UFMG
UNIVERSIDADE FEDERAL
DE MINAS GERAIS

Design and Analysis of Experiments

02 - The Role of Experimentation

Version 2.11

Felipe Campelo
http://www.cpdee.ufmg.br/~fcampelo

Graduate Program in Electrical Engineering

Belo Horizonte
March 2015

“*There may be some beliefs that cannot be decided by data, but such beliefs are dogmas that lie (double entendre intended) beyond the reach of evidence.*”

John K. Kruschke

American mathematician and cognitive psychologist

Image: `http://www.amazon.com/John-K.-Kruschke/e/B00474BQRQ`

An experiment can be characterized as a test (or a series of tests) wherein changes are introduced in the state of a system or process, enabling the observation and characterization of effects that can occur as a result of these changes.

Usually performed with an objective in mind:

- Uncovering influential variables in a given system or process;
- Determining desired values for certain parameters
- Characterize behavior of the system or process under study.

Experiments

Data gathering

- Retrospective study;
- Observational study;
- Designed experiment;

Characteristics

- Use of historical data;
- Investigating correlations;

Problems

- Data representativeness;
- Availability of data;

Experiments
Data gathering

- Retrospective study;
- Observational study;
- Designed experiment;

Characteristics

- Observation of the system with minimal disturbance;
- Investigation of usual behaviors;

Problems

- Low representativeness of extreme cases;
- Low variability can affect observation of interesting effects;

Experiments

Data gathering

- Retrospective study;
- Observational study;
- Designed experiment;

Characteristics

- Introduction of deliberate changes in the system;
- Inference on the *causality* of the effects;

Problems

- Requires rigorous experimental design and data analysis;
- Usually more expensive.

Experimentation strategies

Educated guessing

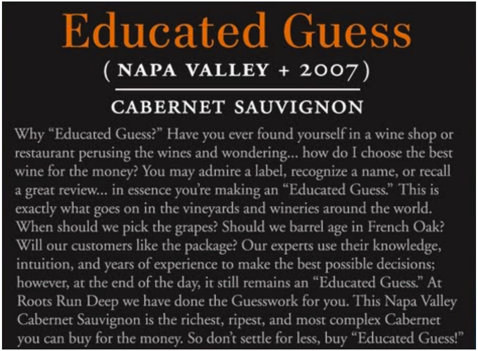
- Select arbitrary combination of levels for the factors;
- Test and observe behavior;
- Change one or two factors at a time, then re-test;

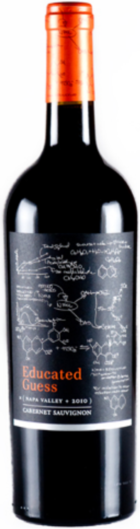
- Widely used in engineering;
- Can achieve good results, but has a lot of limitations;

Experimentation strategies

Educated guessing

- Select arbitrary combination;
- Test an observe;
- Change and re-test;



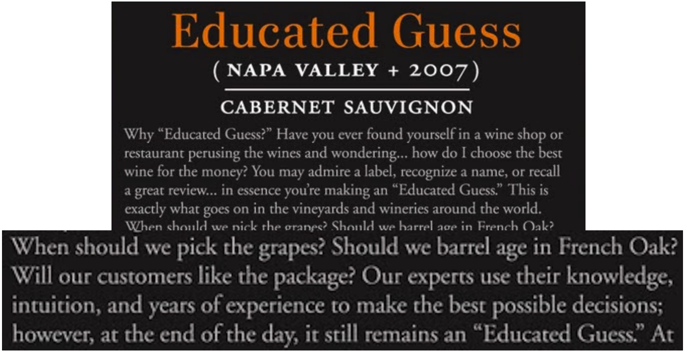


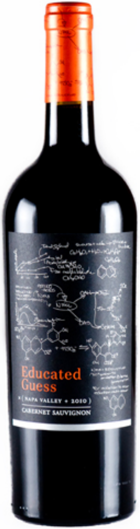
Images (c) Roots Run Deep Winery: `http://www.rootsrundeep.com/educated_guess.html`

Experimentation strategies

Educated guessing

- Select arbitrary combination;
- Test an observe;
- Change and re-test;





Images (c) Roots Run Deep Winery: `http://www.rootsrundeep.com/educated_guess.html`

Experimentation strategies

COST: Change One Separate factor at a Time

- Select a reference point;
- Change each factor individually, keeping all others constant;

- Widely used in practice;
- Can achieve good results as long as there are no interaction effects;

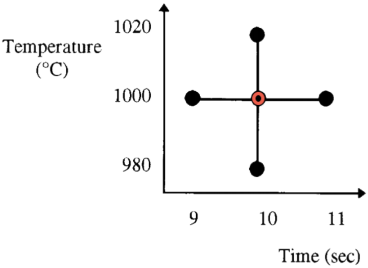


Image: (c) D.C. Montgomery

Experimentation strategies

Factorial designs

- Select **levels** for each factor;
- Vary the factors simultaneously, in a systematic way;

