Statistique Bayésienne: projet

Anna Simoni

November 9, 2022

To validate the course of Statistique Bayésienne you have to return two homeworks (Devoir Maisons) and this project. The final grade will be obtained from the grades of the two homeworks and the project.

Here are the instructions for the project:

(1) The project must be done **by groups of two students**. (2) The project length has to be about 5 pages (plus code), with reasonable font size and margins. (3) It is an applied project. Each group has to choose a statistical application with real or simulated data and then analyse it with Bayesian methods related to the material seen in class (or extensions of it). You are free to choose to work either with a real data set or with simulated data. You are free to choose a statistical problem that you are interested in or that you have already analysed with frequentist procedures (for instance, in another class or during an internship). Alternatively, you can consider a statistical problem related to an academic article (below I propose a partial list or you can propose an article). (4) **Before starting to work on a project it must be approved by me. So, please email me (anna.simoni@ensae.fr): the names of the participants of the group and a short description of the project so that I can validate it.** The project has to be chosen and validated by me before November 30, 2022 so that you will have enough time to work on the project. (5) You can use any language among Python, R or Matlab for the programming. (6) The due date of the project is January 15, 2023. By the due date please upload

- your report as a pdf named `SurnameStudent1_SurnameStudent2`,
- a zipped folder `SurnameStudent1_SurnameStudent2` containing your code and a detailed readme file with instructions to run the code

on `https://app.compilatio.net/v5/document-submission/Y6P-76C-46D`

A possible structure of the project is: (1) definition of the problem under study and explanation of why it is interesting, (2) choice of the appropriate Bayesian technique (you should explain the methodology used and the motivation why you have chosen it), (3) description of the computational method used and difficulties encountered, (4) explanation and interpretation of the results. You can also compare the results you find with the results that you find with a frequentist approach.
If you choose to develop a project based on an academic article you can, for instance, either replicate and extend the simulations in the paper, or find a real data set and apply the method of the paper to it, or develop a simulation study if the paper does not contain it. An article can be chosen only by one group on the first-in-first-out basis. So, if you want to choose an article among the ones proposed below please email me the titles of three articles in order of preference and I will assign you the first article in your list that is still available.

Proposed academic articles

Theory:

1. Ma, Y. and J.S. Liu (2022). On Posterior Consistency of Bayesian Factor Models in High Dimensions, *Bayesian Analysis*, Vol. 17, 901 – 929.

2. Hamura, Y., Irie, K., and S. Sugasawa (2022). On Global-Local Shrinkage Priors for Count Data, *Baysian Analysis*, Vol. 17, 545 – 564.

3. Bontemps, D. (2011). Bernstein-von Mises theorems for Gaussian regression with increasing number of regressors, *Annals of Statistics*, 39, 2557 – 2584.

4. Gustafson, P. and L.Wasserman, (1995). Local Sensitivity Diagnostics for Bayesian Inference, *The Annals of Statistics*, 23, 2153 – 2167.

5. Ma, Y. and J. S. Liu (2022). On Posterior Consistency of Bayesian Factor Models in High Dimensions, *Bayesian Analysis*, 17, 901 – 929.

Computational:

6. Chan, J. C. C., Jacobi, L. and D. Zhu (2022). An automated prior robustness analysis in Bayesian model comparison, *Journal of Applied Econometrics*, Vol. 37, 583 – 602.

7. Benoit, D. F., Van Aelst, S., and D. Van den Poel (2016). Outlier-robust Bayesian Multinomial Choice Modeling. *Journal of Applied Econometrics*, 31, 1445 – 1466.

8. Ghosh, J., Li, Y. and R. Mitra (2018). On the Use of Cauchy Prior Distributions for Bayesian Logistic Regression. *Bayesian Analysis*, 13, 359 – 383.

9. McCulloch, R. and P.E. Rossi, (1994). An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics*.

10. Ando, T. and Zellner, A. (2010). Hierarchical Bayesian Analysis of the Seemingly Unrelated Regression and Simultaneous Equations Models Using a Combination of Direct Monte Carlo and Importance Sampling Techniques. *Bayesian Analysis*, 5, 65 – 96.

Partial Identification:

11. van Hasselt, M., Bollinger, C. R. and J. W. Bray (2021). A Bayesian approach to account for misclassification in prevalence and trend estimation, 37, 351 – 367.

