EC220 Introduction to Econometrics

Dr Canh T. Dang

Lecture Notes

Cedric Tan

September 2019

Concept
Concept. [1]

Concept
Concept. [2]

Contents

1 Introduction to Econometrics: Michaelmas Term

The initial focus of MT is applied econometrics, particularly causal questions such as "what-if" questions. Mislabelling causality as correlation can be a critical error that people make when analysing data. The course intends to teach you how to analyse data and answer economic questions using "econometrics" and data.

Examples of what-ifs are below:

- What happens to a country if it withdraws from a trade agreement?
- What is the effect of parents' education on children's education?
- What is the impact on your health if you go to a hospital?

1.1 Causality

A causes B: A contributes (or influences) to the occurrence of event B. The *cause* A is partly responsible for the *effect* B, and the effect B is partly dependent on the cause A. A can be necessary for the occurrence of B, but A can simply lead to fluctuations in B, this is still a causal relationship.

So we can take two definitions for causality:

- A is a necessary condition for B to occur
- A can cause fluctuations in B

Labels are also necessary for the structure of causation:

- Event A is called Treatment
- Event B is called the Outcome
- A third variable that causes the two events to happen is called a Confounder

We can have reverse causality - A causes B but also B causes A. An example would be *Umbrellas* and *Rain* where bringing umbrellas is caused by the possibility of rain. Sometimes timing helps to establish causality - as *Rain* happens before *Umbrellas*, *Umbrellas* cannot cause *Rain*.

Reasons that A and B are correlated:

1. A causes B (direct causation)
2. B causes A (reverse causation)
3. A and B are consequences of a common cause but do not cause each other (confounder)
4. A causes B and B causes A (bidirectional causation)
5. A causes C which then causes B (indirect causation)
6. No connection between A and B, the correlation is pure coincidence

All statistical techniques only establish associations, causation requires interpretation.
Correlation: the extent to which A and B tend to decrease and increase at the same time.

Causation can occur without correlation, here is an example for medicine:
Illness (A) can cause death (B), but nowadays healthcare (C) can eliminate the correlation between common illness and death.

1.2 Why Causality?

This is the economist's comparative advantage, the ability to infer causality from correlation.

Examples of causality are listed below with classifications like the above.

- Direct Causation:
- Reverse Causation:
- Confounder Problem:
- Bidirectional Causation:
- Indirect Causation
- Pure Coincidence:

2 Counterfactuals

Insured or not insured. Here we will see the use of counterfactuals to infer causality.

2.1 Health Insurance

What is the effect of health insurance on health? Health insurnace started as a voluntary scheme. The Obama "Affordable Care Act": attempted to compel the whole of America to buy health insurance. Similarly, is the NHS a good policy?

Question: What is the effect of health insurance on health expenditures and on health outcomes?

2.1.1 Health Insurance: Causal Questions

We want to compare:

- The health of someone with insurance Y_i^T
- The health of the same person without insurance Y_i^C

Though this is not a realistic possibility simply because we cannot have the exact same person simultaneously adopting both approaches. Although the data might show that there is a correlation between Health Insurance and Health, there could be other factors. For example, the United States spends more of its GDP on health care than do other developed nations, yet Americans are surprisingly unhealthy. However, America is also unusual in that is has no universal health insurance scheme. Perhaps there is a causal connection here.

- The causal effect of insurance: heaving health insurance may lead to better health because of better health care
- The reverse causal effect: the less healthy are more likely to buy insurance
- Confounder effect: the more educated tend to buy insurance more often and they know how to live healthier (possibly)

- Pure coincidence: another possibility

Many of the working, prime-age poor, however, have long been uninsured. In fact, many uninsured Americans have chosen not to participate in an employer-provided insurance plans. Using the National Health Interview Survey (NHIS), that rates health from poor (1) to excellent (5), we can gauge, with characteristics and analysis of whether or not someone has health insurance or not, the possibility of a causal effect.