- Estimation of main effects and interactions;
- Greater precision in the effect estimates;
- More efficient use of resources (information/observation);

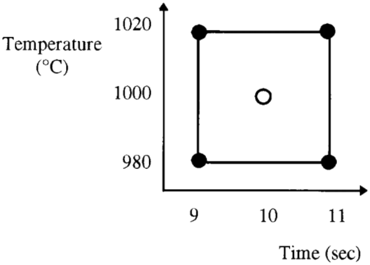


Image: (c) D.C. Montgomery

Fundamental principles

Design of experiments (DoE)

Process of designing data gathering protocols to enable accurate analyses by statistical tools, capable of supporting sound and objective conclusions.

- Applicable to systems and processes subject to noise, experimental errors, uncertainties, etc.
- Necessary for the conclusions to have a quantifiable meaning;
- Helpful in avoiding errors due to personal biases or other artifacts of experimentation and analysis.

Design of the experiment

- Scientific/technical question of interest;
- Selection of variables and values;
- Definition of the desired confidence level;
- Sample size calculations;
- Determination of protocols for data gathering;

Statistical analyses of the data

- Calculation of a test statistic;
- Validation of the assumptions of the statistical model;
- Calculation of the magnitude of effects;
- Drawing of conclusions and recommendations;

Fundamental principles

Design of experiments (DoE)

- Repetition and replication;
- Randomization;
- Blocking.

- Repeated measurements - estimation of within-group variability;
- Replication - estimative of the experimental error;
- Greater precision in estimating the model parameters;

Fundamental principles
Design of experiments (DoE)

- Repetition and replication;
- Randomization;
- Blocking;

- Avoids contamination of the data by order-dependent effects such as:
  - Heating effects;
  - Wear and tear effects;
  - External interferences;

Fundamental principles
Design of experiments (DoE)

- Repetition and replication;
- Randomization;
- Blocking;

- Isolation of nuisance variables (those that influence the response, but are not interesting for the analyses) that can be controlled;
- Improvement in the estimation of effects for the factors of interest;
- Reduction or eliminations of inconvenient factor effects;

Fundamental principles

The role of experimental design

Experimental design is useful for avoiding the influence of spurious factors and personal biases on the results, by performing experiments in a impartial and objective way.

> "*Never have too much love for your hypotheses.*"

> "*The great tragedy of Science - the slaying of a beautiful hypothesis by an ugly fact.*"
> – Thomas H. Huxley

Image: `http://www.iep.utm.edu/huxley/`

Example

Jacques Benveniste and the memory of water

- Nature (1988);
- Investigation committee: Maddox, Stewart, Randi;
- Retracted by Nature due to evidence of misconduct.

Methodological problems

- Experimenter bias (absence of proper blinding);
- Cherrypicking (selective recording of results);
- Unaccounted sampling errors;
- Possible contamination;
- Complete lack of prior physical/ chemical plausibility;
- **Non-reproducibility**.

Image: `http://www.jacques-benveniste.org/inmemoriam/photo1.html`

Structure of Experimental Design

Main points

To enable the use of a scientific approach in the design of an experiment, it is important to have on a solid understanding of:

- The field where the experiment is to be conducted;
- The strategy for data collection;
- The way the data should be analyzed (at least qualitatively).

Structure of Experimental Design

Guidelines for a good design

- Pre-experimental design:
  - Identification and definition of the problem;
  - Selection of experimental and response variables of interest;
  - Choice of experimental protocols;
- Choice of the experimental design;
- Collection of the data;
- Statistical data analyses;
- Conclusions and recommendations;

Pre-experimental design

Before we start

- Is the investigation relevant?
- Would the results be interesting for the research community?
- Practical relevance?
  - Employ exploratory experiments;
- Placement within the literature;
  - Avoid repetition and irrelevance.

"*Sometimes one should do a completely wild experiment,
like blowing the trumpet to the tulips every morning
for a month. Probably nothing would happen, but what if it did?*"

– Sir George Howard Darwin

Girl playing: `http://www.film.queensu.ca/tulips/default.html`

Pre-experimental design

Definition of hypotheses

- Exploratory experimentation $\neq$ careless experimentation;
- Data collection guided by questions of interest;
    - Difference on average/best/worst performance;
    - Robustness / reliability;
- The translation *scientific question* $\rightarrow$ *test hypothesis* requires special attention, and a solid knowledge of the technical area in which the experiment is being performed;

Choice of Experimental Design

Experimental design

- (Relatively) simple, as long as the pre-experimental part is well done;
- Dependent on what is being tested (statistical question);
- A sound design tends to determine the analyses technique to be used, at least qualitatively;
- Involves considerations about:
  - Sample size;
  - Ordering of observations;
  - Determination of restrictions to the randomization and the use of blocks, etc.
- Available in several statistical/mathematical packages;

Choice of Experimental Design

Problem-dependent

- Depending on the experimental question, different experimental designs are required
- A solid, statistically sound design tends to determine which statistical tests must be employed in the analysis step, at least qualitatively.
- Quantification of the proportion between intra-groups and inter-groups variability;

Actual Experiment
Data gathering

- Must be consistent with design, otherwise the validity of the results may be compromised - data collection must always follow the plan:
    - No premature stops;
    - *No-peeking rule*[a];
- Use of pilot experiments:
    - Gathering of preliminary information;
    - Practice with the experimental conditions;

[a] Except when planned, of course.

