Relationship Between Variables

Siddharth Kalla69.9K reads

It is very important to understand relationship between variables to draw the right conclusion from a statistical analysis. The relationship between variables determines how the right conclusions are reached. Without an understanding of this, you can fall into many pitfalls that accompany statistical analysis and infer wrong results from your data.

There are several different kinds of relationships between variables. Before drawing a conclusion, you should first understand how one variable changes with the other. This means you need to establish how the variables are related - is the relationship linear or quadratic or inverse or logarithmic or something else?

Suppose you measure a volume of a gas in a cylinder and measure its pressure. Now you start compressing the gas by pushing a piston all while maintaining the gas at the room temperature. The volume of gas decreases while the pressure increases. You note down different values on a graph paper.

If you take enough measurements, you can see a shape of a parabola defined by xy=constant. This is because gases follow Boyle's law that says when temperature is constant, PV = constant. Here, by taking data you are relating the pressure of the gas with its volume. Similarly, many relationships are linear in nature.

Relationships in Physical and Social Sciences

Relationships between variables need to be studied and analyzed before drawing conclusions based on it. In natural science and engineering, this is usually more straightforward as you can keep all parameters except one constant and study how this one parameter affects the result under study.

However, in social sciences, things get much more complicated because parameters may or may not be directly related. There could be a number of indirect consequences and deducing cause and effect [1] can be challenging.

Only when the change in one variable actually causes the change in another parameter is there a causal relationship. Otherwise, it is simply a correlation [2]. Correlation doesn't imply causation [3]. There are ample examples and various types of fallacies in use.

A famous example to prove the point: Increased ice-cream sales shows a strong correlation to deaths by drowning. It would obviously be wrong to conclude [4] that consuming ice-creams causes drowning. The explanation is that more ice-cream gets sold in the summer, when more people go to the beach and other water bodies and therefore increased deaths by drowning.

Positive and Negative Correlation

Correlation between variables can be positive or negative. Positive correlation implies an increase of one quantity causes an increase in the other whereas in negative correlation, an increase in one variable will cause a decrease in the other.

It is important to understand the relationship between variables to draw the right conclusions. Even the best scientists can get this wrong and there are several instances of how studies get correlation and causation mixed up.

Links
[1] https://explorable.com/cause-and-effect
[2] https://explorable.com/statistical-correlation
[3] https://explorable.com/correlation-and-causation
[4] https://explorable.com/type-I-error