

Simple Linear Regression Analysis

PHILHOON OH

October 5, 2016

Abstract

In this report, I will reproduce the main results displayed in section 3.1 *Simple Linear Regression* (chapter3) of the book **An introduction to Statistical Learning**.

Introduction

The overall goal is to provide advice on how to improve sales of the particular product. More specifically, the idea is to determine whether there is an association between advertising and sales, and if so, develop an accurate model that can be used to predict sales on the basis of the three media budgets

Data

The Advertising data set consists of the Sales (in thousands of units) of a particular product in 200 different markets, along with advertising budgets (in thousands of dollars) for the product in each of those markets for three different media: TV, Radio, and Newspaper.

Methodology

We consider one media from the data set, TV, and study its relationship with Sales. For this purpose, we use a simple linear model:

$$Sales = \beta_0 + \beta_1 TV$$

To estimate the coefficients β_0 and β_1 we fit a regression model via the least squares criterion.

Results

We compute the regression coefficients

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.03	0.46	15.36	0.00
tv	0.05	0.00	17.67	0.00

Table 1: Information about Regression Coefficients

-Under the null hypothesis, p-value of intercept is significant to reject the null. The estimate of the intercept is 7.0325935 with sd 0.4578429

-Under the null hypothesis, p-value of slope is significant to reject the null. The estimate of the slope is 0.4578429 with 0.0026906

More information about the least squares model is given in the table below:

	Quantity	Value
1	R2	0.61
2	RSE	3.26
3	F-stat	312.14

Table 2: Regression Quality Indices

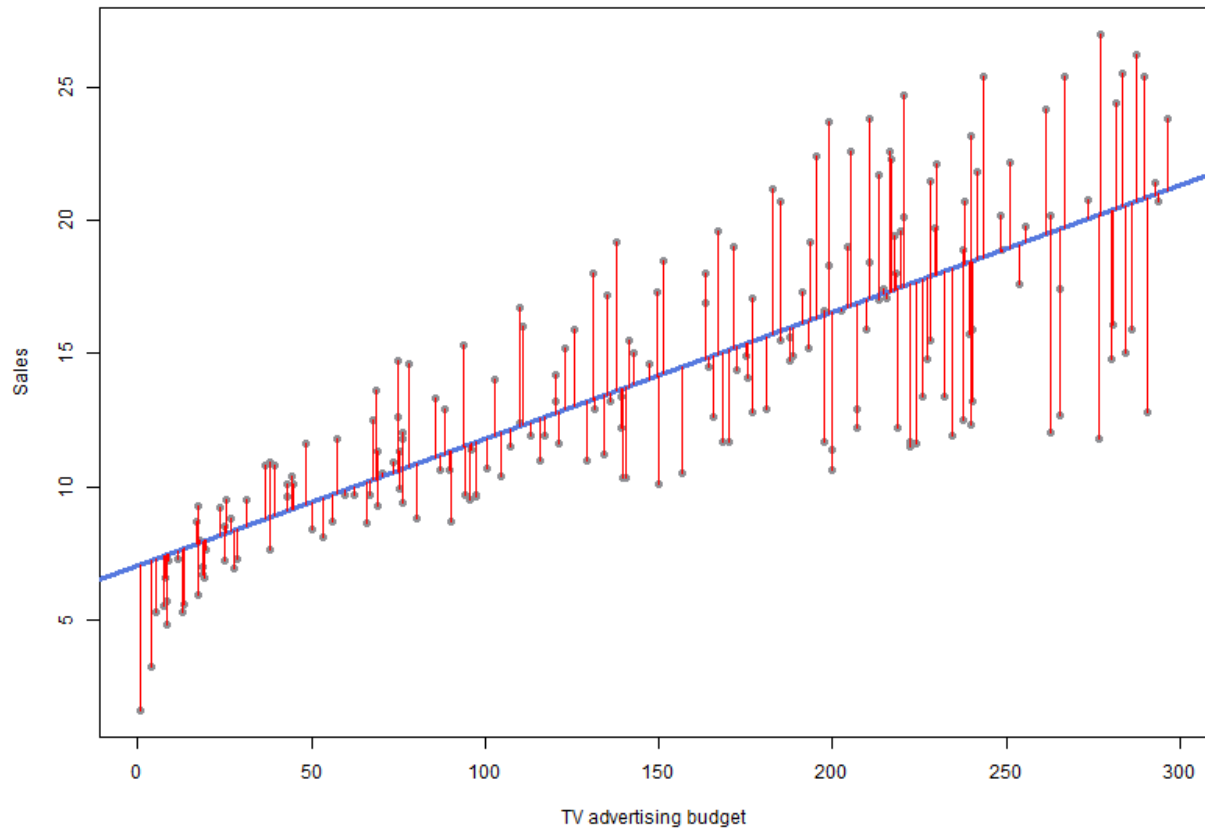


Figure 1: scatterplot-tv-sales

Conclusions