Linear Regression – 2021/20 GCE AS Statistics Further Mathematics A

1. Nov/2021/Paper_Y532/01/No.3

(a) Using the scatter diagram in the Printed Answer Booklet, explain what is meant by **least squares** in the context of a regression line of y on x. **[2]**

(b) A set of bivariate data (t, u) is summarised as follows.

$n = 5$ $\quad \Sigma t = 35$ $\quad \Sigma u = 54$

$\Sigma t^2 = 285$ $\quad \Sigma u^2 = 758$ $\quad \Sigma tu = 460$

(i) Calculate the equation of the regression line of u on t. **[3]**

(ii) The variables t and u are now scaled using the following scaling.

$v = 2t, w = u + 4$

Find the equation of the regression line of w on v, giving your equation in the form $w = f(v)$. **[2]**

2. Nov/2020/Paper_Y532/01/No.3

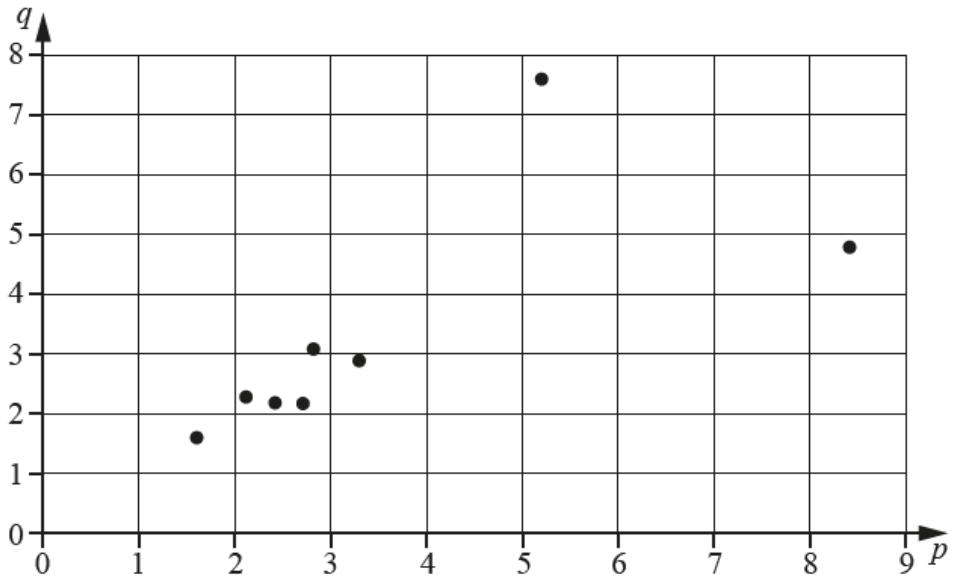
An investor obtains data about the profits of 8 randomly chosen investment accounts over two one-year periods.

The profit in the first year for each account is p% and the profit in the second year for each account is q%.

The results are shown in the table and in the scatter diagram.

Account	A	B	C	D	E	F	G	H
p	1.6	2.1	2.4	2.7	2.8	3.3	5.2	8.4
q	1.6	2.3	2.2	2.2	3.1	2.9	7.6	4.8

$n = 8$ $\quad \Sigma p = 28.5$ $\quad \Sigma q = 26.7$ $\quad \Sigma p^2 = 136.35$ $\quad \Sigma q^2 = 116.35$ $\quad \Sigma pq = 116.70$



(a) State which, if either, of the variables p and q is independent. **[1]**

(b) Calculate the equation of the regression line of q on p. **[3]**

(c) **(i)** Use the regression line to estimate the value of q for an investment account for which $p = 2.5$. **[1]**

(ii) Give two reasons why this estimate could be considered reliable. **[2]**

(d) Comment on the reliability of using the regression line to predict the value of q when $p = 7.0$. **[2]**