A Mini-conference on Bayesian Methods at Lund University
10th of February, 2015
Room EC3:108, Ekonomicentrum, Tycho Brahes väg 1, Lund.
www.lucs.lu.se/bayes-at-lund-2015/

Program

09.15–11.00 Pre-conference hands on session Introduction to Bayesian Hierarchical Modelling. (also in room *EC3:108*)

11.15–12.30 Welcome & Session 1: Foundations, Performance and Pollen

▷ *Bayes from a frequentist point of view*, Krzysztof Podgorski, Departement of Statistics, Lund University

▷ *Performance of Bayesian prediction of treatment differences using a two-factor linear mixed-effects model*, Johannes Forkman, SLU

▷ *Lindley's paradox*, Bengt Ringnér, Mathematical Statistics, Lund University

▷ *Pollen based spatial reconstruction of past land cover: Estimating latent GMRFs with Dirichlet observations*, Behnaz Pirzamanbin et al., Mathematical Statistics, Lund University

12.30–13.15 Sandwich Lunch in the lobby (free, but registration is required)

13.15–14.15 Keynote speaker: Mattias Villani, Linköping University - *Bayesian model inference – why, what and how*?

14.30–15.45 Session 2: Computation, Cells and Socks

▷ *Tiny data, approximate Bayesian Computation and the socks of Karl Broman*, Rasmus Bååth, Lund University Cognitive Science

▷ *Data-Cloning ABC for (approximate) maximum likelihood estimation*, Umberto Picchini, Mathematical Statistics, Lund University

▷ *Joint cell population identification through Bayesian hierarchical modeling*, Kerstin Johnsson, Centre for Mathematical Sciences, Lund University

▷ *Distributing a collapsed sampler for topic models*, Måns Magnusson et al., Linköping University

15.45–16.15 Coffee and Cake

16.15–17.30 Session 3: Belief, Money and the Moose Population

▷ *Estimation of local moose population using Bayesian hierarchical modelling*, Jonas Wallin, Matematiska vetenskaper, Chalmers

▷ *Teaching Bayesian data analysis in psychology*, Geoffrey R. Patching, Department of Psychology, Lund University

▷ *Bayesian estimation of optimal portfolio*, Stepan Mazur, Department of Statistics, Lund University

▷ *How Bayesian belief networks can help save the world*, Ullrika Sahlin, Centre for Environmental and Climate Research, Lund University

Keynote Presentation: Bayesian model inference - why, what and how?

Mattias Villani, Division of Statistics and Machine Learning, Linköping University, mattias.villani@gmail.com

Practical Bayesian work is based on probability models for the data. In most real-world problems there is uncertainty about the correct model to use, and more than one model is typically under consideration. I will introduce the basics of Bayesian model inference, including model comparison and model averaging. Methods for implementing Bayesian model inference will be briefly discussed. I will also raise concerns about Bayesian model inference, and criticise the way this theory is used in a large portion of practical applications.

Session 1: Foundations, Performance and Pollen

Bayes from a frequentist point of view

Krzysztof Podgorski, Statistics, Lund University, krzysztof.podgorski@stat.lu.se

After discussing the basis for the 'unreconciled' differences between the so-called 'frequentists' and 'Bayesians', we argue that a frequentist statistician can benefit from the Bayesian methodology. The prior does not need to be taken as a subjective opinion but rather as a 'not-so-good' estimate of the truth. Then the posterior information is a way of correcting it using the data. A simple example of the frequentist Bayesian method is presented. It is shown to perform very well as measured by the frequentist criteria.

Performance of Bayesian prediction of treatment differences using a two-factor linear mixed-effects model

Johannes Forkman, SLU, Johannes.Forkman@slu.se

Randomized complete block experiments are traditionally studied using analysis of variance, which corresponds to fitting a linear model with fixed effects of blocks and treatments. However, we consider a linear model with fixed effects of blocks and random effects of treatments. This model provides shrinkage estimates of treatment differences, which is advantageous in terms of mean-square error. Through simulation, Bayesian prediction of treatment differences was compared with best linear unbiased prediction. Bayesian prediction gave slightly smaller root-mean-square error in predictions and more accurate coverage of intervals than best linear unbiased prediction.

