

EXERCISE ON REGRESSION

Philip Pennance

February 2010

1. Find the coefficient of linear correlation between the concentration of ozone Y in ppm and the number of cars X in thousands. If:

Y	2	5	7	8	11
X	10	12	15	20	21

2. The table below shows lobster yield Y versus water temperature X .

Yield Y :	2	4	6	8	10
Temperature X :	4	5	6	7	8

Find

- (a) $S_{XY} = \sum (X - \bar{X})(Y - \bar{Y})$
 - (b) $S_{XX} = \sum (X - \bar{X})^2$
 - (c) The slope of the regression line.
 - (d) The intercept of the regression line.
 - (e) The predicted yield when $X = 10$.
3. Let X and Y be random variables and ρ the coefficient of linear correlation between X and Y . Show that $-1 \leq \rho \leq 1$.

4. Write a program in SAS which generates 1000 values of a standard normal random variable X and calculates the coefficient of linear correlation between X and $Y = X^2$.

- (a) Is there evidence of linear correlation at the .05 level of significance?

- (b) Are X and Y correlated?

Random numbers can be generated using the SAS function RANNORM (seed). Use your student ID as "seed".

5. Consider the model $Y_i = \beta X_i + \epsilon_i$

- (a) What is the least squares estimate of β_1 ?

- (b) Give an example in which the model might apply.

6. In problem 5. find

- (a) Eb_1

- (b) $\text{Var}(b_1)$

- (c) EY_h given $X = x_h$

- (d) $\text{Var}(Y_h)$ given $X = x_h$

- (e) The point of smallest variability for $\hat{Y}_h = b_1 x_h$