ABC In Sydney

Date: 3rd-4th July, 2014 University of New South Wales

Location: Room 149, Old Main Building, UNSW. (Enter the Red Centre building through the glass doors marked School of Mathematics and Statistics. Take the stairwell for 1.5 floors and then cross through the corridor to the Old Main Building. Continue to the end of the corridor, past the stairwell, and turn left. Room 149 will then be on your right.)

Approximate Bayesian Computation (ABC) refers to a class of statistical methods that allow the implementation of an approximate Bayesian analysis when evaluation of the likelihood function is computationally prohibitive or just unavailable. In the past ten years, ABC has moved from a niche idea in population genetics to a mainstream statistical procedure.

This workshop is held as a satellite to the joint IMS-ASC international conference, also held in Sydney, on 7-10 July, 2014. It consists of invited talks on the latest innovations and applications of ABC and related methods. It will also commence with a (crowd-sourced!) short course introducing ABC methods to those new to the area.

As such, the workshop is targeted to all students, researchers and academics with an interest in learning more about ABC, as well as those researchers who are working directly in the area.

Sponsors: We gratefully acknowledge the support of the *Statistical Society of Australia*, the *ARC Centre for Mathematical and Statistical Frontiers* and the *Faculty of Science, UNSW* in providing a free to attend workshop.

Thur 3rd July		
1:20-1:30	Welcome	Scott Sisson (UNSW)
1:30-2:00	Introduction to ABC	Matthew Moores (QUT)
2:00-2:30	Regression/Marginal adjustments	Scott Sisson (UNSW)
2:30-3:00	How to choose summary statistics	Dennis Prangle (Reading)
3:00-3:30	**Afternoon tea**	
3:30-4:00	MCMC and SMC algorithms	Chris Drovandi (QUT)
4:00-4:30	Model validation and choice	Scott Sisson (UNSW)
4:30-5:00	An application in epidemiology	Zach Aandahl (UNSW)
5:00	Close	

Fri 4th July

9:00-9:10	Welcome	Scott Sisson (UNSW)
9:10-10:00	Lazy ABC	Dennis Prangle (Reading)
10:00-10:50	Bayesian indirect inference using a parametric auxiliary model	Chris Drovandi (QUT)
10:50-11:20	**Morning Tea**	
11:20-11:55	Quantifying uncertainty in parameter estimates of stochastic cell spreading models using ABC	Brenda Nho Vo (QUT)
11:55-12:45	An introduction to particle approximate Bayesian computation methodologies	Pierre Del Moral (UNSW)
12:45-2:15	**Lunch**	
2:15-2:50	Functional regression ABC for Gaussian process density estimation	Guilherme Rodrigues (UNSW)
2:50-3:40	Approximate Bayesian computation in state space models	Gael Martin (Monash)
3:40-4:10	**Afternoon Tea**	
4:10-5:00	Pre-processing for ABC-SMC with undirected graphical models	Matthew Moores (QUT)
5:00	**Close**	Scott Sisson (UNSW)

Abstracts

Dennis Prangle, University of Reading, UK

Lazy ABC

ABC involves producing a large number of model simulations, which demands considerable computing time. Simulated datasets are then accepted only if they are sufficiently close to the observations. However, it is often clear before a simulation ends that it is unpromising: it is likely to produce a poor match or require excessive time. This talk is on "lazy ABC", an ABC algorithm which saves time by sometimes abandoning such simulations. This is accomplished in such a way as to leave the target distribution unchanged from that of standard ABC i.e. no further approximation is introduced.

Chris Drovandi, Queensland University of Technology

Bayesian Indirect Inference using a parametric auxiliary model

In this talk I will present a rather general framework for likelihood-free Bayesian inference problems called Bayesian Indirect Likelihood (BIL). I will then focus on some specific instances of the BIL framework that use in some way a parametric auxiliary model, which is an alternative model that possesses a tractable likelihood function. These methods are referred to as parametric Bayesian Indirect Inference (pBII) methods. One class of pBII methods uses the score or parameter of the auxiliary model to form summary statistics for ABC. A different class (called parametric BIL, pBIL) uses the likelihood (either at the full-data or summary statistic level) of the auxiliary model as a replacement to the likelihood of the true model. The theoretical aspects of pBIL will be explored and contrasted against the ABC approach that uses summary statistics formed from the auxiliary model. The theoretical results and the performance of the methods will be demonstrated on examples of varying complexity, including estimating the parameters of a stochastic process for macroparasite population evolution. This is joint work with Prof. Tony Pettitt, Dr Anthony Lee and Leah South.

