

Tutorial Topic 8

Simple Linear Regression

Introduction

In this topic we will be looking at Simple Linear Regression where we attempt to quantify the relationship between variables with one of those variables being the dependant variable. Dependant variable is the variable we are investigating and an Independent variable is an explanatory variable in that it helps explain the variation in the Dependant variable. So simple regression involves a single dependent variable and a single independent variable. In the next topic when we look at multiple regression, will have multiple independent variables that will attempt to explain variations in the dependant variable.

A simple regression business example could be Sales Revenue for a product. So in my example I'd make sales revenue my dependant variable because that is the variable I'm investigating and therefore my independent variable could be advertising expenditure. I could plot the sales revenue for various levels of advertising expenditure on a scatter diagram. From that data, I can create a regression equation. There are some pitfalls here, so I do need to check some assumptions first.

Therefore, the aims of this tutorial are to:

- conduct a simple regression and interpret the meaning of the regression coefficients b_0 and b_1
- use regression analysis to predict the value of a dependent variable based on an independent variable
- assess the adequacy of your estimated model
- evaluate the assumptions of regression analysis
- make inferences about the slope
- make inferences about the correlation coefficient
- estimate mean values and predict individual values

Textbook Questions

- 12.9 A concert promoter is interested in the relationship between ticket price and merchandise sales. She obtains data from the previous year. [Dataset: CONCERT.XLSX]

Merchandise Sales (\$)	Ticket Price (\$)
3,402	115
7,069	109
20,269	79
1,113	125
15,116	85
16,954	85
18,317	80
13,623	95
24,131	65
24,073	55
21,806	75
10,687	100

- a. Construct a scatter diagram.

- b. Assuming a linear relationship, use the least-squares method to find the regression coefficients, b_0 and b_1 .
 - c. Interpret the meaning of b_0 and b_1 in this problem.
 - d. Predict merchandise sales if the ticket price is \$100.
- 12.27 In the previous problem, 12.9, a concert promoter used the ticket price to predict merchandise sales. Perform a residual analysis for these data. Based on these results:
- a. Determine the adequacy of the fit of the model.
 - b. Evaluate whether the assumptions of regression have been seriously violated.
- 12.45 A question has arisen about the best location for police breath-testing units on public roads in Victoria. One argument is that testing should be done in high-risk locations such as the sites of major events and outside night entertainment areas. Another argument is that the number of people tested and found to be above the 0.05 drink-driving limit is simply a function of the number of drivers tested, regardless of location. Data are easily tabulated for an assumed number of charges laid and the number of drivers tested over a two-week period. (Use Excel Data Analysis: Correlation.)

Number of drivers tested	Number of drivers charged
250	20
288	19
330	24
68	5
145	6
310	17
220	10
88	6
71	4
169	8
121	11
115	9
196	16
243	22

- a. Calculate the coefficient of correlation, r .
- b. At the 0.05 level of significance, is there a significant linear relationship between the number of drivers tested and the number of drivers charged for drink-driving?

TEXTBOOK REFERENCE:

Basic Business Statistics: Concepts and Applications. *Berenson, M.L. Levine, D.M. Szabat, K.A. O'Brien, M. Jayne, N. Watson, J.* 5th edition. 2019. Pearson Australia Group Pty Ltd. ISBN 9781488617249. Chapter 12, sections 12 to 12.5 and 12.7 to 12.9.