Write your name here

Surname	Other names

Pearson Edexcel GCE

Centre Number | Candidate Number

AS and A level Further Mathematics
Further Statistics 2

Practice Paper

Correlation and regression

You must have:
Mathematical Formulae and Statistical Tables (Pink)

Total Marks

Instructions

- Use black ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B).
- Fill in the boxes at the top of this page with your name, centre number and candidate number.
- Answer all the questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions in the spaces provided – there may be more space than you need.
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Inexact answers should be given to three significant figures unless otherwise stated.

Information

- A booklet 'Mathematical Formulae and Statistical Tables' is provided.
- There are 6 questions in this question paper. The total mark for this paper is 82.
- The marks for each question are shown in brackets – use this as a guide as to how much time to spend on each question.
- Calculators must not be used for questions marked with a * sign.

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.
- If you change your mind about an answer, cross it out and put your new answer and any working underneath.

1. A clothes shop manager records the weekly sales figures, £ s, and the average weekly temperature, t °C, for 6 weeks during the summer. The sales figures were coded so that

$$w = \frac{s}{1000}$$

The data are summarised as follows

$$S_{ww} = 50 \qquad \sum wt = 784 \qquad \sum t^2 = 2435 \qquad \sum t = 119 \qquad \sum w = 42$$

(*a*) Find S_{wt} and S_{tt}

(3)

(*b*) Write down the value of S_{ss} and the value of S_{st}

(2)

(*c*) Find the product moment correlation coefficient between s and t.

(2)

The manager of the clothes shop believes that a linear regression model may be appropriate to describe these data.

(*d*) State, giving a reason, whether or not your value of the correlation coefficient supports the manager's belief.

(1)

(*e*) Find the equation of the regression line of w on t, giving your answer in the form $w = a + bt$

(3)

(*f*) Hence find the equation of the regression line of s on t, giving your answer in the form $s = c + dt$, where c and d are correct to 3 significant figures.

(2)

(*g*) Using your equation in part (*f*), interpret the effect of a 1°C increase in average weekly temperature on weekly sales during the summer.

(1)

(Total 14 marks)

---

2. A meteorologist believes that there is a relationship between the height above sea level, h m, and the air temperature, t °C. Data is collected at the same time from 9 different places on the same mountain. The data is summarised in the table below.

h	1400	1100	260	840	900	550	1230	100	770
t	3	10	20	9	10	13	5	24	16

[You may assume that $\sum h = 7150$, $\sum t = 110$, $\sum h^2 = 7171500$, $\sum t^2 = 1716$, $\sum th = 64\,980$ and $S_{tt} = 371.56$]

(*a*) Calculate S_{th} and S_{hh}. Give your answers to 3 significant figures.

(3)

(*b*) Calculate the product moment correlation coefficient for this data.

(2)

(*c*) State whether or not your value supports the use of a regression equation to predict the air temperature at different heights on this mountain. Give a reason for your answer.

(1)

(*d*) Find the equation of the regression line of t on h giving your answer in the form $t = a + bh$.

(4)

(*e*) Interpret the value of b.

(1)

(*f*) Estimate the difference in air temperature between a height of 500 m and a height of 1000 m.

(2)

(Total 13 marks)

3. A large company is analysing how much money it spends on paper in its offices every year. The number of employees, x, and the amount of money spent on paper, p (£ hundreds), in 8 randomly selected offices are given in the table below.

x	8	9	12	14	7	3	16	19
p (£ hundreds)	40.5	36.1	30.4	39.4	32.6	31.1	43.4	45.7

(You may use $\Sigma x^2 = 1160$ $\Sigma p = 299.2$ $\Sigma p^2 = 11\ 422$ $\Sigma xp = 3449.5$)

(*a*) Show that $S_{pp} = 231.92$ and find the value of S_{xx} and the value of S_{xp}.

(5)

(*b*) Calculate the product moment correlation coefficient between x and p.

(2)

The equation of the regression line of p on x is given in the form $p = a + bx$.