Analysis of the experimental data

A consequence of design

- Analysis techniques are generally relatively simple, but *the devil is in the details*;
- Use of existing statistical tools and frameworks, such as
- Free, versatile, good graphical capabilities, relatively simple (but with one hell of a learning curve);

Analysis of the experimental data

Statistical modeling

- General procedure for testing the experimental hypotheses:
  - Definition of a *null-model* (absence of effects) and of a desired level of significance;
  - Determination of $P(data|null\text{-}model)$;
  - Decision by rejection (or not) of the null hypothesis;
  - Validation of model assumptions;
  - Estimation of the *magnitude* of differences - **practical significance**;

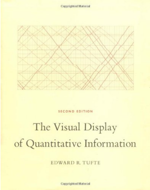
> *Statistical methods do not **prove** anything, but they allow an objective definition of margins of plausibility for certain statements.*

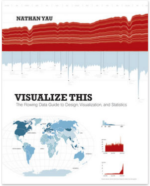
Reporting of results

Presentation

Combine textual, numeric and graphical elements to tell a story with your data. It simplifies the understanding and analysis of the results.

- Strive to achieve graphical excellence;
- Coherence of notation - special attention to figures and tables;
- Display simultaneous confidence intervals and other graphical indicators of effect size.





Other great resources on graphical excellence:

Flowing Data (`http://flowingdata.com/`)

Information is Beautiful (`http://www.informationisbeautiful.net`)

Conclusions

Drawing and reporting conclusions

- Conclusions should be based on solid evidence from the data;
- Be conservative - it is common to exaggerate the generality of the results;
- Report significance levels and the assumptions under which the results are valid;
- *Suggest explanations* to the observed results;
- Careful with *anomaly hunting*;

> *Always let the science drive the statistics. If you get a statistically significant result, go back and describe what it means in the scientific context.*
>
> – Aaron Rendahl

Discussion

Some more relevant points

- Use of previous knowledge, theoretical or empirical;
- Iterative experimentation;
- Statistical $\times$ practical significance;
- Use of additional experiments to validate conclusions.

Bibliography

Required reading

1. Dept. Biochemistry & Cell Biology, Rice University , *Common Errors in Student Research Papers.* - http://www.ruf.rice.edu/~bioslabs/tools/report/reporterror.html
2. T. Brady, *Reviewer's quick guide to common statistical errors in scientific papers.* http://goo.gl/96Qfk
3. R.M. Szydloa et al., *Sign of the Zodiac as a predictor of survival for recipients of an allogeneic stem cell transplant for chronic myeloid leukaemia (CML): an artificial association.* Transplantation Proceedings 42 (8):3312-3315, 2010.
4. D.C. Montgomery, *Design and Analysis of Experiments*, Chapter 1. 5th ed., Wiley, 2005

Recommended reading

1. S.C. Stearns, *Some Modest Advice for Graduate Students* - http://goo.gl/jtT4hA
2. A. Rendahl, *Experimental design and statistical analysis: questions to consider as you write your grant.* - http://goo.gl/yT0TK7
3. J. Maddox et al., *"High-dilution" experiments a delusion.* Nature 334, 287-290, 1988
4. V. Czitron, *One-Factor-at-a-Time Versus Designed Experiments.* The American Statistician, 53(2) 126-131, 1999.
5. B. Dunning, *An Enthusiast's Primer on Study Types.* Skeptoid Podcast, Skeptoid Media, Inc., 2013. - http://skeptoid.com/episodes/4381

About this material

Conditions of use and referencing

Please reference this work as:

Felipe Campelo (2015), *Lecture Notes on Design and Analysis of Experiments*. Online: https://github.com/fcampelo/Design-and-Analysis-of-Experiments
Version 2.11, Chapter 2; Creative Commons BY-NC-SA 4.0.

```
@Misc{Campelo2015-01,
  title={Lecture Notes on Design and Analysis of Experiments},
  author={Felipe Campelo},
  howPublished={\url{https://github.com/fcampelo/Design-and-Analysis-of-Experiments}},
  year={2015},
  note={Version 2.11, Chapter 2; Creative Commons BY-NC-SA 4.0.},
}
```