2]

$y = 1.01x + 0.816$

d) Another golfer scored 70 in Round 1. Calculate an estimate of his score in Round 2. [2]

71.5

e) Interpret the meaning of the gradient and y-intercept in the context of the data. [2]

y-int: Score Rd.2, w/ 0 in Rd.1

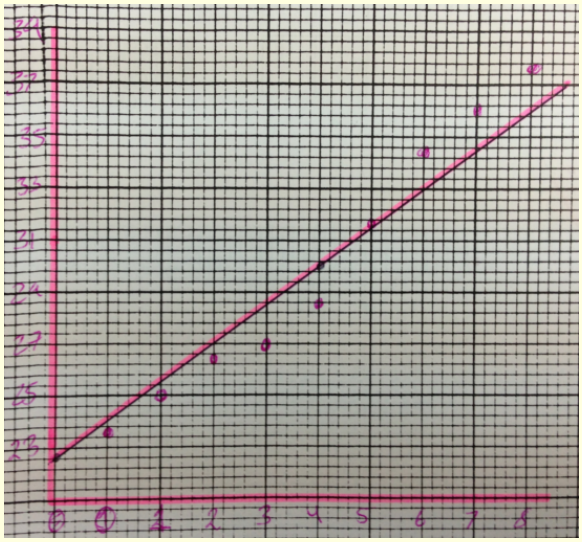
gradient: Score increases by Rd.1 → Rd.2

IB Practice D

In an experiment, a vertical spring was fixed at its upper end. It was stretched by hanging different weights on its lower end. The length of the string was then measured. The following readings were obtained.

Load (kg) x	0	1	2	3	4	5	6	7	8
Length (cm) y	23.5	25	26.5	27	28.5	31.5	34.5	36	37.5

a) Plot the data on a scatter diagram. Use a scale of 1 cm to represent 1 kg on the horizonal axis and 1 cm to represent 2 cm on the vertical axis.



b) Find the mean and standard deviation:

i) of the load $\bar{x}=4, \sigma=2.58$

ii) of the length $\bar{y}=30, \sigma=4.78$

c) Write down the equation of the regression line for y on x.

$y=1.825x+22.7$

d) Draw the regresson line for y on x on the scatter diagram.

e) Estimate the length of the spring when a load of 5.4 kg is applied.

$y=32.6$

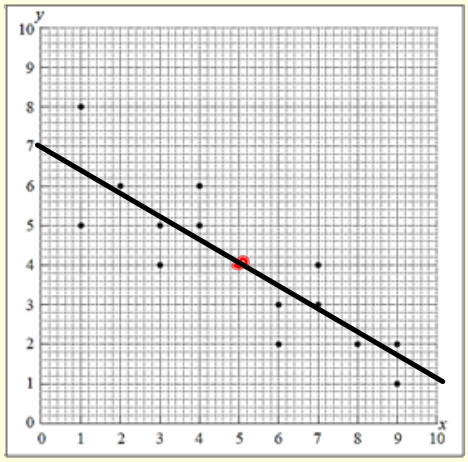
f) Malcolm uses the equation to claim that a weight of 30 kg would result in a length of 62.8 cm. Comment on his claim.

No, because extrapolation

IB Practice E

Consider the following values of x and y and the scatter diagram which represents the information given in the table.

x	1	1	2	3	3	4	4	b	6	6	7	7	8	9	9
y	5	a	6	4	5	5	6	4	2	3	3	4	2	1	2



a) Write down the value of a and of b. [2]

$a=8 \quad b=5$

b) Find the value of the correlation coefficient. [2]

-0.868

c) Write down the equation of the regression line for y on x. [2]

$y=-0.08x+7.04$

d) Draw the regresson line for y on x on the scatter diagram. [2]

$(\bar{x},\bar{y})=(5,4)$