

ENG-467

Experimental design and data analysis with R

Raupach Timothy Hugh, Schlaepfer Rodolphe

Cursus	Sem.	Type
Civil & Environmental Engineering		Obl.
Environmental Sciences and Engineering	MA2, MA4	Opt.

Language of teaching	English
Credits	2
Session	Summer
Semester	Spring
Exam	Written
Workload	60h
Weeks	14
Hours	2 weekly
Courses	1 weekly
Exercises	1 weekly
Number of positions	

Summary

Linking together the elements of a research project. Basic principles of designing experiments and observational studies. Statistical model of Multiple regressions and Analysis of variance, as special cases of the general linear model, Data analysis with the statistical software R.

Content

1. Introduction (goal of the course, prerequisite, what is R)
2. An example: The Jura Gradient Experiment
3. An introduction to basic coding in R
4. Designing experiments and observational studies (Basic Principles, Power Analysis and Number of Replications, Some Types of Experimental Designs, Some Types of Sampling Designs for Observational Studies)
5. Statistical models (linear models, linear models with quantitative explanatory variables, linear models with categorical explanatory variables)
6. Principles of data analysis (Hypotheses to be tested, analysis of multiple regression, analysis of ANOVA, including Model Checking)
7. Analysing experiments (completely randomized experiment with one and two factors, complete randomized blocks, split-plot experiments)
8. Analysing observational studies (simple random sampling, systematic sampling, stratified sampling)
9. Special Issues (model assumptions not fulfilled, unbalanced designs, pseudo-repetitions, repeated measures, mixed effects, effect size, power of an experiment, contrasts, multivariate situations)

Keywords

Experimental design, sampling design, linear models, multiple regression, analysis of variance, data analysis, statistical software R.

Learning Prerequisites**Required courses**

Probability and Statistics, Prof. Victor Panaretos, Bachelor semester 2" or another course with a similar content (statistical distributions, expected value, error types one and two, parameter estimation, testing hypotheses, statistical significance, simple linear regression, one way analysis of variance)

Recommended courses

Ecologie numérique, ENV 521, Dr Vincent Jassey, Prof. Alexandre Buttler

Important concepts to start the course

Scientific method: from research questions to reporting, through data collection and data analysis

Learning Outcomes

- The participants can
- Interpret in a coherent way the main elements of the research process: "Research goal and questions", "Design of experiment and/or observational study", "Data collection", "Formulating the statistical model", "Hypothesis to be tested", "Elaborating the R Code", "Data analysis with R" and "Interpreting the results", "Reporting".
- Design simple experiments (purely randomized experiment (one and two factors), complete randomized block experiments and split-plot experiments) and simple observational studies (simple random -, systematic and stratified sampling)
- Use the concept of general linear models (GLM) to formulate statistical models for studying relationships between response variables and explanatory variables (quantitative and categorical)
- Implement the basic concept of data analysis for developing R codes for analysing multiple regressions and simple ANOVA models.
- special issues like "model assumptions not fulfilled", "repeated measures", "unbalanced designs", "pseudo-replications", "effect size" and "multivariate situations" and how to handle them.

Transversal skills

- Access and evaluate appropriate sources of information.

Teaching methods

Lectures
Exercises

Expected student activities

attendance at the lectures
completing exercises
reading written material (given documents, documents on the web)

Assessment methods

Written exam during the examination period

Supervision

Office hours	Yes
Assistants	Yes

Resources

Bibliography

- **Borcard, Daniel; Gillet, François; Legendre, Pierre. 2011. Numerical Ecology with R. Springer.**
- **Cochran William G. 1977. Sampling Techniques. Third Edition. Wiley. 474 pp.**
Note: This edition is freely available on internet. Cochran's book is one of the fundamental work on sampling.
- **Crawley Michael J. 2015. The R Book. Second Edition. Wiley. 1051 pp.**
Note: The first edition of this book is freely available on internet.
- **Crawley Michael J. Statistics. An introduction using R. Second Edition. 359 pp. Is a it-ebook (see www.it-ebooks.info). Free available on Internet**
- **Davison A. C. & Kuonen, D. (2013). Probabilités et Statistique pour Sciences de l'Environnement. Polycopié disponible à la "Librairie Polytechnique" de l'EPFL. (Edition 2013 modifiée par V.M. Panaretos).**
- **Lawson John. 2015. Design and Analysis of Experiments with R. CRC Press. 506 pp.**
- **Montgomery Douglas C. 2013. Design and Analysis of Experiments. Eighth Edition. Wiley. 730 pp.**

Note: Montgomery is one of the leading experts in Experimental Design. The fifth edition of his book is freely available on Internet.

- Quinn Gerry P., Keough Michael J. 2002. **Experimental Design and Data Analysis for Biologists**. Cambridge. 537 pp., is freely available on Internet

- Sutherland William J. 2006. **Ecological Census Techniques. A handbook. Second Edition**. Cambridge

Ressources en bibliothèque

- [Numerical Ecology with R](#)
- [Sampling Techniques](#)
- [The R book](#)
- [Design and Analysis of Experiments with R](#)
- [Design and Analysis of Experiments](#)
- [Experimental Design and Data Analysis for Biologists](#)
- [Ecological Census Techniques. A handbook](#)

Websites

- [http://many useful Websites will be given during the lectures](#)