Experiment Planning and Design

Lecture 8: Course review and Scientific Communication

Claus Aranha
caranha@cs.tsukuba.ac.jp

Department of Computer Science

2015-06-23

Class 1: What is Science

- Method to learn more about the world
  - “Quest for truth”
  - Social implications
- Scientific and Unscientific Questions
- Scientific Method and Scientific Community
- Science vs Engineering
- Science vs Theory
- Computer Science

Class 2: Experimental Science Concepts

- What is an experiment: A methodical way to obtain facts about the world
- Experimental Factors - controllable and uncontrollable variation;
- Goals and Methodologies of an experiment;
- Types of experiment:
  - Historical
  - Observation
  - Controlled
- Experimental Rigour and Experimental reproducibility;
  - Fairness
  - Clarity
  - Completeness

Class 3: Point and Interval Estimators

- Experimental Data can be seen as a distribution (Sample vs Population)
- Point Estimators: functions that describe this population;
  - Examples: Mean, Median, Variance, Skew
  - Point Estimators have their own distributions, variances, etc.
- Interval Estimators: Estimate the range of values produced by a point estimator;
  - Confidence Interval;
  - Prediction Interval;
  - How to calculate these;
- In general we want to describe our data as intervals, because it includes the value of a point estimator, and a prediction of how much it can vary;
- The role of the Central Limit Theorem

Class 4: Hypothesis testing for one Mean

- Hypothesis testing as a mean to draw conclusions from data
- How to formulate a Null Hypothesis and an Alternative Hypothesis
- Hypothesis as statements about probabilities (α, p, critical values)
- Hypothesis over the mean of a population
- The Student T Test
- Assumptions of the T-test: Normality, Variance, Independence of Residuals;
- Errors Type I and Type II

Class 5: Special cases of Hypothesis testing I

- Paired Design;
- Non-parametric testing;
- Testing on the difference of two means;
- Proportion Testing
- Calculating α, β and sample size;

Class 6: Special cases of Hypothesis testing II

- Multiple samples and ANOVA;
- Experimental Models;
  - Completely Randomized Model and the Latin Hypercube
  - Blocking Model
  - Factorial Model

Case Studies

Let's see how these concepts fit together in two case studies

Case 1: Comparison of Drilling Riser Configuration

files/CS04.pdf

Case 1: Comparison of Drilling Riser Configuration

How to approach this problem?

- Question of interest 1: Is there a riser configuration (from options A, B, C) with a higher Meant Time to Failure (MTTF) than the standard configuration?
- Question of interest 2: If the answer to the question of interest 1 is YES, which configuration option has the highest Mean Time to Failure?

Main Concerns

- Testing a riser configuration is expensive. The standard configuration has historical measures recorded, but each new observation will cost us money.
- Therefore, it is important to calculate beforehand the minimum number of observations that will be required.

Case 1: Comparison of Drilling Riser Configuration

What tests to execute?

