MATU9D2 : PRACTICAL STATISTICS

Spring 2017

PRACTICAL SESSION 8

- Hand Calculations
  : Testing Correlation
  : Simple Linear Regression

- Handout 1 of 2

C.A. Howie
Computing Science & Maths

University of Stirling

ANSWER THE FOLLOWING QUESTIONS USING PEN, PAPER AND CALCULATOR - NOT COMPUTER

1. Take the data given below and construct a scatter diagram. **Note that you plotted this data and calculated summary statistics and r for this data in Practical 7.**

x	10	12	14	16	18	20	22	24	26	28
y	25	24	22	20	19	17	13	12	11	10

(a) Is the correlation significantly different to zero?

(b) Find the regression line of y on x for the data and estimate the error variance.

(c) Add the fitted line (regression line) to your graph.

(d) Is the slope equal to zero?

2. A farmer wishes to predict the number of tons per acre of crop which will result from a given number of applications of fertiliser. Data has been collected and is shown below:

Fertiliser applications	1	2	4	5	6	8	10
Tons per acre	2	3	4	7	12	10	7

(a) Plot the data

(b) Find a suitable regression relationship to help the farmer in making the prediction.

(c) Predict the number of tons per acre will result from 7 fertiliser applications.

(d) Calculate a 95% Confidence Interval for the mean yield given 7 fertiliser applications

(e) Calculate a 95% Prediction Interval for the yield given an individual farmer applies fertiliser 7 times.

FORMULAE

The test of the Null hypothesis $H_o : \rho = 0$ against the alternative $H_1 : \rho \neq 0$ has test Statistic

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Using significance level of 0.05, this is compared with $\pm$t(n-2; 0.025).

FORMULAE (continued)

$$S_{XX} = \sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}$$

$$S_{XY} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$$

$$S_{YY} = \sum y_i^2 - \frac{\left(\sum y_i\right)^2}{n}$$

Inference in Simple Linear Regression

The residual variance σ^2 is estimated by

$$\hat{\sigma}^2 = \frac{S_{YY} - \dfrac{S_{XY}^2}{S_{XX}}}{n-2}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

$$\hat{\beta} = \frac{S_{XY}}{S_{XX}}$$

A test of the Null Hypothesis $H_0 : \beta = 0$ against the alternative $H_1 : \beta \neq 0$ has the test statistic

$$T = \frac{\hat{\beta}}{\sqrt{\dfrac{\hat{\sigma}^2}{S_{XX}}}}$$

Using a significance level of 0.05, this is compared with $\pm$ t(n -2; 0.025).

Parameter	Estimate	Estimated Standard Error
α	$\hat{\alpha}$	$\sqrt{\hat{\sigma}^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{XX}}\right)}$
β	$\hat{\beta}$	$\sqrt{\frac{\hat{\sigma}^2}{S_{XX}}}$
$\alpha + \beta x$	$\hat{\alpha} + \hat{\beta}x$	$\sqrt{\hat{\sigma}^2\left(\frac{1}{n} + \frac{(x-\bar{x})^2}{S_{XX}}\right)}$

In each case interval estimates, having confidence 0.95, are given by

$$Estimate \quad \pm \quad t(n-2;0.025) \times Estimated\ Standard\ Error$$

The formula for a 95% prediction interval for an individual observation is

$$\hat{\alpha} + \hat{\beta}x \quad \pm t(n-2;0.025)\sqrt{\hat{\sigma}^2\left(1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{S_{XX}}\right)}$$