(*c*) Show that, to 3 significant figures, $b = 0.824$ and find the value of a.

(4)

(*d*) Estimate the amount of money spent on paper in an office with 10 employees.

(2)

(*e*) Explain the effect each additional employee has on the amount of money spent on paper.

(1)

Later the company realised it had made a mistake in adding up its costs, p. The true costs were actually half of the values recorded. The product moment correlation coefficient and the equation of the linear regression line are recalculated using this information.

(*f*) Write down the new value of

(i) the product moment correlation coefficient,

(ii) the gradient of the regression line.

(2)

(Total 16 marks)

4. The table shows data on the number of visitors to the UK in a month, v (1000s), and the amount of money they spent, m (£ millions), for each of 8 months.

Number of visitors v (1000s)	2450	2480	2540	2420	2350	2290	2400	2460
Amount of money spent m (£ millions)	1370	1350	1400	1330	1270	1210	1330	1350

You may use

$S_{vv} = 42587.5$ $S_{vm} = 31512.5$ $S_{mm} = 25187.5$ $\Sigma v = 19390$ $\Sigma m = 10610$

(*a*) Find the product moment correlation coefficient between m and v.

(2)

(*b*) Give a reason to support fitting a regression model of the form $m = a + bv$ to these data.

(1)

(*c*) Find the value of b correct to 3 decimal places.

(2)

(*d*) Find the equation of the regression line of m on v.

(2)

(*e*) Interpret your value of b.

(2)

(*f*) Use your answer to part (*d*) to estimate the amount of money spent when the number of visitors to the UK in a month is 2 500 000.

(2)

(*g*) Comment on the reliability of your estimate in part (*f*). Give a reason for your answer.

(2)

(Total 13 marks)

5. A biologist is studying the behaviour of bees in a hive. Once a bee has located a source of food, it returns to the hive and performs a dance to indicate to the other bees how far away the source of the food is. The dance consists of a series of wiggles. The biologist records the distance, d metres, of the food source from the hive and the average number of wiggles, w, in the dance.

Distance, *d* m	30	50	80	100	150	400	500	650
Average number of wiggles, *w*	0.725	1.210	1.775	2.250	3.518	6.382	8.185	9.555

[You may use $\sum w = 33.6$ $\sum dw = 13833$ $S_{dd} = 394600$ $S_{ww} = 80.481$ (to 3 decimal places)]

(*a*) Show that $S_{dw} = 5601$.

(2)

(*b*) State, giving a reason, which is the response variable.

(1)

(*c*) Calculate the product moment correlation coefficient for these data.

(2)

(*d*) Calculate the equation of the regression line of w on d, giving your answer in the form $w = a + bd$.

(4)

A new source of food is located 350 m from the hive.

(*e*) (i) Use your regression equation to estimate the average number of wiggles in the corresponding dance.

(ii) Comment, giving a reason, on the reliability of your estimate.

(2)

(Total 11 marks)

6. The age, t years, and weight, w grams, of each of 10 coins were recorded. These data are summarised below.

$$\sum t^2 = 2688 \qquad \sum tw = 1760.62 \qquad \sum t = 158 \qquad \sum w = 111.75 \qquad S_{ww} = 0.16$$

(*a*) Find S_{tt} and S_{tw} for these data.

(3)

(*b*) Calculate, to 3 significant figures, the product moment correlation coefficient between t and w.

(2)

(*c*) Find the equation of the regression line of w on t in the form $w = a + bt$.

(4)

(*d*) State, with a reason, which variable is the explanatory variable.

(2)

(*e*) Using this model, estimate

(i) the weight of a coin which is 5 years old,

(ii) the effect of an increase of 4 years in age on the weight of a coin.

(2)

It was discovered that a coin in the original sample, which was 5 years old and weighed 20 grams, was a fake.

(*f*) State, without any further calculations, whether the exclusion of this coin would increase or decrease the value of the product moment correlation coefficient. Give a reason for your answer.

(2)

(Total 15 marks)

TOTAL FOR PAPER: 82 MARKS