Experiment Design for Computer Sciences (01CH740)

Topic 00 - Course Introduction

Claus Aranha
caranha@cs.tsukuba.ac.jp

University of Tsukuba, Department of Computer Sciences

April 29, 2020

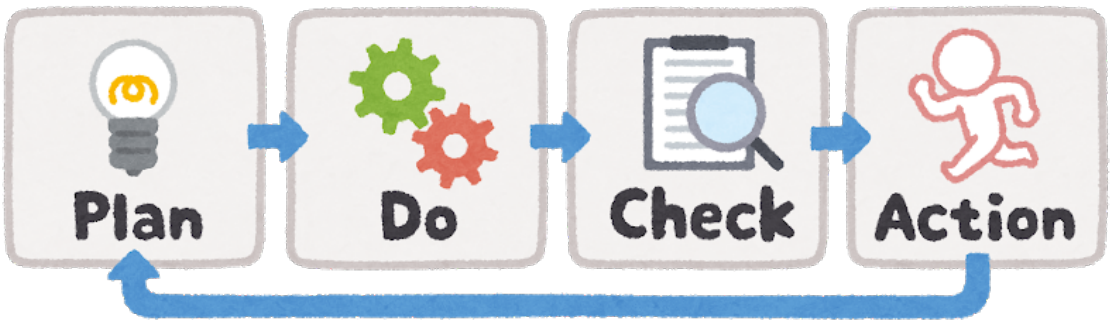
Version 2020.1

What is this course about?

From the syllabus

The collection and analysis of data through experiments is one of the cornerstones of the scientific method. In this course, we study the general philosophy and methods behind experimentalism: Why do we perform experiments, what is a good/rigorous experiment, how to plan and design a rigorous experiment, and how to perform statistical analysis on experimental data.

TL;DR: We study how to properly plan, execute, analyse and interpret an experiment (PDCA cycle!).



Motivation

In my years in Tsukuba, I have seen students commit the same errors many times in their presentations:

- Experiments that do not control for noise factors;
- Comparative experiments that are not fair;
- Experiments that are not reproducible;
- Experiments that are not falsifiable;
- Unclear experimental conditions and assumptions;

The final result is the same: The experimental result does not support the conclusions of the student.

Motivation

These problems are not limited to students

- Lack of rigorous experimentation protocol is common in Computer Science – It is not only students that do it!
- The situation has gotten much better in the past five years – you will usually have a paper rejected if you don't do proper statistical analysis of your experiments;
- But we still have a long way to go, specially compared to Bio and Psych areas;

We need to be careful to avoid a **reproducibility crisis** in the computer sciences!

What will you learn in this course?

- How to design experiments for your own research;
- How to analyse your own experimental data in a robust way;
- How to read and criticize scientific experiments for robustness;

If your research topic requires using computational experiments to show the efficacy of your proposed methods, this course will be specially useful for you.

Course Topics

- What is an experiment:
  - Characteristics of a robust, reliable, reproducible experiment;
  - How to design an experiment to answer a scientific question;
  - Pitfalls to avoid when designing an experiment;
- Statistical tools for analyzing experimental data:
  - Basic statistics for data analysis and visualization;
  - Statistical Inference
  - Statistical testing for single, paired, and multiple sample testing;

Course Materials

The materials of this course can be found on the "manaba" system: `https://manaba.tsukuba.ac.jp/ct/course_1354290`. If you are a member of the University of Tsukuba, but not enrolled in this course, you can still access the manaba course using the self registration code: 5685116.

If manaba is not available, or you are not part of University of Tsukuba, you can access most of the materials on the github repository for this course: `https://caranha.github.io/ExperimentDesignCS/`. However, report submission and some of the recommended readings are only available through manaba.

Course Materials

Lecture Notes

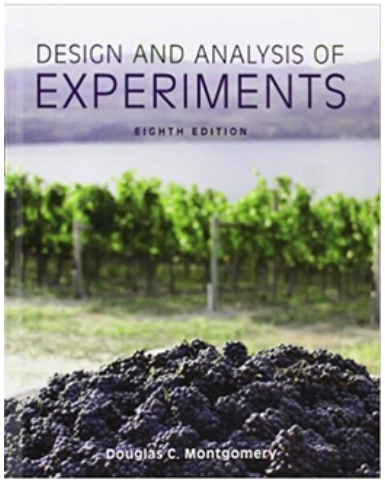
The main materials for this course are these lecture notes. Make sure to read them and ask questions if anything is unclear!

