

## Introduction to Econometrics

Econometrics is most concerned about **what actually happens in practice** in an imperfect world

Its goal is to determine the **relationship between two or more variables**

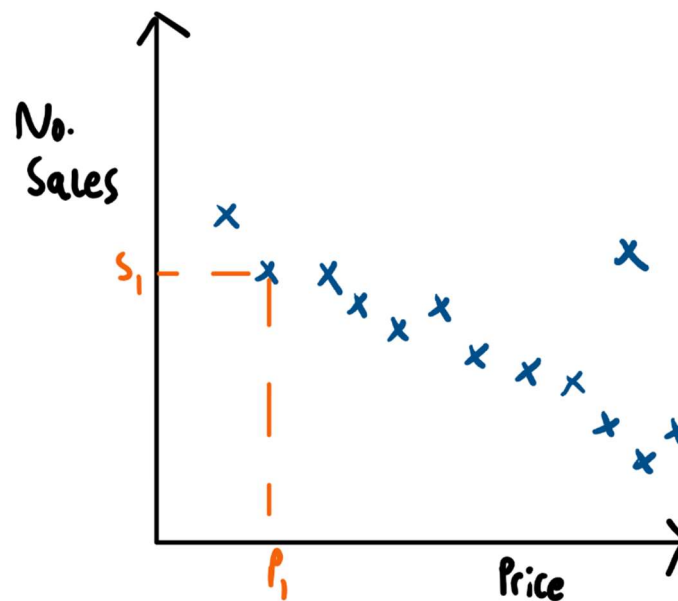
E.g. analysis on the relationship **between price and number of sales** in a business setting

### Independent + Dependent Variables

The field being **investigated** for its determinants (i.e. number of sales) is the **dependent variable**

The issue being assessed as a **possible cause** (i.e. product price) is the **independent variable**

For every possible value of the independent variable (price) there will be a **corresponding output for the dependent variable** (number of sales), as shown:



This diagram shows **different levels of a product's price** that **determines the number of sales** as result

At price  $P_1$ ,  $S_1$  amount of sales result, and every other price has its **unique level** of resulting sales

This diagram illustrates a **general negative relationship** between price + sales (i.e. higher prices link with lower levels of sales, as predicted by **standard economic theory**)

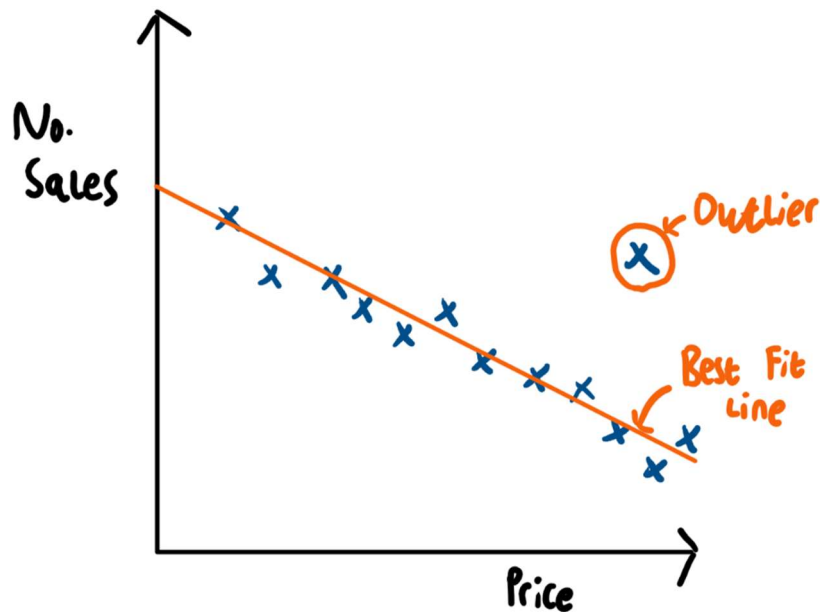
It's important to note **two things** here:

- There will **never be a definitive answer to the relationship between 2 variables**, with **exceptions** always occurring to the trend (i.e. some consumers are willing to buy the product no matter the price)
- This diagram illustrates **actual recorded data** of price compared with the number of sales from a sample, so **no assumptions or predictions need to be made here and hence no errors are present**

### Best Fit Line

We can draw a best fit line to **illustrate the relationship and determine sales for prices not recorded**

The diagram below shows the **overall trend between a hypothetical product price and the number of resulting sales** through a **best fit line** drawn based on actual data recorded in first diagram



Note that the **outlier is highlighted and removed from consideration** when determining this line

The best fit line allows for **predictions to be constructed** for other values of price + sales not recorded

Helpful for a business wanting to **determine the expected sales for a theoretical price**, allowing the firm to prepare **supply stock** accordingly

The best fit line provides a **greater understanding of cause + effect**

### Constructing the Best Fit Line

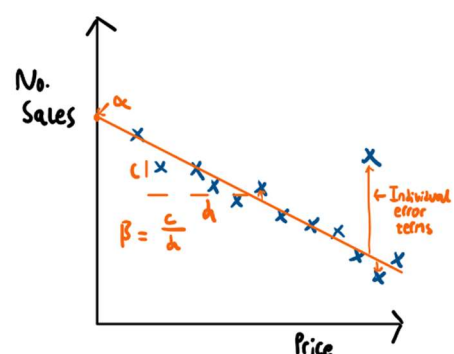
Econometrics essentially deals with **finding the best fit line to give the relationship between 2 variables**, requiring **3 pieces** of information:

- **The intercept constant ( $\alpha$ )**
  - This is the point at which the **best fit line intercepts the y axis** (i.e. what is the value of y when x is 0)
- **The slope ( $\beta$ )**
  - May be **positive or negative** depending on the **direction of the relationship**
  - May range from **horizontal to vertical** depending on the **strength of the relationship**
- **Vertical error term ( $\epsilon$ )**
  - The error term is used to show **how far the actual dependant variable data is from the best-fit line on average**

These three factors can be **summarized into an equation** for a **linear regression model** to represent the best fit line:

$$y = \alpha + \beta x + \epsilon$$

The first part of the equation ( $y = \alpha + \beta x$ ) **gets us onto the best fit line**, whilst the individual vertical error terms **take us**



**to the data point** ( $\epsilon$  represents the **average** magnitude of these individual error terms)

The error term ( $\epsilon$ ) gives the **average deviation of a data point from the best fit line**, and the **smaller** this value is, the **more accurate the best fit line is** in representing the data.