

6.1: Introduction to Statistical Learning

Navona Calarco

The University of Toronto

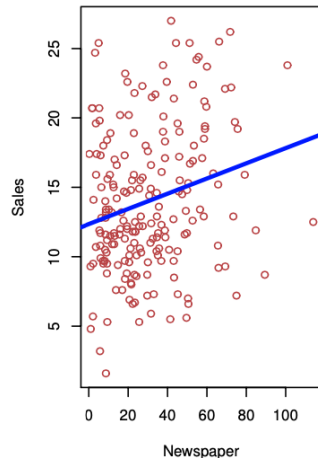
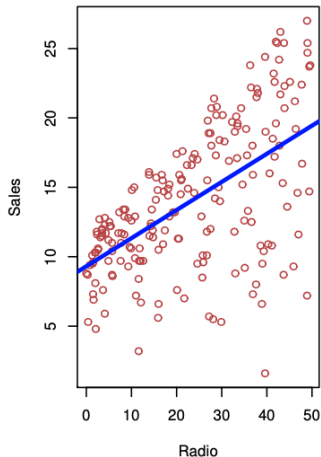
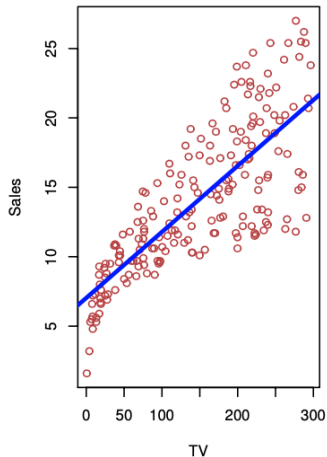
What is Statistical Learning?

Suppose we want to figure out the **association between the allocation of advertising budgets and sales** in order to increase sales for a client.

- There are three types of advertising: TV, radio, and newspaper which can be labelled X_1 , X_2 , and X_3 respectively.
- The advertising budgets are the independent variables, or **predictor variables**.
- The sales is the dependent variable, or **response variable** which we label Y . When you click the **Render** button a document will be generated that includes:
- Each observation in the dataset has a value for the predictors X_1 , X_2 , X_3 , and the response Y .

What is Statistical Learning?

The sales in relation to each of the advertising budgets are shown along with a simple fitted line for the relationships.



What is Statistical Learning?

We want to find the relationship between the predictor variables and the response variable.

We can assume that there is a relationship between $X = (X_1, X_2, X_3)$ and Y which can be written in the general form

$$Y = f(X) + \epsilon$$

- f is a fixed unknown function of the predictor variables.
- ϵ is a random error term which has mean zero.

Statistical learning is summarized by the set of approaches which are used to estimate f .

