Exam of course **Bayesian Data Analysis II – Part 1**

February 3 2022

Prof. Emmanuel Lesaffre

I-Biostat, K.U.Leuven, Leuven

Included is the description of your exam project.

Practical arrangements:

1. Establish groups of size of **about 4-5** students!

2. Deliver your exam project on a **pdf** file. Each document should have a title page with heading: **2022 BDA II KU Leuven Part 1,** the names of the members of the group and their email addresses. Send the document to me **by email** (email address: **emmanuel.lesaffre@kuleuven.be**) at the latest on **Sunday, January 9, 2022, at 1 PM (Belgian time).**

3. Give your pdf file **the name of the first member of the group!**

4. **Send also your programs** to me in the same email, so that I can check that the programs really work! **Make sure that your program works, otherwise your score on the exam project is 0!**

5. Use `OpenBUGS`/`JAGS`/`R2OpenBUGS`/`R2jags`/`Nimble` and additional `R` software such as `CODA` or `BOA`, **but not SAS (PROC MCMC).**

6. For clarity, **first repeat the question** in your document and then give the solution.

7. Describe the flow of your procedures and the reasoning behind. Be clear! **PUT PAGE NUMBERS** on project! **PUT PAGE NUMBERS** on project! **PUT PAGE NUMBERS** on project! (understood?)

8. Annotate your output, but not everything that you have done needs to be put in the report. **Limit your report to 20 pages (excluding title page, but including tables and figures)!**

9. Note that this year the oral defense is done **ONLINE**. So be prepared and have a printed version of your exam project nearby.

10. **Study also the course material**; you will get general questions on Bayesian statistics at the oral defense.

11. **Good luck!**

Analysis: Modelling white grub worms' survival – biological control

1. Description of the data

White grub worms damage lawns by feeding on the roots and eventually turn into adult beetles that eat the leaves of other prized plants in one's garden. Many homeowners treat their lawn with a chemical grub control product, or grub killer, once or twice a year. However, several natural methods can effectively eliminate grub worms, such as entomopathogenic nematodes (EPNs). EPNs are a group of nematodes (threadworms) that cause insects' death and provide a safer alternative for the biological control of insect pests.

Nematodes emerging from a wax moth larva cadaver

In a gel-soil mixture, an experiment was conducted in cell culture plates to assess the survival of white grubs exposed to two types of EPN species, namely *S. sacchari* (SS) and *H. baujardi* (HB). White grub mortality was recorded over bi-daily intervals for 12 days. The EPNs were evaluated using seven grubs per plate and replicated ten times, thus 70 grubs per EPN species. The following table provides a summary of the experiment's dataset:

The data set **Grubs_Nematodes.csv** contains the following covariates:

VARIABLE	`GROUP`	`GRUBSIZE`	`UREPID`	`LOWERLIM`	`UPPERLIM`
DESCRIPTION	EPN species: 1 = S. sacchari (SS) 2 = H. baujardi (HB)	Grub size (cm)	Unique replicate plate ID (1-20). Each record denotes data of a single white grub in the experiment. Hence, 140 observations (white grubs) in total.	The lower bound of interval (days) during which the grub died.	The upper bound of interval (days) during which the grub died. A missing value denotes no limit. In particular, grubs that did not die during the experiment were right-censored at Day 12.

It is noted that the survival time of grubs is interval-censored as per variables `LOWERLIM` and `UPPERLIM`.

2. Research questions

The following research questions are of interest:

1. **Which of the grub worms' survival distributions, lognormal and Weibull distributions, best fits the data?**
2. **Which of the EPN species acts the fastest in killing the white grubs?**
3. **Does the grubs' size predict the survival of the worm?**

Tasks:

- First take the midpoint of the lower and upper limit of the interval-censored survival time and ignore the clustering. For the worms that survive, you take the mid-point of the lower limit and day 12. Then you:
    - look at the median death time per EPN species to assess which of the two EPN species acts the fastest in killing the white grubs
    - evaluate the effect of the covariate grubs' size
    - consider the clustering by including a random effect, i.e., the variation of location parameters across replicate plates. What does the random effect's variance component estimate suggest about the "clustering" of responses?
    - check whether there are outlying/influential observation(s). If so, give possible reasons for the outlying observation(s)
    - check the distribution of the random effect and propose a robust distribution for the random effect
- Secondly, you repeat the analyses considering the interval-censored character of the survival times. Are your conclusions still the same?

Note that for all analyses:

- Check convergence using classical diagnostics.
- Perform posterior predictive checks to evaluate the chosen model.
- Give interpretations for the parameter(s) of interest.

GOOD LUCK!