Statistics and regression - Practice test 1

Name: _____ Date: _____

10 questions | 55 marks | 70 minutes

Read paired data, fit and assess a line, then use normal probabilities and cutoffs.

Use a graphics calculator or appropriate digital tool for labelled tasks.

1. A class records weekly revision time and a quiz score for six volunteers. [5 marks]

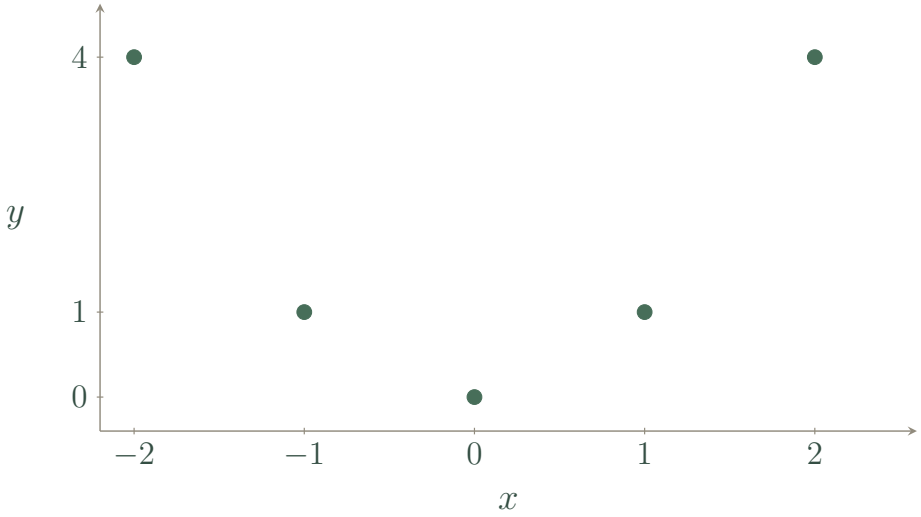
Paired measurements

Time (hours/week)	Score (%)
1	42
2	47
3	51
4	50
5	61
6	65

(a) For predicting score from revision time, identify the explanatory and response variables. [2 marks]

(b) Plot the paired observations on labelled axes and describe the broad direction. [3 marks]

2. Consider points $(-2, 4), (-1, 1), (0, 0), (1, 1), (2, 4)$. [4 marks]



(a) Describe their form. [2 marks]

(b) Their correlation is zero. Explain why that does not mean the variables are unrelated. [2 marks]

3. An observational survey reports that students who sleep longer tend to report better concentration. [5 marks]

(a) Give a cautious conclusion. [2 marks]

(b) Identify a possible confounder and describe a design improvement. [3 marks]

4. For a paired sample, $S_{xx} = 40, S_{yy} = 90$ and $S_{xy} = 48$. [5 marks]

(a) Calculate r. [2 marks]

(b) Find and interpret r^2. [3 marks]

5. A sample has $\bar{x} = 8, \bar{y} = 30, s_x = 2, s_y = 5$ and $r = 0.8$. [5 marks]

(a) Find the least-squares line predicting y from x. [3 marks]

(b) Predict y at $x = 10$. [2 marks]

6. A courier fits $C = 7 + 2.4m$, where cost C is in dollars and parcel mass m is in kg. Data covered 1-12 kg. [6 marks]

(a) Interpret the slope. [2 marks]

(b) Predict cost at 8 kg and classify the prediction. [2 marks]

(c) Explain why the intercept is a model coefficient, not proof that a zero-mass parcel can be sent. [2 marks]

7. The model $\hat{y} = 5 + 3x$ is used with observations $(1, 9), (2, 10), (3, 16)$. [7 marks]

(a) Calculate the three residuals. [3 marks]

(b) Calculate the sum of squared residuals. [2 marks]

(c) Why need these residuals not sum to zero? [2 marks]

8. One observation is unusually far above a scatter pattern. A report deletes it and announces a higher r^2. [5 marks]

(a) Explain why the deletion is not justified by the higher r^2 alone. [2 marks]

(b) Describe a transparent analysis. [3 marks]

9. Packet mass X is normal with mean 250 g and standard deviation 3 g. [8 marks]

(a) Find $P(X < 247)$. [3 marks]

(b) Find $P(247 < X < 256)$. [3 marks]

(c) For 400 packets, find the expected number below 247 g. [2 marks]

10. Scores follow a normal model with mean 60 and SD 9. [5 marks]

(a) Find the lower cutoff for the highest 15%. [3 marks]

(b) State one reason this cutoff should not be used for a markedly skewed score distribution. [2 marks]