University of Iceland

School of Engineering and Sciences

Faculty of Physical Sciences

STAE529M Bayesian Data Analysis

Fall 2021

Outline of project - Parts I and II

Part I - **Assigned:** Friday October 22th 2021.

Due: Monday November 8th 2021.

Part II - **Assigned:** Friday October 22th 2021.

Due: Monday November 22th 2021.

Part I

Part I involves obtaining, viewing, exploring, error checking and describing the data. Based on this work the relevant scientific questions will be formed. Further, in this part, statistical models for the data are constructed. Also, relevant references are found and referred to in the report. Part I will be turned in as a report on the given due date and it should involve the following.

Project description

This is an introduction to the problem. Here the story behind the data and the scientific goals of the project are given. Formulate the scientific questions that stem from the scientific goals.

Description of the data

The data need to be described in some details with tables and/or graphs. The structure of the data and what is being measured should also be described here. The source of the data needs to be specified.

Statistical models

Here the statistical model/models for the data and any unobserved quantities are given. All distributions are described, that is, both sampling distributions, prior distributions, and distributions for unobserved processes or unobserved variables.

A list with all parameters and variables in the project

Create a table with all variable names (observations, parameters, unobserved variables) along with a short description of each variable, e.g., μ_{jt}: the mean temperature at station j for year t.

References

A list of references found at this point is included in the report. At least one reference to a

paper on statistical models similar to the one used here should be given.

Part II

The second part of the project involves answering the questions asked in Part I. It also involves a description of the Bayesian analysis and computer programs, and writing down results followed by a discussion and conclusions. The report for Part II is the final report. It includes Part I along with the sections described below.

Bayesian estimation

The estimation of parameters and other quantities with the Bayesian approach is given here. The posterior distribution is derived. The MCMC algorithm is described. If the Gibbs sampler is used then the conditional distributions are given. If a Metropolis algorithm or a Metropolis–Hastings algorithm are used then these need to be described.

Computer programs

The computer programs that are written should be included in the appendix.

Results

Report estimates of parameters and other quantities in tables and figures and describe the results.

Discussion and Conclusions

A discussion on the project. This could involve answering the following questions; What scientific questions were answered? What conclusions can be drawn from the analysis of the data? How well did the statistical model/models describe the data. What steps can be taken in the future to further improve the analysis and the statistical modeling?

References

Collect all references that were used in the final report and put them in a section after the Discussion and Conclusion section and before the appendix.

Appendix

The code written for the project is placed in the appendix.

The final report

Below is an outline of the final report.

Introduction

1. the story
2. the data
3. the scientific goals / scientific questions

Data

1. description of the data
2. tables
3. graphs

Statistical models

1. data distributions
2. prior distributions
3. list of variables

Bayesian inference

1. the posterior distribution
2. MCMC algorithm (Gibbs sampler / conditional distributions / Metropolis algorithms / Metropolis–Hastings algorithms)

Results

1. the results in the form of tables and graphs
2. description of the results
3. evaluation of the proposed model

Discussion and Conclusions

1. answers to scientific questions
2. conclusions based on the data
3. how well did the proposed model describe the data
4. future steps

References

1. a journal like list of references

Appendix

1. the code for the project

The presentation

The presentation should be 10 minutes long and it should consist of about 10 slides. The slides should include;

1. introduction
2. description of the data
3. the statistical model/models
4. the Bayesian inference
5. results
6. discussion and conclusions