# Multiple Regression

*Robert Edwards*

*15 February 2019*

## 1 Introduction

When fitting regression models with multiple explanatory variables, the interpretation of an explanatory variable is made in association with the other variables. For example, if we wanted to model income then we may consider an individual's level of education, and perhaps the wealth of their parents. Then, when interpreting the effect an individuals level of education has on their income, we would also be considering the effect of the wealth of their parents simultaneously, as these two variables are likely to be related.

## 2 Modelling with Two Continuous Covariates

The regression model we will be considering contains the following variables:

- the continuous outcome variable y, the credit card balance of an individual; and
- two explanatory variables x1 and x2, which are an individual's credit limit and income (both in thousands of dollars), respectively.

### 2.1 Exploratory Data Analysis

Table 1: Summary statistics on credit scores.

| Variable | Mean | SD | Minimum | 1st quartile | Median | 3rd quartile | Maximum |
|---|---|---|---|---|---|---|---|
| Balance | 520.01 | 459.76 | 0 | 68.75 | 459.5 | 863 | 1999 |
| Income | 45.22 | 35.24 | 10.35 | 21.01 | 33.12 | 57.47 | 186.63 |
| Limit | 4735.6 | 2308.2 | 855 | 3088 | 4622.5 | 5872.75 | 13913 |

**What is the mean credit Limit?**

- Mean Credit Limit = 4735.6

**What is the median credit Balance?**

- Median Credit Balance = 459.5

**What is the percent credit card holders with income greater than $57,470?**

- Mean Credit Limit = 0.25

**What is the correlation coefficient for the linear relationship between Balance and Limit?**

- Cor(Balance, Limit) = 0.8616973

Table 2: Correlation coefficient for the linear relationship between Balance and Limit

| | Balance | Limit | Income |
|---|---|---|---|
| Balance | 1.0000000 | 0.8616973 | 0.4636565 |
| Limit | 0.8616973 | 1.0000000 | 0.7920883 |
| Income | 0.4636565 | 0.7920883 | 1.0000000 |