

STA 3180 Statistical Modelling: Causal Inference

STA 3180 Statistical Modelling: Lecture Notes on Causal Inference

Introduction

Causal inference is the process of inferring cause-and-effect relationships between variables. It is a key component of statistical modelling and is used to draw conclusions about the effects of one variable on another. This lecture will cover the basics of causal inference, including key concepts, definitions, coding examples, and practice multiple choice questions.

Key Concepts

The following are some key concepts related to causal inference:

* **Counterfactuals**: Counterfactuals are statements about what would have happened in the absence of a particular event or condition. For example, "If I had not taken the bus, I would have been late for work."

* **Observational Studies**: Observational studies are studies that observe the relationship between two variables without manipulating either one. These studies are useful for understanding the relationship between variables, but they cannot be used to draw causal conclusions.

* **Experimental Studies**: Experimental studies are studies that manipulate one variable to observe the effect on another. These studies can be used to draw causal conclusions.

* **Randomized Controlled Trials (RCTs)**: RCTs are experiments in which participants are randomly assigned to different treatment groups. These experiments are the most reliable way to draw causal conclusions.

Definitions

The following are some key definitions related to causal inference:

* **Cause**: A cause is an event or condition that produces an effect.

* **Effect**: An effect is an event or condition that is produced by a cause.

* **Correlation**: Correlation is a measure of the strength of the relationship between two variables. It does not necessarily imply causation.

* **Causality**: Causality is the relationship between a cause and its effect.

*****Confounding Variable**:** A confounding variable is a third variable that is related to both the cause and the effect. It can lead to spurious correlations and make it difficult to draw causal conclusions.

Coding Examples

Start of Code

The following code shows how to calculate the correlation between two variables using the ``corr()`` function in R:

```
```\n\nx <- c(1, 2, 3, 4, 5)\ny <- c(2, 4, 6, 8, 10)\n\ncorr(x, y)\n```\n
```

End of Code

### ## Practice Multiple Choice Questions

Q1: Which of the following is NOT a key concept related to causal inference?

- A) Correlation
- B) Counterfactuals
- C) Causality
- D) Randomization

Answer: D) Randomization

Explanation: Randomization is a key concept related to experimental studies, but it is not a key concept related to causal inference.