Name:

Date:

Regression, Correlation and Hypothesis Testing

AS-Level

Edexcel

Mathematics

Mark	Score (%)
43	

Materials

For this paper you must have:

- Ruler
- Pencil, Rubber, Protractor and Compass
- Scientific calculator, which you are expected to use when appropriate

Instructions

- Answer all questions
- Answer questions in the space provided
- All working must be shown
- Do all rough work in this book. Cross out any rough work you don't want to be marked

Information

- The marks for the questions are shown in brackets

Leave blank

1 The table shows, for each of a random sample of 8 paperback fiction books, the number of pages, x, and the recommended retail price, £y, to the nearest 10p.

x	223	276	374	433	564	612	704	766
y	6.50	4.00	5.50	8.00	4.50	5.00	8.00	5.50

(a) Calculate the value of the product moment correlation coefficient between x and y. **(3)**

(b) Interpret your value in the context of this question. **(2)**

(c) Suggest one other variable, in addition to the number of pages, which may affect the recommended retail price of a paperback fiction book. **(1)**

(d) The same 8 books were later included in a book sale. The value of the product moment correlation coefficient between the number of pages and the sale price was 0.959, correct to three decimal places.
What can be concluded from this value? **(2)**

(Total for question 1 is 8 marks)

2 The table shows the length, in centimeter, and maximum diameter, in centimeters, of each of 10 honeydew melons selected at random from those on display at a market stall.

Length	24	25	19	28	27	21	35	23	32	26
Maximum diameter	18	14	16	11	13	14	12	16	15	14

(a) Calculate the value of the product moment correlation coefficient. **(3)**

(b) Interpret your value in the context of this question. **(2)**

(Total for question 2 is 5 marks)

Leave blank

3 The scatter diagrams below were drawn by a student.

Diagram *A*

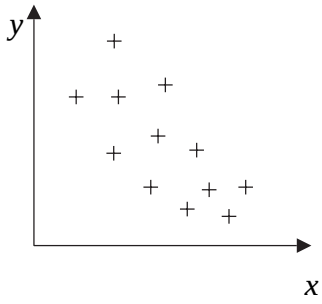


Diagram *B*

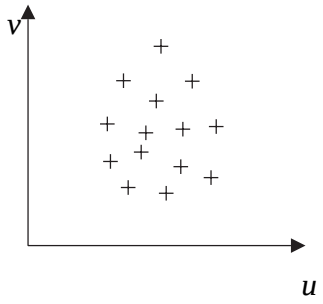
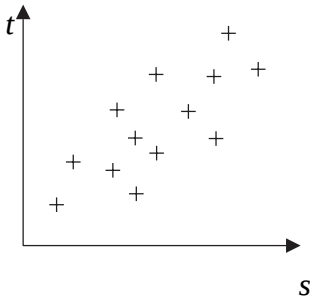


Diagram *C*



The student calculated the value of the product moment correlation coefficient for each of the sets of data.

The values were

$$0.68 \quad -0.79 \quad 0.08$$

Write down, with a reason, which value corresponds to which scatter diagram. **(6)**

(Total for question 3 is 6 marks)

Leave blank

4 Francine lives in Perth. She is investigating the relationship between the daily total rainfall, x, and the daily mean wind speed, y. She takes a random sample of eigth days from weather data collected in 2015, as shown in Table 2.

Daily total rainfall, x mm	6.6	35.0	9.8	2.0	12.2	18.6	13.0	15.8
Daily mean wind speed, y knots	6.4	14.1	9.4	5.2	9.3	10.1	8.6	8.8

(a) State the type of data recorded for daily total rainfall. **(1)**

(b) Use your calculator to find the value of the product moment correlation coefficient for these data. **(1)**

Francine claims that the value of the product moment correlation coefficient shows that there is a positive correlation between the daily total rainfall and the daily mean wind speed.

(c) Test Francine's claim, using a 1% level of significance. State your hypotheses clearly **(3)**

Francine calculates the equation of the regression line of y-on-x to be

$$y = 5.3872 + 0.255x$$

(d) State, with a reason, whether Francine is justified in using a regression line for these data. **(1)**

(e) Use Francine's regression line to predict the daily mean wind speed when there is 20mm of rainfall. **(1)**

(f) Explain why Francine should not use her regression line to predict the daily mean wind speed when there is 50mm of rainfall. **(1)**

(Total for question 4 is 8 marks)

Leave blank

5 An ornithologist believes that there is a relationship between the tail length, t mm, and the wing length, w mm, of female hook-billed kites. A random sample of size 10 is taken from a database of these kites and the relevant data is given in the table below.

t (mm)	191	197	208	180	188	210	196	191	179	208
w (mm)	284	285	288	273	280	283	288	271	257	289

The ornithologist plans to use a linear regression model based on these data and interpolate or extrapolate as necessary to estimate the wing length of other female hook-billed kites from their tail length.

(a) (i) Explain what is meant by extrapolation. **(1)**

(ii) Explain the dangers of extrapolation. **(1)**

The ornithologist attempts to calculate the product moment correlation coefficient, r, and obtains a value of 1.3

(b) Explain how she should be able to identify that this is incorrect without carrying out any further calculations. **(1)**

(c) Use your calculator to find the correct value of the product moment correlation coefficient, r. **(1)**

(d) Stating your hypotheses clearly test, at the 1% significance level, whether or not there is evidence that the product moment correlation coefficient for the population is positive. **(3)**

(e) Explain what your test in part (d) suggests about female hook-billed kites. **(1)**

(Total for question 5 is 8 marks)

Leave blank

6 The product moment correlation coefficient is denoted by r

(a) Sketch separate scatter diagrams, with five points on each diagram, to show $r = 1$ **(3)**

Two judges rank seven collie dogs in a competition. The collie dogs are labelled A to G and the rankings are as follows

Rank	1	2	3	4	5	6	7
Judge 1	*A*	*C*	*D*	*B*	*E*	*F*	*G*
Judge 2	*A*	*B*	*D*	*C*	*E*	*G*	*F*

(b) Stating your hypotheses clearly, test, at the 5% level of significance, whether or not the judges are generally in agreement. **(5)**

(Total for question 6 is 8 marks)