Bivariate Data: exam practice

10 questions · 43 marks · worked solutions and markschemes

Name _____ Date _____ Score / 43

- Answer every question before you look at the solutions. They start after the last question, with the marks for each line: M for method, A for accuracy, R for reasoning.
- Online, with more practice: ibmathrevision.com/slai/bivariate-data-exam-practice

1. **3 marks · Paper 1 style · calculator allowed**

In a photography contest, Judge A and Judge B rank 5 participants. The sum of the squared differences between their ranks, $\sum d^2$, is found to be 12. Calculate the exact value of Spearman's rank correlation coefficient, r_s, for this data.

2. **3 marks · Paper 1 style · calculator allowed**

Two judges rank six dancers.

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

(a) Find Spearman's rank correlation coefficient, r_s. **[2]**

(b) Comment on the level of agreement between the judges. **[1]**

3. **3 marks · Paper 1 style · calculator allowed**

For data with x between 2 and 20, the regression line of y on x is $y = 2.4x + 15$ and $r = 0.93$.

(a) Estimate y when $x = 10$. **[2]**

(b) Explain why the line should not be used to estimate y when $x = 50$. **[1]**

4. **3 marks · Paper 1 style · calculator allowed**

Five pairs of values are recorded.

x	3	7	2	9	5
y	5	9	4	8	6

(a) Rank each set of values. **[2]**

(b) Find Spearman's rank correlation coefficient. **[1]**

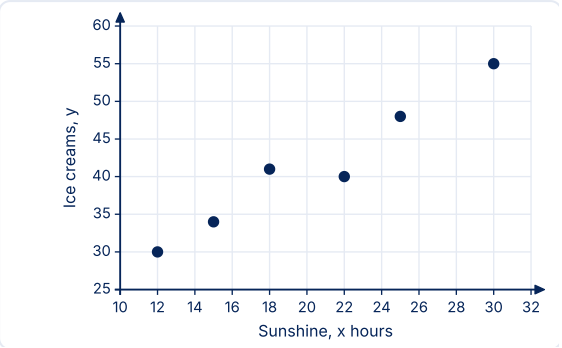
5. **4 marks · Paper 2 style · calculator allowed**

The table shows the hours of sunshine, x, and the number of ice creams sold, y, at a kiosk on six days.

Sunshine, x	12	15	18	22	25	30
Ice creams, y	30	34	41	40	48	55

(a) Find Spearman's rank correlation coefficient. **[3]**

(b) Find Pearson's product-moment correlation coefficient. **[1]**



Sunshine and ice creams sold

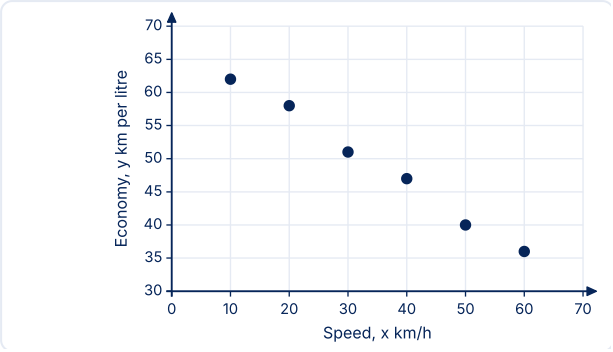
6. **5 marks · Paper 2 style · calculator allowed**

The table shows the speed of a car, x km h^{-1}, and its fuel economy, y km per litre, on six test runs.

Speed, x	10	20	30	40	50	60
Economy, y	62	58	51	47	40	36

(a) Find the value of r and describe the correlation. **[2]**

(b) Find the regression line of y on x and estimate y when $x = 45$. **[3]**



Speed and fuel economy

7. **6 marks · Paper 2 style · calculator allowed**

An environmentalist models the relationship between the distance from a highway (d, in metres) and the concentration of a pollutant (C, in parts per million). The equation of the regression line is $C = -0.04d + 8.5$.

1. Estimate the concentration of the pollutant at a distance of 50 metres from the highway.
2. The actual concentration measured at 50 metres was 7.1 parts per million. Calculate the percentage error of the estimate.

8. **7 marks · Paper 2 style · calculator allowed**

The table shows the age, x years, and the price, y thousand euros, of eight cars of the same model.

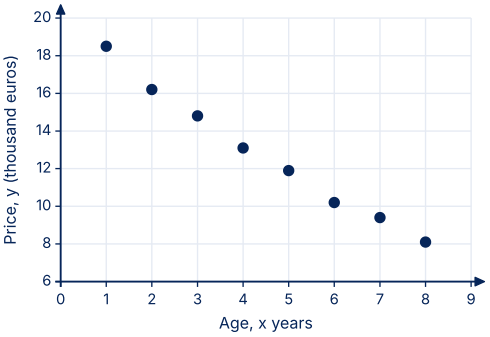
Age, x	1	2	3	4	5	6	7	8
Price, y	18.5	16.2	14.8	13.1	11.9	10.2	9.4	8.1

(a) Find Pearson's product-moment correlation coefficient, r. **[2]**

(b) Find the equation of the regression line of y on x. **[2]**

(c) Interpret the gradient of the line in context. **[1]**

(d) Estimate the price of a car that is 5.5 years old. **[2]**



Age and price of eight cars

9. **4 marks · Paper 2 style · calculator allowed**

A data set is shown.

x	1	2	3	4	5	6
y	2	4	8	16	32	64

(a) Write down the value of Spearman's rank correlation coefficient. **[1]**

(b) Find Pearson's product-moment correlation coefficient. **[1]**

(c) Explain why the two values are different. **[2]**

10. **5 marks · Paper 1 style · calculator allowed**

A linear regression line $y = mx + c$ is fit to a set of data points. One of the data points is $(5, 12)$. The equation of the regression line is found to be $y = 1.8x + 2.5$.

