A) Talks

1. Alquier, Pierre

   Title: Concentration and Robustness of Discrepancy-Based ABC via Rademacher Complexity

   Abstract: Classical implementations of approximate Bayesian computation (ABC) often employ summary statistics to measure the discrepancy among the observed data and the synthetic samples generated from each proposed value of the parameter of interest. However, finding suitable summary statistics, that are close to sufficiency, is challenging for most of the complex models for which ABC is required. In this talk we address these aspects by developing a novel unified and constructive framework, based on the concept of Rademacher complexity, to study ABC posteriors based on integral probability semi--metrics, which includes routinely-implemented discrepancies such as MMD and Wasserstein distance. In particular, we prove that the posterior will asymptotically concentrate to the best parameter under minimal assumptions. These ABC posteriors also enjoy very strong robustness properties. All these theoretical properties are supported by numerical experiments. This is a joint work with Sirio Legramanti (University of Bergamo) and Daniele Durante (Bocconi University, Milan).

2. Antoniano-Villalobos, Isadora

   Title: A Multidimensional Objective Prior Based on Scoring Rules

   Abstract: When it comes to multidimensional parameter spaces, the derivation of objective priors is, at best, challenging. A common practice is to assume prior independence and set up the joint prior as the product of marginals, obtained via ``standard'' objective methods, such as Jeffreys or reference priors. However, the assumption of independence a priori is not always reasonable and whether it is objective is still open to discussion. By extending a previously proposed objective approach based on scoring rules, we propose a novel objective prior for multidimensional parameter spaces which does not require the independence assumption. The proposed prior has the appealing property of being proper and it does not depend on the chosen model, but only on the parameter space considered.

3. Berger, Jim

   Title: Reflections on Intrinsic Priors

   Abstract: Twenty-six years ago, Luis Pericchi introduced intrinsic priors for hypothesis testing and model uncertainty. