Lecture 1.1

Course Overview and Evidence from Data

"*I keep saying that the sexy job in the next 10 years will be* **statisticians**, *and I'm not kidding.*"

Hal Ronald Varian, Chief Economist at Google
August 6, 2009, The New York Times

Course Overview

Topics: basic probability theory, distributions and properties, sampling methods, EDA, estimation, hypothesis tests, regression, experimental design, transform methods, model construction, and joint distributions.

Assumed Background: The course assumes a basic understanding of linear algebra and calculus, as covered in MATH1051. It is a student's own responsibility to fill in any gaps in their assumed knowledge.

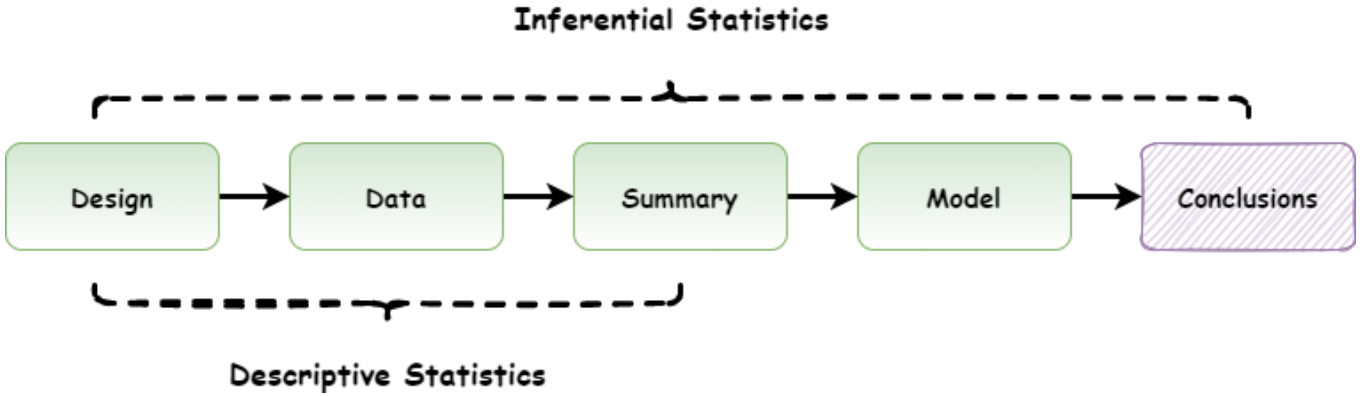
Learning Resources

- Lectures (in-person and via zoom)
- Lecture slides, notes, chapter exercises, practicals, etc...
- Practicals start in week 2
- RStudio used for data analysis
- Ed discussion board
- Office-hours via Zoom

Assessment

- Three assignments (10% each)
- Final exam (70%) (in-person and held on campus).

Statistical Study



- **Design** the study
  - observational or experimental?
  - formulate a statistical question
  - what data to collect
  - how to collect it/who to collect it from
- Perform the study to **collect data**
- **Summarize/visualize** the data
- Make a **model** for the data
- Draw **conclusions** on the population

Improving 3D Spatial Reasoning

A recent study investigated whether games like Minecraft can be used to improve a persons 3D spatial reasoning skills. [1]

Spatial reasoning skill was assessed with the Mental Rotation Test (MRT) giving a score out of 40.

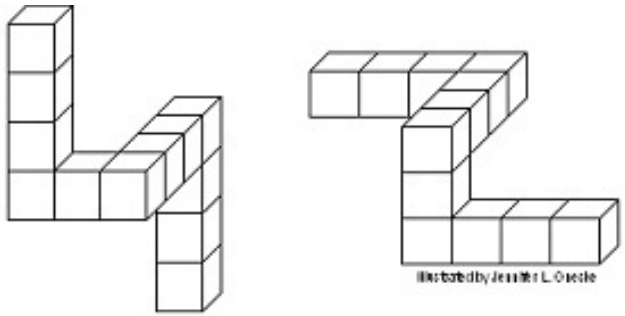


Figure 1: Based on Shepard & Metzlar's 'Mental Rotation Task'

[1] Carbonell-Carrera et al. (2021) *Minecraft as a block building approach for developing spatial skills*, Entertainment Computing 38: 100427

Study Design

34 participants were recruited to the study and randomly allocated to either Group A or Group B.

All participants completed the MRT at the beginning of the study.

Two weeks later participants in Group A completed a number of tasks in Minecraft. The session lasted for approximately 4 hours. Participants in Group B did not.

All participants complete the MRT again at the end of the study and the change in their score recorded.

Summary of Terminology

- An *experiment* involves actively applying treatments to subjects and observing their responses.
- An experimental *treatment* is a combination of factors at different levels. In other words, a factor is a general type or category of treatments. Different treatments constitute different levels of a factor.
- The variables describing the treatments are the *explanatory variables* in the study.
- The *response* from an experiment is the variable/s of interest.

Comparative Experiments

- *Comparative* experiments (treatment group vs. control group) are desirable to eliminate the *placebo* effect.
- *Randomization* helps remove possible bias in a comparative experiment.
- A *blind* experiment is one in which the subjects do not know which treatment they are getting.
- In a *double-blind* experiment the experimenters also do not know which treatment the subjects are receiving.

Working with Data in RStudio

Let's use RStudio to compare the mean scores of the two groups.

Does a mean difference of 2.47 points give evidence that the Minecraft session can improve a persons MRT score?

Evidence

Two possible explanations/hypotheses:

H_0 : The Minecraft session really has no effect on MRT scores and the observed difference of 2.47 points was just due to the variability in MRT scores.

H_0 : The difference of 2.47 points arose because the Minecraft session does improve MRT scores.

Randomisation Test

Suppose the first explanation is correct, i.e., there is really no difference. In other words, whether a person was in Group A or B is completely unrelated to the change in their MRT score.

In that case, if we randomly divide the participants into two groups, irrespective of whether they played Minecraft or not, then we expect to see a difference in mean that is similar to, maybe even larger than, 2.47 (but certainly not much smaller).

However, if the second explanation is correct, i.e., playing Minecraft improves the MRT score, if we divide people into two groups, irrespective of whether they played Minecraft or not, then we expect to see a difference in mean that is smaller than 2.47.

Hypothesis Testing

- The randomisation test is one example of statistical *hypothesis testing*
- The first explanation we gave is called the *null hypothesis*, H_0, of the test ("status quo", "no association", "no change", or "no effect")
- If H_0 is true, the probability of obtaining such unusual (or more unusual) data is the *p-value* of the test [here, $p \approx 0.09$]
- We cannot say *"the probability that H_0 is true is $p = 0.09$..."*

P-values

- A small p-value suggests H_0 may be wrong, giving evidence for the second explanation
- A large p-value suggests that the data are consistent with the null explanation, giving inconclusive evidence of an effect

Strength of Evidence