1. Calculate the residual for the point $(5, 12)$.
2. The mean of the x values is $\bar{x} = 4.2$. Calculate the mean of the y values, $\bar{y}$.

Answers and worked solutions

Question 1 — 3 marks

Answer: 0.4

Spearman formula with $n = 5$:

$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$

Substitute:

$r_s = 1 - \frac{6(12)}{5(25-1)} = 1 - \frac{72}{120}$ **M1**

$\frac{72}{120} = 0.6$ **A1**

$r_s = 1 - 0.6 = 0.4$ **A1**

Question 2 — 3 marks

Answer: (a) $r_s = 0.829$

(a)

Pearson's r of the ranks, GDC **(M1)**

$r_s = 0.829$ **A1**

(b)

strong agreement (strong positive correlation of ranks) **A1**

Question 3 — 3 marks

Answer: (a) 39

(a)

$2.4 \times 10 + 15$ **(M1)**

39 **A1**

(b)

50 is outside the range of the data (extrapolation) **R1**

Question 4 — 3 marks

Answer: (b) $r_s = 0.9$

(a)

x: $2, 4, 1, 5, 3$ **A1**

y: $2, 5, 1, 4, 3$ **A1**

(b)

$r_s = 0.9$ **A1**

Question 5 — 4 marks

Answer: (a) $r_s = 0.943$; (b) $r = 0.974$

(a)

ranks of x: $1, 2, 3, 4, 5, 6$; ranks of y: $1, 2, 4, 3, 5, 6$ **(M1)(A1)**

$r_s = 0.943$ **A1**

(b)

$r = 0.974$ **A1**

Question 6 — 5 marks

Answer: (a) $r = -0.997$, strong negative; (b) 43.6 km per litre (line $y = -0.537x + 67.8$)

(a)

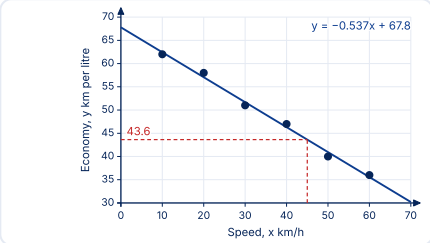
$r = -0.997$ **A1**

strong negative **A1**

(b)

$y = -0.537x + 67.8$ **A1A1**

43.6 **A1**



Estimate at $45\ \text{km/h}$: 43.6 km per litre

Question 7 — 6 marks

Answer: (a) 6.5; (b) 8.45

1. Substitute $d = 50$ into the regression line. **(M1)**
   $C = -0.04(50) + 8.5$.
   $C = -2.0 + 8.5 = 6.5$ ppm
   . **A1**
2. Use the percentage error formula:
   $\left|\frac{\text{Approximate}-\text{Exact}}{\text{Exact}}\right| \times 100$
   . **(M1)**
   $\text{Error} = \left|\frac{6.5-7.1}{7.1}\right| \times 100$.
   (A1)
   $\text{Error} = \left|\frac{-0.6}{7.1}\right| \times 100 \approx 8.45\%$
   . **A1A1**

Question 8 — 7 marks

Answer: (a) $r = -0.994$; (b) $y = -1.45x + 19.3$; (d) 11.3 thousand euros

(a)

$r = -0.994$ **A2**

(b)

$y = -1.45x + 19.3$ **A1A1**

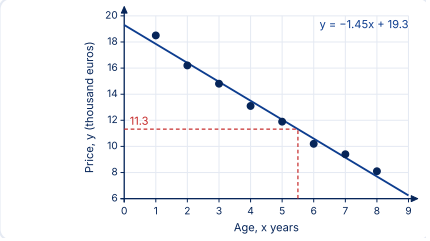
(c)

the price falls by about €1450 for each extra year of age **A1**

(d)

$-1.45 \times 5.5 + 19.3$ **(M1)**

11.3 thousand euros **A1**



Estimate at 5.5 years: 11.3 thousand euros

Question 9 — 4 marks

Answer: (a) $r_s = 1$; (b) $r = 0.906$

(a)

$r_s = 1$ **A1**

(b)

$r = 0.906$ **A1**

(c)

y always increases as x increases, so the ranks agree exactly **R1**

but the relationship is not linear, so $r < 1$ **R1**

Question 10 — 5 marks

Answer: (a) 0.5; (b) 10.06

1. Predicted y for $x = 5$:
   $y = 1.8(5) + 2.5 = 9 + 2.5 = 11.5$
   . $M1$**A1**
   Residual $=$ Actual $-$ Predicted $=$
   $12 - 11.5 = 0.5$. **A1**
2. The regression line must pass through the mean point $(\bar{x}, \bar{y})$. **M1**
   Substitute $\bar{x} = 4.2$:
   $\bar{y} = 1.8(4.2) + 2.5 = 7.56 + 2.5 = 10.06$
   . **A1**

Every solution is also on ibmathrevision.com/slai/bivariate-data-exam-practice
Our own original questions, written and checked by IB Math Revision. IB Math Revision is independent and is not affiliated with, endorsed by or connected to the International Baccalaureate Organization. This work/product/service has been developed independently from and is not endorsed by the International Baccalaureate Organization. International Baccalaureate, Baccalauréat International, Bachillerato Internacional and IB are registered trademarks owned by the International Baccalaureate Organization.