The lecture notes were produced based on the "Design and Analysis of Experiments" material produced by Felipe Campelo. You can reach the original lecture notes on: `https://github.com/fcampelo/Design-and-Analysis-of-Experiments`

All good ideas are thanks to Felipe (and other contributors) all errors are my own :-) (Please submit errors as github issues!)

Course Materials

Books and Links



Many topics in this course are explored in much more depth on "Design and Analysis of Experiments", by Douglas C. Montgomery.

In manaba there will be a much larger list of resources for extra study available. Make sure to check it out!

Course Limitations

This course is just an **introduction** to the ideas and tools of statistical analysis. Each experiment is different, and the specifics of your experiment might require different tools that are not described here.

I expect that students will see this course as a starting point to get familiar with the basic concepts of statistical testing, and look for more specific information for their unique needs.

Weekly Schedule

The course is composed of **regular lectures** and **review and discussion lectures**.

- **Regular lectures**: Download the Lecture Notes from manaba or github. See the video link for a detailed discussion. Study the recommended resources.
- **Review and discussion**: Lecture will take place using an online meeting tool (TBD). Students may ask questions and receive feedback about their reports.

Use the "Forums / 掲示板" area on manaba to ask questions too!

Course Calendar

- 5/1 : Introduction (Today)
- 5/8 : Point and Interval Indicators
- 5/15 : Inference Testing I - **Deadline Report 1**
- 5/22 : Inference Testing II
- 5/28 : Inference Testing III
- 6/5 : Case Study, Review and Discussion
- 6/12 : Power Analysis and Sample Size - **Deadline Report 2**
- 6/19 : Anova and Post-hoc testing
- 6/26 : Blocking and Parameter Selection
- 6/27 : Case Study, Review and Discussion
- 7/3 : **Deadline Report 3**

Grading

The grading of this course is based on three reports. (No final examination).

- Report 1: Submit before lecture 3 (30%)
- Report 2: Submit before lecture 7 (30%)
- Report 3: Submit 1 week after lecture 10 (40%)

Each report is a "mini paper", including an experiment description, data acquisition, data analysis and discussion of the results.It must use the techniques and ideas introduced in this course.

Report requirements

- The reports must be submitted in English;
- Besides the report text (in PDF format), the reports must include all information necessary for reproducing the analysis.
  - Data aquisition protocol and/or code for data generation;
  - Code for reproducing statistical analysis and figures, including data files[1];
- The report will be graded by the correctness of the analysis, and the quality of the discussion, but **not** on the results of the experiments (negative results are ok!)
- You are encouraged (not required) to write reports based on your research!

[1] It is easy to provide this data if you generate your report using R Markdown, but this is not reqired

Computer Science English Program (CSE)

The (CSE) supports a master degree fully in English. If you plan to take many classes in English, I encourage you to enroll in the this program.

Send an e-mail with your name (Roman letters and Kanji) and student ID to
s-g30@cs.tsukuba.ac.jp

For more information, see the orientation material at the "New Student Orientation" course on manaba (xx20011-002).

About the Lecturer

- Name: Claus Aranha;
- Country: Brazil;
- Research Topics:
  - Evolutionary Algorithms;
  - Artificial Life;
- Hobbies:
  - Game Programming;
  - Astronomy;
- webpage:

  `http://conclave.cs.tsukuba.ac.jp`

About these Slides

These slides were made by Claus Aranha, 2020. You are welcome to copy, re-use and modify this material.

These slides are a modification of "Design and Analysis of Experiments (2018)" by Felipe Campelo, used with permission.

Individual images in some slides might have been made by other authors. Please see the references in each slide for those cases.

Image Credits I

[Page 2] PDCA icon from `https://www.irasutoya.com`
[Page 9] "Design and Analysis of Experiments" book cover image from Amazon.
[Page 10] Sprout image from `https://www.irasutoya.com`