Bayesian Data Analysis – Assignment 8

General information

- The recommended tool in this course is R (with the IDE R-Studio). You can download R **here** and R-Studio **here**. There are tons of tutorials, videos and introductions to R and R-Studio online. You can find some initial hints from **RStudio Education pages**.
- Instead of installing R and RStudio on you own computer, see **how to use R and RStudio remotely**.
- When working with R, we recommend writing the report using R markdown and the provided **R markdown template**. The remplate includes the formatting instructions and how to include code and figures.
- Instead of R markdown, you can use other software to make the PDF report, but the the same instructions for formatting should be used. These instructions are available also in **the PDF produced from the R markdown template**.
- Report all results in a single, **anonymous** *.pdf -file and return it to **peergrade.io**.
- The course has its own R package `aaltobda` with data and functionality to simplify coding. To install the package just run the following (upgrade="never" skips question about updating other packages):
  1. `install.packages("remotes")`
  2. `remotes::install_github("avehtari/BDA_course_Aalto",`
     `subdir = "rpackage", upgrade="never")`
- Many of the exercises can be checked automatically using the R package `markmyassignment`. Information on how to install and use the package can be found **here**. There is no need to include `markmyassignment` results in the report.
- Recommended additional self study exercises for each chapter in BDA3 are listed in the course web page.
- Common questions and answers regarding installation and technical problems can be found in Frequently Asked Questions (FAQ).
- Deadlines for all assignments can be found on the course web page and in peergrade. You can set email alerts for trhe deadlines in peergrade settings.
- You are allowed to discuss assignments with your friends, but it is not allowed to copy solutions directly from other students or from internet. You can copy, e.g., plotting code from the course demos, but really try to solve the actual assignment problems with your own code and explanations. Do not share your answers publicly. Do not copy answers from the internet or from previous years. We compare the answers to the answers from previous years and to the answers from other students this year. All suspected plagiarism will be reported and investigated. See more about the **Aalto University Code of Academic Integrity and Handling Violations Thereof**.
- Do not submit empty PDFs or almost empty PDFs as these are just harming the other students as they can't do peergrading for the empty or almost empty

submissions. Violations of this rule will be reported and investigated in the same way was plagiarism.

- If you have any suggestions or improvements to the course material, please post in the course chat feedback channel, create an issue, or submit a pull request to the public repository!

Information on this assignment

This assignment is related to Chapter 7. The maximum amount of points from this assignment is 6.

Note! This assignment build upon assignment 7, so be sure that assignment 7 is correct before you start with this assignment. You can ask TAs for help to check that your assignment 7 answer is good.

Reading instructions: Chapter 7 in BDA3, see reading instructions **here**. Also read about the PSIS-LOO that in **this article** or **the corresponding preprint**.

Grading instructions: The grading will be done in peergrade. All grading questions and evaluations for assignment 8 can be found **here**

Reporting accuracy: For posterior statistics of interest, only report digits for which the Monte Carlo standard error (MCSE) is zero. *Example:* If you estimate $E(\mu) = 1.234$ with $\text{MCSE}(E(\mu)) = 0.01$, you should report $E(\mu) = 1.2$.

Installing and using `rstan`: See the Stan demos on how to use Stan from R. The university Ubuntu desktops have the necessary libraries installed so there should be no need to install anything. To install Stan on your laptop, see the instructions below.

In R, install package `rstan`. Installation instructions on Linux, Mac and Windows can be found at `https://github.com/stan-dev/rstan/wiki/RStan-Getting-Started`. Additional useful packages are `loo`, `bayesplot` and `shinystan` (but you don't need these in this assignment). For Python users, the `Arviz` library may be relevant.

Stan manual can be found at `https://mc-stan.org/users/documentation/`. From this website, you can also find a lot of other useful material about Stan.

Model assessment: LOO-CV for factory data with Stan (6p)

Use leave-one-out cross-validation (LOO-CV) to assess the predictive performance of the pooled, separate and hierarchical Gaussian models for the factory dataset (see the second exercise in Assignment 7). To read in the data, just use:

```
> library(aaltobda)
> data("factory")
```

PSIS-LOO is a recently developed method for approximating the exact LOO and is thus not in BDA3. For more information, see the lecture slides and the original paper **here** or **here**.

Use Stan for fitting the models, and the `loo` R package for computing the approximate LOO-CV given the posterior samples provided by Stan. You can install the package as

```
> install.packages("loo")
```

Python users can use PSIS-LOO implementation in ArviZ library.

The report should include the following parts.

1. Fit the models with Stan as instructed in Assignment 7. To use the `loo` or `psisloo` functions, you need to compute the log-likelihood values of each observation for every posterior draw (i.e. an S-by-N matrix, where S is the number of posterior draws and $N = 30$ is the total number of observations). This can be done in the `generated quantities` block in the Stan code; for a demonstration, see the Gaussian linear model `lin.stan` in the R Stan examples that can be found **here**.

2. Compute the PSIS-LOO elpd values and the $\hat{k}$-values for each of the three models.

   Hint! It will be convenient to visualize the $\hat{k}$-values for each model so that you can easily see how many of these values fall in the range $\hat{k} > 0.7$ to assess the reliability of the PSIS-LOO estimate for each model. You can read more about the theoretical guarantees for the accuracy of the estimate depending on $\hat{k}$ from the original article (see **here** or **here**), but regarding this assignment, it suffices to understand that if all the $\hat{k}$-values are $\hat{k} \lesssim 0.7$, the PSIS-LOO estimate can be considered to be reliable, otherwise there is a concern that it may be biased (too optimistic, overestimating the predictive accuracy of the model).

3. Compute the effective number of parameters p_{eff} for each of the three models.

   Hint! The estimated effective number of parameters in the model can be computed from equation (7.15) in the book, where $\text{elpd}_{\text{loo-cv}}$ is the PSIS-LOO value (sum of the LOO log densities) and lpd is given by equation (7.5) in the book.

4. Assess how reliable the PSIS-LOO estimates are for the three models based on the $\hat{k}$-values.

5. An assessment of whether there are differences between the models with regard to the $\text{elpd}_{\text{loo-cv}}$, and if so, which model should be selected according to PSIS-LOO.

6. Both the Stan and R code should be included in your report.