

Introduction to maximum likelihood and Bayesian statistics for ecologists (2017)

Petr Keil, petr.keil@idiv.de, www.petrkeil.com

Summary

The course is an introduction to applied statistics that uses either maximum likelihood, or Bayesian inference. The course has two objectives. First, it will introduce in the critical concepts unifying all parametric statistical techniques; these are: likelihood, probability distributions, and the distinction between model, data, parameters, and predictions. Second, the course will show that one can do statistics outside of the classical categories such as regression or ANOVA, and that statistics can work as a modular kit in which several simple blocks can be used to analyze problems of any complexity. Participants will learn to specify and fit models using likelihood maximization in R, and using MCMC in OpenBUGS, JAGS, or STAN. The course emphasizes practical issues over philosophical ones. We will use mainly ecological examples.

What should you bring?

- Your own **laptop** with **R** and **JAGS** installed.
- **How to install JAGS?**
 - Windows and macOS users download the installers at: <http://goo.gl/tbw7Lt>
 - Ubuntu users: `sudo apt-get install jags`
- JAGS is not part of R, it's a stand-alone application, but we will command it from R. For this you will need to install a few extra R packages by typing (in R):
`install.packages(c("R2jags", "rjags", "coda"))`
- Basic familiarity with R is required: You should know how to set a working directory, how to import and export data, you should know how R indexing (the square brackets `[,]`), and you should be familiar with the following types of objects: `vector`, `data.frame`, and `list`. Knowledge of loops will be an advantage.

Is there an exam?

No, you will be credited for attendance and for performing a simple Bayesian analysis.

Day 1 – Introducing ML and Bayes

Wed 1/3/2017, Symbiosis, Interim II

9:30 - 12:00

- Course contents, pros and cons of ML and Bayes, necessary skills.
- Intro to likelihood: Probability density, probability mass, cumulative probability, likelihood, and deviance on Normal example.

14:00 - 16:00

- Poisson and Binomial distributions. Likelihood maximization.
- Posterior probability. The principle of MCMC sampling. Simplest model in JAGS.
- Basic literature.

Day 2 – Bayesian generalized and mixed-effects models

Thu 2/3/2017, Symbiosis, Interim II

9:30 - 12:00

- Generalized linear models - linear regression, logistic regression, Poisson regression,
- ANOVA etc., all that in BUGS.

14:00 - 16:00

- Hierarchical (mixed-effects) models, random effects vs. fixed effects, latent variables, informative vs. non-informative priors.

Day 3 – Custom models

Fri 3/3/2017, Symbiosis, Interim II

9:30 - 12:00

- Exercises from Days 1 and 2 – participants work on their own.
- Custom models 1: Time, space, phylogenies, survival, occupancy modelling, imperfect detection

13:00 - 15:00

- Custom models 2: Time, space, phylogenies, survival, occupancy modelling, imperfect detection
- Model selection, model evaluation, information theory criteria, handling uncertainty, Bayesian credible intervals, prediction intervals.

Literature

- Bolker (2008) *Ecological Models and Data in R*. Princeton Univ. Press.
- Gelman et al. (2004) *Bayesian Data Analysis*. Chapman & Hall.
- Kéry (2010) *Introduction to WinBUGS for Ecologists*. Academic Press.
- Kruschke (2014) *Doing Bayesian Data Analysis*. Academic Press.
- McCarthy (2007) *Bayesian Methods for Ecology*. Cambridge Univ. Press.
- Gelman & Hill (2007) *Data analysis using regression and multilevel/hierarchical models*. Cambridge Univ. Press.
- Plummer (2012) JAGS Version 3.3.0 user manual.