Applied Analysis of Variance and Experimental Design

401–0625–01L

Marianne Müller
mlm@ethz.ch

15. September 2014

Introduction

Introduction

1. Why is Experimental Design important?
2. Course organisation
3. Principles of experimental design

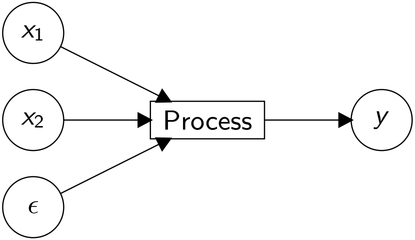
Women's health initiative (WHI)

Does hormone treatment improve women's health?

- Population: healthy, post-menopausal women in the U.S.
- Outcomes:
  - coronary heart disease
  - invasive breast cancer
  - dementia, osteoporosis
- Explanatory variables:
  - estrogen treatment, yes/no
  - demographic variables (age, race, etc.)
  - other health-related variables (diet, diseases)

Observational studies and experiments

- Observational study: data (explanatory variables and outcomes) are collected from a sample of the population.
- Experiment: conditions can be controlled, at least to some extent



Observational study starting in 1991

- Observational population: 93,676 women; tracked over eight years on average. Some women took hormones, others not.
- Results: Good health/low rates of CHD associated with estrogen treatment.
- Conclusion: Estrogen treatment positively affects health, such as CHD

Question: Is the conclusion justified?

RCT: Randomized controlled trial

- Experimental population: 373'092 women determined to be eligible, 18'845 provided consent to be in experiment, 16,608 included in the experiment.
- Women were randomized to either the experimental group (estrogen treatment) or the control group (no estrogen treatment).
- Women were grouped (blocked) together by age and clinic. Then within each age $\times$ clinic block, 50% of the women were randomly assigned to the experimental treatment, the remaining to the control treatment This type of random allocation is called a randomized block design.

Results of the RCT

- Women on treatment had lower incidence rates for hip fracture
- but higher incidence rates for CHD and breast cancer.
- Conclusion: Estrogen isn't a viable preventative measure for CHD in the general female population.

Question: Why lead the two studies to different conclusions?

Observational studies can only find associations. In an experiment we can make causal inferences. Nevertheless, observational studies are useful as well. Observational data might already be available, or there might be ethical constraints on experimentation.

Application of experimental design

- Agriculture and biology
- Medicine
- Engineering and industry
- Market research
- Psychology

Introduction

Main topics

- Principles of experimental design
- 1-Factor Anova
- Block designs
- Factorial designs
- Fractional factorials
- Split plot designs

Material

- Course outline, literature, organisational details, slides on
`stat.ethz.ch/education/semesters/as2014/anova`
- Recommended textbook: Montgomery, D.C (2013). *Design and Analysis of Experiments*, 8th edition, Wiley, New York.
- Lecture notes and any additional material on `ILIAS`, accessible via „myStudies“ and „course catalog“.
- Exercises, datasets, solutions see website

Exercises

- Assistant:
  Sylvain Robert, robert@stat.math.ethz.ch
- Introduction into R on 6/10/14 1-2 pm in HG F3 and 2-3 pm in a computer room to be announced.
- Afterwards every two weeks according to course outline in HG F3.
- External auditors who need an account send an email to an assistant.

Exam

- No confirmation required, keine Testatbedingung
- Session examination:
  - written exam
  - open book, simple pocket calculator
  - duration 120 minutes
  - 4 credits

Introduction

Planning of experiments

1. Statement of problem, empirically testable hypotheses
2. Collecting information (relevant background)
3. Choice of outcome, response variable
4. Determine sources of variation in response:
    - factors of interest
    - nuisance factors (blocking, randomisation)
    - factors to be held constant
5. Choice of design and randomisation scheme

Important principles of experimental design

- Replication: several observations under the same conditions to reduce variability
- Randomisation: Random allocation of treatments to experimental units to reduce systematic bias with regard to other influential variables, confounding variables
- Blocking: make blocks of similar units and compare treatments within blocks, to control nuisance factors and improve precision

RCT for heart disease patients

50 people get heart drugs only (group 1), another 50 people get diet instructions and heart drugs (group 2), response variable is the regularity of heart beat one month later.

Design 1: 50 women for group 1 and 50 men for group 2.

Design 2: 100 male patients, group 1 is treated in hospital 1, group 2 in hospital 2.

Design 3: 100 patients in hospital 1, the first 50 patients are treated with drugs only, the remaining 50 patients get drugs and diet instructions.

Question: Which design do you choose?

Example: Coronary Drug Project (1980)

	Medication	Placebo
5 Year Mortality	20%	21%

Medication group:

	Compliance	
	high	low
5 Year Mortality	15%	25%

Placebo group:

	Compliance	
	high	low
5 Year Mortality	15%	28%

Randomisation

Random allocation of units to experimental conditions to

- avoid systematic differences

  Example: Planting of varieties A and B

poor design:	A	A	A	A	A	B	B	B	B	B
better design:	A	A	B	A	B	A	B	B	B	A

- get valid t and F tests even if the normal distribution assumption is not satisfied.

Summary

Enough replication!

Block what you can and randomize what you cannot!