Experiment Design for Computer Sciences

00 - Course Intro

Claus Aranha

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

Computer Science Department

Version 2.11.T

Tsukuba, 2016

"*The combination of some data and an aching desire for an answer does not ensure that a reasonable answer can be extracted from a given body of data.*"

John W. Tukey (1915 – 2000)
American mathematician

Image source: `http://www-history.mcs.st-andrews.ac.uk/Mathematicians/Tukey.html`

Course Overview

Objectives

- To develop advanced skills in designing experiments, defining and testing hypotheses, and performing statistical data analyses within one's field of interest;
- By the end of this course, the student should be able to:
  - Plan experiments related to his/her work;
  - Perform appropriate statistical analyses of the data obtained from the experiment;
  - Develop sound conclusions based on the available data;
  - Identify the problems and limitations of his/her own experiments, and suggest improvements;
  - Perform critical interpretations of other experimental methodologies and results reported in the literature.

Course Overview

Course Structure

- Lectures (10 weeks): discussions about several aspects and techniques for design and analyses of experiments. Theory, application examples and *computational case studies*;
- One Team Project/Presentation;
- One Individual Report;
- One Case Study;

Course Overview

Course Structure

Evaluation criteria

Item	Type	Value
Case studies	Individual Report	40
Research Report	Individual Report	30
Research Project	Report and presentation	30

Other relevant Information

- Lectures slides, example R files, data, etc. available at `https://git.io/vVpvC/`
- MANABA site for submitting reports and questions. Code: 5718071
- Software/services used: R (`http://cran.r-project.org/`), GitHub (`http://github.com/`).

Course Overview

Course Bibliography

Main:

- *Lecture Notes on Design and Analysis of Experiments*. Online: `http://git.io/vVpvC` Version 2.11.t; Creative Commons BY-NC-SA 4.0.
- D.C. Montgomery, G.C. Runger (2010), *Applied Statistics and Probability for Engineers*, John Wiley & Sons.
- Michael J. Crawley (2007), *The R Book*, Wiley.

Additional:

- D.C. Montgomery (2012), *Design and Analysis of Experiments*, John Wiley & Sons.
- B. Caffo (2015), *Statistical inference for data science*, LeanPub - `https://leanpub.com/LittleInferenceBook/`
- J.J. Faraway (2002), *Practical Regression and Anova using R* - `http://goo.gl/ewMWL`
- D. Wiens (2005), *Introduction to Design and Analysis of Experiments* - `http://goo.gl/hZXg1`

Course Overview

Required / Desired background

This is a course on *applied* experimental design and analysis. As such, a large portion of the course is dedicated to case studies in which the student will design experiments, collect (simulated) data, perform inference and report his or her analysis.

It is **strongly reccomended** that the student should complete the free online course *R Programming*[b] **before the end of the second week** of the semester (except if the student is already fluent with R).

It is also **strongly reccomended** that the student should complete the free online course *Reproducible Research*[c] **before the end of the first month** of the semester (except if the student is already fluent with writing reports using R Markdown).

---

[b] `https://www.coursera.org/course/rprog`

[c] `https://www.coursera.org/course/repdata`

About this material

Conditions of use and referencing

This work is a derivative work of
Felipe Campelo (2015), *Lecture Notes on Design and Analysis of Experiments*.
Online: `https://github.com/fcampelo/Design-and-Analysis-of-Experiments`
Version 2.11;

I'd prefer if you referenced the original work above, but this work can be referenced as:

Claus Aranha (2016), *Lecture Notes on Experiment Design for Computer Sciences*.
Online: `https://github.com/caranha/Design-and-Analysis-of-Experiments`
Version 2.11.t; Creative Commons BY-NC-SA 4.0.

About this material

Acknowledgments

Special thanks to:

- Prof. Felipe Campelo, PhD. from UFMG, for organizing the original material, and making it available under a Creative Commons License