CHAPTER NINE

CORRELATION AND SIMPLE LINEAR REGRESSION

9.0 Introduction

The analysis of correlation can determine the strength of the relationship between two variables. From the correlation, we can further convert the relationship in term of mathematical equation which is known as simple linear regression.

9.1 Practical Steps to Solve Simple Linear Regression Problem

$$Y = a + bX$$

Step 1: Draw the scatter plot of the (X,Y) data for visual inspection of the relationship that may exist between X and Y.

Step 2: Construct the following table to facilitate computation.

k	X	Y	X^2	Y^2	XY
1	x_1	y_1	x_1^2	y_1^2	x_1y_1
2	x_2	y_2	x_2^2	y_2^2	x_2y_2
:	:	:	:	:	:
n	x_n	y_n	x_n^2	y_n^2	x_ny_n
Sum	$\sum_i x_i$	$\sum_i y_i$	$\sum_i x_i^2$	$\sum_i y_i^2$	$\sum_i x_i y_i$

Step 3: Calculate the linear regression parameters (a, b) using.

$$SS_{xy} = \sum_i x_i y_i - \frac{1}{n}(\sum_i x_i)(\sum_i y_i) \quad \text{and} \quad SS_{XX} = \sum_i x_i^2 - \frac{1}{n}(\sum_i x_i)^2$$

where

$$b = \frac{SS_{XY}}{SS_{XX}} \qquad \text{and} \qquad a = \bar{y} - b\bar{x}$$

Step 4: The linear regression model of the data is given by

$$Y = a + bX$$

9.2 Calculate the linear regression parameters

$$SS_{xy} = \sum xy - \frac{\left(\sum x\right)\left(\sum y\right)}{n}$$

$$SS_{xx} = \sum x^2 - \frac{\left(\sum x\right)^2}{n}$$

$$SS_{yy} = \sum y^2 - \frac{\left(\sum y\right)^2}{n}$$

$$b = SS_{xy} / SS_{xx}$$

$$a = \bar{y} - b\bar{x}$$

9.3 Sample Estimate of Correlation Coefficient

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx} SS_{yy}}}$$

9.4 Making Inferences about the Slope, *B*

$$s_e = \sqrt{\frac{SS_{yy} - bSS_{xy}}{n-2}}$$

$$SSE = \sum e^2 = \sum (y - \hat{y})^2$$

$$SST = \sum y^2 - \frac{\left(\sum y\right)^2}{n}$$

$$SSR = SST - SSE$$

$$r^2 = bSS_{xy} / SS_{yy}$$

$$b \pm t_{\alpha/2} s_b$$

$$s_b = s_e / \sqrt{SS_{xx}}$$

9.5 Hypothesis Testing About *B*

$$t = \frac{b - B}{s_b}$$