# Simple Regression Analysis

Kevin Liao 10/7/2016

## Abstract

In this report, we reproduce the main results displayed in section 3.1 Simple Linear Regression (chapter 3) 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  $\beta_0$  and  $\beta_1$  we fit a regression model via the least squares criterion.

0.04753664 0.002690607 17.66763 1.46739e-42

## Results

```
# All the numbers could be retrieved from regression.RData
load("../data/regression.RData")
regression$coefficients

## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 7.03259355 0.457842940 15.36028 1.40630e-35
```

We compute the regression coefficients

## Table 1: Information about Regression Coefficients

Coefficients	Estimate	Std. Error	t-statistic	p-value
Intercept	7.0325	$0.4578 \\ 0.0027$	15.36	<0.0001 (2.2e-16)
TV	0.0475		17.67	<0.0001 (2.2e-16)

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

Table 2: Regression Quality Indices

Quantity	Value
Residual Standard Error	3.259
R-squared	0.612
F-statistic	312.14

# Regression Analysis of Sales and TV

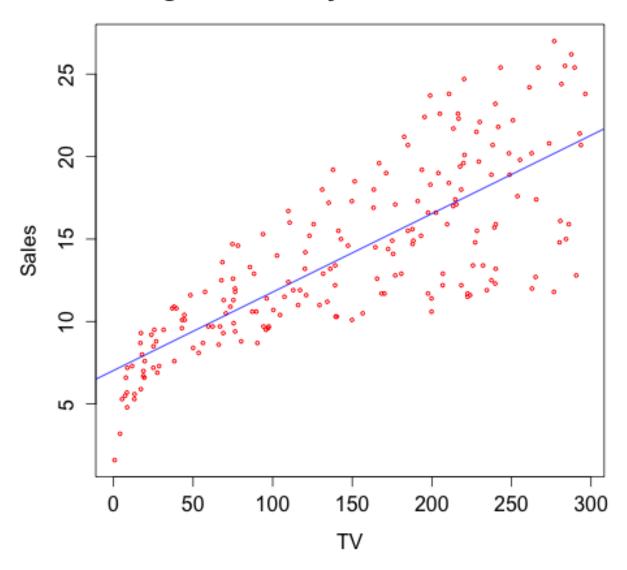


Figure 1: Scatter Plot with fitted Regression Line

## Conclusions

I have successfully reproduced the results and concluded that there must exist linear relationship between TV and sales.