Lindley's paradox

Bengt Ringnér, Mathematical Statistics, Lund University,
Bengt.Ringner@matstat.lu.se

In 1957 Lindley came up with a situation in which testing whether or not a certain effect is present would give the following results:

• Frequentists say: I am seldom wrong, and I say that there is an effect, although, maybe, too small to have a realistic explanation.

• Bayesians say: Given the observations, absence of effect is way more probable than presence of a realistic one.

I cannot see any contradiction here. Instead my point is:

• When testing $\theta = 0$ or $\theta \approx 0$ the choice of prior distribution is crucial, even for a large sample size.

Explicit examples will be given.

Pollen based spatial reconstruction of past land cover: Estimating latent GMRFs with Dirichlet observations Behnaz Pirzamanbin et al.

Behnaz Pirzamanbin et al., Mathematical Statistics, Lund University,
behnaz@maths.lth.se

When assessing past earth surface-climate feedbacks and past human impact on climate, an accurate estimates of past land cover are required. Fossil pollen records extracted from lakes and bogs provide good insight into the local land cover around each site. For use in climate modelling these sparse observations have to be interpolated; creating continuous maps of past land cover at regional and sub-continental scales. We construct a hierarchical model based on latent Gaussian Markov Random Fields with Dirichlet observations. The model is used to reconstruct past land cover across Europe for five time periods - centred around 1950, 1800, 1300 CE and 1000,

4000 BCE, by combining estimates of past human land use and output from a dynamic vegetation model with pollen based local land-cover estimates. To estimate the model a block updated MCMC, which includes an adaptive Metropolis adjusted Langevin step, is used. Model results are evaluated by comparing results for the 1950-time period to a European forest map from 2006. The lack of good historic land-cover data makes the evaluation challenging.

Session 2: Computation, Cells and Socks

Tiny data, approximate Bayesian computation and the Socks of Karl Broman

Rasmus Bååth, Lund University Cognitive Science, rasmus.baath@lucs.lu.se

Big data is all the rage, but sometimes you don't have big data. Sometimes you don't even have average size data. Sometimes you only have eleven unique socks. This is the story about how a tweet by esteemed biostatistician Karl Broman got me thinking about socks for over a week. During the ride we will learn about Approximate Bayesian Computation, which is much easier than it sounds, and which is perhaps the *most* general method of fitting statistical models.

Data-Cloning ABC for (approximate) maximum likelihood estimation

Umberto Picchini, Mathematical Statistics, Lund University, umberto@maths.lth.se

In the last decade approximate Bayesian computational methods (ABC) have enabled the use of Bayesian inference for previously intractable models. Essentially ABC circumvents the requirement of a having a readily available likelihood function (or a computationally expensive approximation thereof) by repeatedly simulating samples x from the data-generating model. When samples x are approximately matching the actual measurements we consider the conditioning parameter \theta as generated from the approximated posterior distribution on \theta. However the threshold introduced in ABC algorithms, which regulates the quality of closeness of simulated data to measurements, in practice cannot be set to be arbitrarily small as we otherwise incur into high rejection rates.

We propose to compensate for the inability to decrease the threshold "as much as wanted" with a procedure that samples from an increasingly peaked version of the approximate posterior. Sampling from the peaked posterior

can be achieved using so-called MCMC "data cloning" (e.g. Lele et al. 2007) which returns maximum likelihood estimates and the Fisher information. In this preliminary study we enable data-cloning for an ABC-MCMC sampler: the procedure returns approximate maximum likelihood estimators, useful for complex models for when a typical ABC approximation is able to locate the correct maxima in the posterior but not a sufficiently good approximation for the whole posterior surface. A simple linear regression example will be considered.

Joint cell population identification through Bayesian hierarchical modeling

Kerstin Johnsson, Centre for Mathematical Sciences, johnsson@maths.lth.se

Flow cytometry is a widely used technology in medical research as well as in clinical applications; among other things it can be used to study the immune system and to diagnose blood cancers. An essential part of flow cytometry data analysis is to partition measurements on cells into cell populations. This is hard since technical and biological variation gives the populations different characteristics in different samples. We have developed a Bayesian hierarchical model for joint cell population identification in an entire batch of flow cytometry samples, enabling systematization of this variation and thus facilitating comparisons.