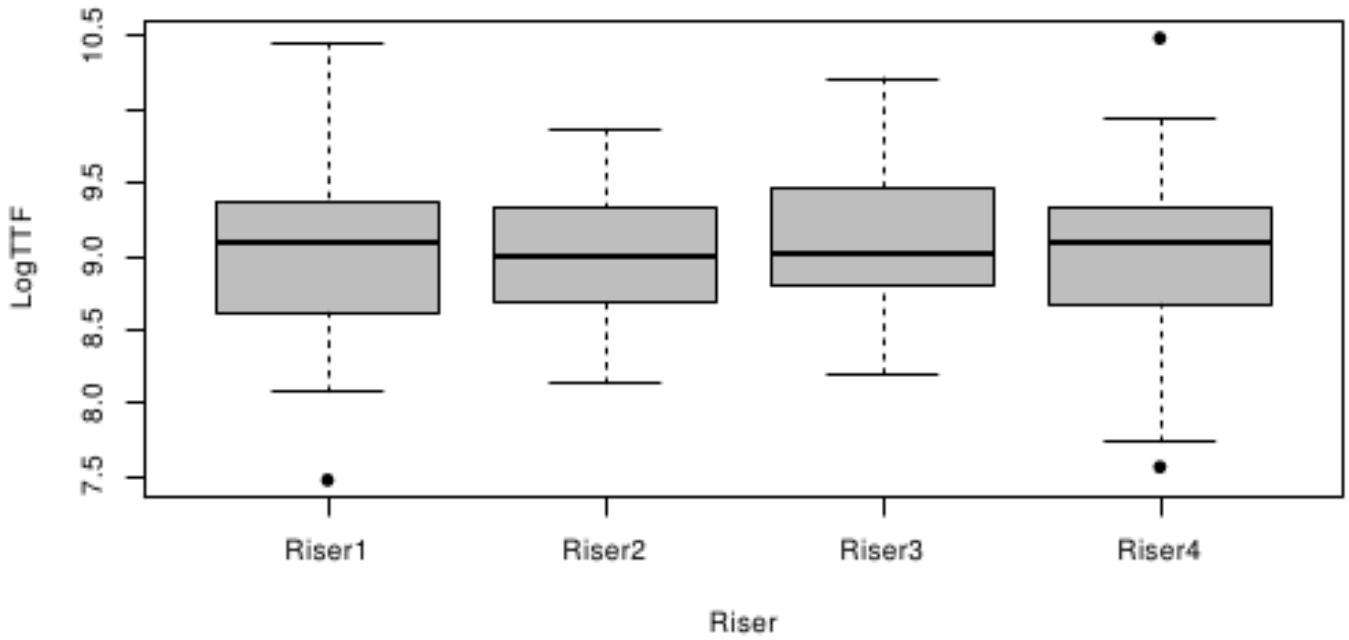
- Because we are interested in the differences between multiple treatments, we initially need an ANOVA test to find out if they are different at all.
- If the ANOVA test indicates a significative difference between the treatments, a follow up “all versus one” test would tell us which treatments are better than the standard.

- Because of the cost associated with each observation, we need to calculate in advance the sample sizes needed for the ANOVA and for the All versus One tests.
- This calculation can be done based on the required values for α, β and γ.

Case 1: Comparison of Drilling Riser Configuration

```
> model <- aov(LogTTF~Riser,data=yriser)
> summary.aov(model)

             Df Sum Sq Mean Sq F value Pr(>F)
Riser         3   0.33  0.1108   0.439  0.725
Residuals   248  62.51  0.2521
```



The ANOVA test tells us that the risers do not significantly differ from the

Scientific Communications

Your science is only as useful as how many people know about it.

So how do we make sure our results are actually used by other people?

Clarity

What do we mean by clarity?

- Experiment Clarity
- Paper Clarity
- Research Clarity

Experiment Clarity

- Is the experiment reproducible? (big one)
- Is the data clear and available?
  - Artificial Data;
  - Real Data;
  - Benchmark Data;
- Are the conclusions clear?

How do we write an experimental section of a paper to ensure clarity?

Paper Clarity

- Beginning from the basics: Do not forget the spell checker!
- Have other people proofread your paper (regardless of the language)
- Follow the templates (unless you have a VERY good reason not to)

Research Clarity

How organized is your research?

- Keeping a research diary
- Research Versioning (vs software versioning)
- Keep old versions of your experiment around!

Reasons for clarity

In Medical sciences, it was found that over 80% of the results were not reproducible after 20 years...

I don’t wanna think about this proportion in computer sciences.

Plagiarism

- Basic Plagiarism
- Small Plagiarism (figures, data, quotes, etc)
- Self-plagiarism

Fraud

- How to avoid fraud by keeping your data open
- Taking care with claims
- Taking care with data and results

Experimenting with people

- Molchan Experiment
- Stanford Prison Experiment
- Facebook Experiment

Final Report

Important Dates:

- June 30th (Tue): Question Class - make questions about your report here!
- July 7th (Tue): Final reports due!

Final Report Format

Upload a Zip file to manaba that include the following files:

- PDF file with your report, including the description of the research, the relevant experimental question, the experimental design, results etc (see the case study in this slides for info)
- R file with the code for any and all calculations that you did in your report.
- Any data files used in your calculations.

Final Report - Extra information:

Format

There is no fixed format, but please remember that I want a well written pdf file. A suggestion for a report template is uploaded to Manaba.

Case Studies

If you cannot use your own research/data for the final report, I have uploaded two case studies to Manaba. Choose one of those for your final report instead.

Required Reading!

"Experimental Computer Science: The Need for a Cultural Change"
Dror G. Feitelson

`http://www.cs.huji.ac.il/~feit/papers/exp05.pdf`

End of this class!

- Let's take some time to do the course evaluation.
- Questionnaries in English and Japanese are available

Next Class

Questions and Report checking!