Terminology to take into account when looking into this case study:

- Health insurance coverage for individual i is described by a binary random variable; we call this the *treatment*

$$D_i = 0, 1$$

- The outcome of interest, a measure of health status denoted by Y_i and is the index proposed by the NHIS i.e. on a scale from 1 to 5
- Potential outcomes: describes what would have happened to someone under the scenarios where they had or had not been insured.
- Those with insurance are called the *treatment group*
- Those without insurance are called the *control group*; a good control group reveals the fate of the treated in a counterfactual world where they are not treated

2.1.2 Fruitless and Fruitful comparisons

The critical thing to be mindful of is keeping things equal i.e. *ceteris paribus*. Comparisons of people with and without health insurance are not apples to apples; such contrasts are apples to oranges, or worse. Among other differences, those with health insurance are better educated, have higher income, and are more likely to be working than the uninsured. Many of the differences in characteristics between the two classes are large and most are statistically precise enough to rule out the hypothesis that these discrepancies are merely chance findings. Thus it won't be surprising that most variables listed are high correlated with health as well as with health insurance status. Some other highly associated characteristics

- Education
- Family income
- Age
- Employed vs Unemployed (less powerful indicator)

We use the Robert Frost metaphor for understanding the potential outcomes.

- For everybody there are two potential outcomes **but only one is actually observed**:

$$Counterfactual = \begin{cases} Y_{1i} & if\ D_i = 0 \\ Y_{0i} & if\ D_i = 1 \end{cases}$$

- The treatment effect is the difference between the actual outcome and the counterfactual given the person has insurance. Thus the effect of having insurance for individual i is $Y_{1i} - Y_{0i}$

To cement this down even further, lets consider the story of two students (K and M) going to MIT who have the option to subscribe to the MIT health insurance plan. Upon reflection, K decides that the MIT insurance is worth paying for since he fears he might get sick. Let's say that for K $Y_{0i} = 3$ and $Y_{1i} = 4$ for $i = K$ For K, the causal effect of insurance is one step up on the NHIS scale:

$$Y_{1,K} - Y_{0,K} = 1$$

M is also coming to MIT but does not fear sickness whatsoever. Not concerned about the winters in Boston, M does not think she will fall sick easily so opts not to buy the MIT insurance. Because $Y_{0,M} = Y_{1,M} = 1$ we have:

$$Y_{1,M} - Y_{0,M} = 0$$

We can summarise this data into a table below:

	K	M
Potential outcome without insurance: Y_{0i}	3	5
Potential outcome with insurance: Y_{1i}	4	5
Treatment (insurance status chosen): D_i	1	0
Actual health outcome: Y_i	4	5
Treatment effect: $Y_{1i} - Y_{0i}$	1	0

There is a selection problem for Jones and Maria. This is because:

- We have:

  $$E[Y_i|D_i = 1] - E[Y_i|D_i = 0]$$

  Which we can call the observed difference in average health.

- This is equivalent to:

  $$E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 0] \text{ (Average treatment effect on the treated) } + \\ E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0] \text{ (Selection bias)}$$

  Which we can call the average treatment effect on the treated plus the selection bias

3 Experiments

How do we solve the selection bias? One method is through random assignment of the treatment factor. That is to say that if D_i is randomly assigned, it is (statistically) independent of potential outcomes such that there is no difference between $E[Y_{0i}|D_i = 1]$ and $E[Y_{0i}|D_i = 0]$. That means we eliminate out selection bias factor $E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0] = 0$.

3.1 Developing the experiment

Remember K and M from the MIT observation. If we were to randomly assign treatment in this case, the sample size would be too small. Further, characteristics between K and M would be too different and we would not be able to check for balance.

Thus, we can use the law of large numbers (LLN) to get a larger sample to test on. The theorem states:

The larger the sample, the closer the sample average will become to the population mean. By making the sample large enough, sample statistics and population can be brought as close to what we want.

By having a large sample, we can wash out the random characteristics within our sample size making our experiment more indicative. Our aim is to remove the selection bias when inferring a causal effect, thus, with enough subjects, we can assure balance across the groups. Balance means no *systematic differences* in characteristics e.g. we hold a similar fraction of men and women in each group. We can balance for observed characteristics but we can only assume randomisation would work for unobserved characteristics, such as motivations, incentives and inner behaviours.