Brenda Nho Vo, Queensland University of Technology

Quantifying uncertainty in parameter estimates of stochastic cell spreading models using ABC

Cell spreading processes involve several mechanisms such as cell motility, cell-to-cell adhesion and cell proliferation and are essential for dermal wound healing, cancerous invasion and the development of an organism from an embryo. Therefore it is crucial to gain an understanding of how such processes evolve. Mathematical models are often used to interpret biological experimental data and to formulate the intrinsic parameters that can be used for prediction. Many approaches in the literature assume deterministic models and only produce point estimates of model parameters. In contrast here, Approximate Bayesian Computation (ABC) has been used to provide a general and principled technique to infer the values of the cell diffusivity, D, and the cell proliferation rate, λ, from a discrete stochastic model of cell spreading and to quantify the uncertainty associated with the parameter estimates, based on experimental data of collective cell spreading of mouse fibroblast cells (3T3 cells) in a circular barrier assay. This ABC analysis was independently conducted using only the area of cell spreading populations enclosed by the leading edge for different combinations of initial cell densities and experimental times in two separate scenarios: cell spreading driven by cell motility only, and cell spreading driven by cell motility and cell proliferation combined.

This study found that D is estimated more precisely for the experiments with larger initial cell populations. D was also found to depend highly on the level of initial cell density and the experimental time, highlighting the importance of considering the level of cell densities and the experimental times when estimating the parameters from collective cell spreading models. While the experiments were performed over relatively short time scales, information from the leading edge alone provides minimal information for estimating λ, highlighting the need to investigate optimal experimental designs within the cell biology context so that parameters can be precisely estimated from the collected data. This is joint work with Chris Drovandi, Tony Pettitt and Matthew Simpson.

Pierre Del Moral, University of New South Wales

An introduction to particle approximate Bayesian computation methodologies

This lecture is concerned with Approximate Bayesian computation methodologies and their particle interpretations. We provide a rather detailed discussion on the theoretical foundations of these methods. We also discuss their application domains, including advanced signal processing, information theory, and computational physics.

Guilherme Rodrigues, University of New South Wales, Sydney

Functional regression ABC for Gaussian process density estimation

In several statistical applications, interest lies on modelling interrelated (similar, but not equal) density functions, each of which related to a different group or subpopulation. The major challenge in this context is to fully explore the data by optimally borrowing strength through the existing groups. In a Bayesian non-parametric approach, we introduce a hierarchically structured density prior, built over a transformation of Gaussian processes. Inference for the infinite-dimensional parameters is performed through a combination of Approximate Bayesian Computation (ABC) and functional regression adjustment. The regressors are defined in terms of the summary statistic, taken here to be the well-known Kernel estimator (computed for each group). By avoiding MCMC methods, the proposed technique provides approximate posterior samples at a considerable lower computational cost. To conclude, we illustrate some interesting properties and give further insight of the method through an instructive example. This is joint work with David Nott and Scott Sisson.

Gael M. Martin, Monash University,

Approximate Bayesian computation in state space models

Exploiting the likelihood-free techniques of approximate Bayesian computation (ABC), a new approach to inference in state space models is proposed. ABC avoids direct evaluation of the likelihood function by matching summary statistics calculated from the observed data with corresponding statistics computed from data simulated from the true process. Exact inference is feasible only if the matching statistics are
sufficient for the unknown parameters. We demonstrate that finite sample sufficiency is not attainable in the state space setting. Hence, we are motivated to seek asymptotic sufficiency via the maximum likelihood estimator (MLE) of the parameters of an auxiliary, or approximating, model. We prove that this auxiliary model-based approach achieves Bayesian consistency, as well as showing that - in a precise limiting sense – the proximity to (asymptotic) sufficiency yielded by the use of the MLE is replicated by the use of the approximate score.
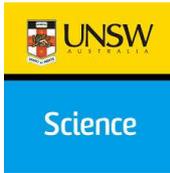
Application of the method in multiple parameter settings is addressed, with a separate treatment of scalar parameters, based on integrated likelihood techniques, advocated as a possible way of avoiding the curse of dimensionality. Some attention is given to a non-linear structure in which the state variable is driven by a continuous time process, with exact Bayesian inference typically infeasible in this case as a result of intractable transitions. The ABC method is demonstrated using the augmented unscented Kalman filter as a fast and simple way of producing an approximation in this setting. A stochastic volatility model for financial returns, based on a square root specification for volatility, is used for illustration, with the ABC method seen to produce quite accurate estimates overall of the exact marginal posteriors, which are accessible in this particular example. This is joint work with Brendan P.M. McCabe, Christian P. Robert and Ole Maneesoonthorn

Matthew Moores, Queensland University of Technology

Pre-processing for ABC-SMC with undirected graphical models

ABC is a useful method for Bayesian inference with intractable likelihoods, but existing algorithms require a large number of simulations of pseudo-data from the generative model. The computational cost of these simulations can be prohibitive for high-dimensional data. In this talk I will show that the scalability of ABC-SMC can be improved by performing a pre-computation step before model fitting. The output of this pre-computation can be reused across multiple datasets. I illustrate the approach using undirected graphical models, including the hidden Potts model for image analysis of large datasets.