Distributing a collapsed sampler for topic models

Måns Magnusson, Leif Jonsson, David Broman, Mattias Villani, Linköping University, mons.magnusson@gmail.com

Latent dirichlet allocation is a model widely used in the probabilistic modeling of text. The purpose is to use a latent representation of topics to reduce the dimensionality of documents and to derive latent topics in documents, where each latent topic can be seen as a dirichlet distribution over words. When MCMC methods are used to make inference regarding the topics of corpuses a collapsed gibbs sampler that samples a topic indicator for each word in the corpus has been popular due to its simplicity and efficiency. The main problem with this sampler is that it is inherently serial in nature, making parallelization di cult and hence limiting the possible scale. At the same time parallelization is necessary for larger corpuses with millions of documents. The AD-LDA sampler in [2] is an attempt to parallelize the collapsed sampler by approximating it with a parallel version without convergence guarantees. We suggest a parallel sampler with convergence guarantees with

almost the same efficiency as the original collapsed sampler but substantially faster computing times.

Session 3: Belief, Money and the Moose Population

Estimation of local moose population using Bayesian hierarchical modelling

Jonas Wallin, Chalmers, Matematiska vetenskaper, jonwal@chalmers.se

Modelling, large animal, populations is a complex problem for several reasons, the main issue is often lack of unbiased data. Typically, the only data is indexes, i.e observers reporting number of seen animals. In literature, the relation between the indexes and the population size is assumed fixed over time. This allows for studying the dynamics of the population using indexes. We define a Bayesian hierarchical model with both unbiased data and indexes. The model shows that it is questionable that indexes, solely, can be used for studying dynamics. Thus showing the fundamental need of unbiased observations.

Teaching Bayesian data analysis in psychology

Geoffrey R. Patching, Department of Psychology, Lund University, Geoffrey.Patching@psy.lu.se

Bayesian methods of data analysis are increasingly being adopted in psychology and across the social sciences. Yet, attempts to teach Bayesian data analysis to the psychology students in Lund has met with many challenges. In this talk, I shall lay out the current progression of statistics teaching from first year undergraduate to master level. Recent attempts to introduce Bayesian methods of data analysis include a day lab, increased emphasis on the software program R, and proposed changes to the course literature. The aim is to provide the fundamental tools for the students to progress, and open discussion about the feasibility of teaching Bayesian data analysis to the students.

Bayesian estimation of optimal portfolio

Stepan Mazur, Department of Statistics, Lund University, mazurstepanm@gmail.com

We consider the estimation of the weights of optimal portfolios from the Bayesian point of view under the assumption that the conditional distribution of the log-returns is normal. The imposed distributional assump-

tion leads to the unconditional non-normal distribution as well as to time-depended asset returns what is usually documented in real data. Using the standard priors for the mean vector and the covariance matrix, we derive the posterior distributions for the weights of the global minimum variance portfolio. Moreover, we reparametrize the model to allow informative and non-informative priors directly for the weights of the global minimum variance portfolio. For almost all models the posterior distributions for the portfolio weights are derived in explicit form. The models are compared by using the coverage probabilities for credible intervals. In an empirical study we analyze the posterior densities of the weights of an international portfolio.

How Bayesian belief networks can help save the world

Ullrika Sahlin, Centre for Environmental and Climate Research, ullrika.sahlin@cec.lu.se

Efficient decision making relies on the success in assessing the consequences of decisions on the systems of concern. Bayesian Belief Networks (BBN) is a widely used tool in risk and decision analysis. Since BBNs are used to evaluate uncertainty in casual relations, and well quantified uncertainty in important casual relations open up for improved science-based decision making, it is possible to say that using BBNs creates relatively better opportunities to save the world compared to not doing it. Now, it is timely to ask what is and what is not a BBN. I will take you through a range of possible answers to this question. Hopefully you can leave this conference knowing if, and in what ways, you are using Bayesian Belief Networks.

Acknowledgement