



DEPARTMENT OF BIOLOGY
SUMMER SESSION 2024

BIO709/BIO809 –Advanced Statistics in Life Sciences

Educational Activity Plan

Training location: Jouvence Resort Center

Professor in charge	:	Guillaume Blanchet
Office	:	D8-1073
Phone	:	821-8000, extension 65130
Email	:	Guillaume.Blanchet@USherbrooke.ca
Instructors participating in the training	:	Andrew MacDonald

Course Schedule

This course is given in an intensive summer school format over five days.

Official Description of the Educational Activity¹

Training Targets

General Objective:

Develop the necessary statistical knowledge to construct statistical models adapted to address a specific problem.

Specific Objectives:

¹<https://www.usherbrooke.ca/admission/fiches-cours/BIO709/statistiques-avancees-en-sciences-de-la-vie/>

- Learn statistical theory to better construct, apply, and interpret various statistical models applied to life sciences.
- Become familiar with primary research in statistical modeling for life sciences.
- Gain experience working collaboratively on issues related to the development and application of statistical methods.

Content

Linear and nonlinear modeling, modeling of complex univariate and multivariate data in life sciences. Implementation of statistical models.

Number of Credits : 3

Workload Distribution : 3-0-6

Detailed Content

The course days will be divided into 90 to 120-minute blocks during which theoretical and practical concepts will be taught. Theoretical concepts will be delivered through lectures interspersed with discussions with students. To ensure the acquisition of theoretical skills, practical computer exercises will be proposed and discussed using the R and Stan programming languages.

Students are strongly encouraged to bring their data to apply the concepts learned in the course to their study systems. From Tuesday to Thursday, the last 30 minutes of the day (5:00-5:30 PM) will be reserved for student presentations of their research projects through short oral presentations.

Day 1 - Basic Concepts

- General introduction
- Probability law
- Frequentist vs Bayesian model
- Basic principles of data simulation
- Data simulation using different probability distributions (e.g., Poisson, binomial, gamma)

Day 2 - Linear Regressions

- Introduction to Stan
- Basic principles of linear algebra
- Linear models
- Generalized linear models

Day 3 - Simple Univariate Hierarchical Models

- Models with a hierarchical component on the intercept
- Models with a hierarchical component on a slope

Day 4 - Complex Hierarchical Model

- Multilevel hierarchical models
- Hierarchical models with multiple terms
- Hierarchical models with constraints
- Gaussian process

Day 5 - Advanced Hierarchical Models

- Phylogenetic hierarchical models
- Using hierarchical model in research

Required Material for the Educational Activity

For this course, a laptop is mandatory on which the R programming language and Stan must be installed.

Reference Books:

Andrew Gelman and Gennifer Hill (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Model*. Cambridge University Press.

Richard McElreath (2020) *Statistical Rethinking – A Bayesian Course with Examples in R and Stan – Second edition*. CRC Press.

Phil Spector (2008) *Data Manipulation with R*. Springer

Christopher K. Wikle, Andrew Zammit-Mangion and Noel Cressie (2019) *Spatio-Temporal Statistics with R*. CRC Press

Noël Cressie and Christopher K. Wikle (2011) *Statistics for Spatio-Temporal Data*. Wiley

Anthony R. Ives (2018) *Mixed and Phylogenetic Models: A Conceptual Introduction to Correlated Data* self published

Douglas Bates, Martin Mächler, Benjamin M. Bolker and Steven C. Walker (2015) Fitting Linear Mixed-Effects Models using `lme4`, *Journal of Statistical Software* 67:1–48.

N. Thompson Hobbs and Mevin Hooten (2015) *Bayesian Models: A Statistical Primer for Ecologists* Princeton University Press.

J. Albert (2009) *Bayesian Computation with R - Second edition* Springer.

Teaching Method

The sessions will consist of lectures and practical work interspersed with discussions.

Documents will be available on a web page dedicated to the course.

For this course, the use of the R, Stan, L^AT_EX and **Markdown** programming languages will be necessary for the various exercises to be completed during the course, to work on non-trivial real cases.

Evaluation

The course will be evaluated on a "Pass"/"Fail" basis.

The evaluation will be based on the work done during the summer school period and participation.

Use of Electronic Devices

The use of computers, cell phones, tablets, or any other similar device during the course is allowed, provided that its use is solely for learning purposes. This authorization may be withdrawn at any time by the instructor if the device is not used solely for learning purposes or if its use leads to abuse or disrupts the course or learning.

Request for Accommodations

All accommodation requests must be made through the student life services of the University of Sherbrooke (<https://www.usherbrooke.ca/bottin/services/services-a-la-vie-etudiante>)