12. Gustafson, P., (2012). On the behaviour of Bayesian credible intervals in partially identified models, *Electronic Journal of Statistics*, 6, 2107 – 2124.

Clustering:

13. Pamminger, C. and S. Frühwirth-Schnatter, (2010). Model-based Clustering of Categorical Time Series, *Bayesian Analysis*, 5, 345 – 368.

14. Frühwirth-Schnatter, S. and S. Kaufmann, (2008). Model-based clustering of multiple time-series, *Journal of Business and Economic Statistics*, 26, 78 – 89.

15. Quintana, F.A. (2006). A predictive view of Bayesian clustering, *Journal of Statistical planning and inference*, 136, 2407 – 2429.

16. Kim, S., Tadesse, M.G., and M. Vannucci, (2006). Variable selection in clustering via Dirichlet process mixture models, *Biometrika*, 93, 877 – 893.

17. Kaufmann, S. (2010). Dating and forecasting turning points by Bayesian clustering with dynamic structure: a suggestion with an application to Austrian data, *Journal of Applied Econometrics*, 25, 309 – 344.

Factor Models:

18. Bolfarine, H., Carvalho, C. M., Lopes, H. F. and J. S. Murray (2021). Decoupling Shrinkage and Selection in Gaussian Linear Factor Analysis, ArXiv.2006.11908.

Finance and Macroeconometrics:

19. Giannone, D., Lenza, M. and G. Primiceri (2019). Priors for the long run, *Journal of the American Statistical Association*, Vol. 114, 565 – 580.

20. Jacquier, E., Polson, N. G. and P. E. Rossi (1994). Bayesian analysis of Stochastic Volatility Models, *Journal of Business and Economic Statistics*, Vol. 12, 371 – 389.

21. Jacquier, E., Polson, N. G. and P. E. Rossi (2004). Bayesian analysis of Stochastic Volatility Models with fat-tails and correlated errors, *Journal of Econometrics*, Vol. 122, 185 – 212.

22. Ghysels, E., McCulloch, R., E., and R. S. Tsay (1998). Bayesian Inference for Periodic Regime-switching Models. *Journal of Applied Econometrics*, 13, 129 – 143.

23. Loaiza-Maya, R., Martin, G. M., and D. T., Frazier (2021). Focused Bayesian prediction. *Journal of Applied Econometrics*, 36, 517 – 543.

Machine Learning and High Dimension:

24. Linero, A. R. (2018). Bayesian Regression Trees for High-Dimensional Prediction and Variable Selection. *Journal of the American Statistical Association*, Vol. 113, 626 – 636.

25. Filipini dos Santos, P.H. and H. F. Lopes (2018). Tree-Based Bayesian Treatment Effect Analysis. ArXiv 1808.09507.

26. Li, Q., Xi, R. and N., Lin (2010). Bayesian Regularized Quantile Regression. *Bayesian Analysis*, 5, 533 – 556.

27. Kyung, M., Gill, J., Gosh (2010). Penalized Regression, Standard Errors, and Bayesian Lassos. *Bayesian Analysis*, 5, 369 – 412.

28. Li, Q. and N. Lin (2010). The Bayesian Elastic Net. *Bayesian Analysis*, 5, 151 – 170.

Other:

29. Lopes, H. F. and N.G. Polson (2014). Bayesian instrumental variables: priors and likelihoods, *Econometric Review*, Vol. 33, 100 – 121.

30. Hahn, R., He, J. and H.F. Lopes (2018). Bayesian factor model shrinkage for linear IV regression with many instruments, *Journal of Business and Economic Statistics*, Vol. 36, 278 – 287.

31. Kim, J. and L. Wang (2019). Hidden group patterns in democracy developments: Bayesian inference for grouped heterogeneity. *Journal of Applied Econometrics*, Vol. 34, 1016 – 1025.

32. Chib, S. and L. Jacobi (2016). Bayesian Fuzzy Regression Discontinuity Analysis and Returns to Compulsory Schooling. *Journal of Applied Econometrics*, Vol. 31, 1026 – 1047.

33. Yin, G. (2009). Bayesian Generalized Method of Moments. *Bayesian Analysis*, 4, 191 – 208.