Confounding Variables

An observational study is when a researcher observes or questions a set of participants, and the participants are stationary.

An experiment is when a researcher manipulates and measures the effect of the manipulation.

A randomized experiment is when a set of participants are randomly assigned to participate in a condition together, or if they do all the conditions in a random order.

An explanatory variable or an independent variable is one that may explain or may cause differences in a response or dependent variable.

A confounding variable is one that

1. Affects the response variable
2. Associates with the explanatory variable

The main problems caused by a confounding variable are

1. Increases variance
2. Introduces bias

A potential confounding variable not identified in study is called a lurking variable.

Confounding variables in randomized experiments usually average out over different treatment groups, so we can conclude that changes in the experimental variable results in a change in the response variable.

Confounding variables in observational studies may explain the observed relationship between the explanatory variable and the dependent, but due to the confounding variable, we can’t conclude causation.

The fundamental rule of inference is that the available data can make inferences about a larger group if the data is considered to be a representative of the larger group.

To rule out confounding variables, we can

1. Randomize the type of treatment
   1. Randomly assigning treatment to the experimental units
2. Randomize order of treatments

A control group is a group treated identically in all respects except for the fact that it does not receive the active treatment.

Blinding is when participants do not know which treatment they received. Double blinding is when neither participants nor the researchers know which participants received the treatment. When treatments cannot be blind, a double dummy is used where group 1 is given a real treatment and a dummy treatment, and another group is given a dummy treatment and then a real treatment.

In a block design, experimental units are usually divided into blocks, and each treatment is assigned to one or more units in each block. This is more efficient if units are variable.

Matched pair designs are when two individuals or the same individual receives each of two treatments. It is important to randomize the order of treatments and use blinding if possible.

Repeated-measures designs are when blocks have varying treatments at repeated time periods.

In a completely randomized experiment, no blocks, no matched pairs and no repeated measures exist. A certain number of units are randomly assigned treatments.

In a randomized block design, units are divided into blocks of similar units, randomly assigned treatments within blocks.

Disadvantages of a good observational study is it is more difficult to establish causal links due to confounding variables.

The advantage of a good observational study is that it is more likely to measure participants in their natural setting.

Types of observational studies

1. Retrospective
   1. Participants called to recall past events
2. Prospective
   1. Participants followed into future events
3. Case control
   1. Sample of cases having an attribute compared to controls which do not to see how they differ for an explanatory variable
4. Cross-sectional
   1. Sample taken, then classified

Advantages of case control over cross sectional

1. Efficiency - may not get enough cases
2. Reduction in potential confounding variables

Difficulties in Experimental and observational studies

1. Confounding variables and their implication
2. Extending results inappropriately

Interacting variables are variables that can interact with an explanatory variable in its relationship with an outcome variable.

The Hawthorne effect occurs when participants in an experiment respond differently than they otherwise should, just because they are in an experiment.

A question of ecological validity and generalizability occurs when variables have been removed from their natural setting and measured in an artificial setting.

Relying on memory and second hand sources causes problems

1. Can be a problem in retrospective observational studies
2. Try to use authoritative reliable sources
3. Use prospective observational studies

The main way to reduce confounding is

1. Randomization
2. Restriction for individual entry
3. Matching of individuals or groups
4. Adjustment
5. Multivariate analysis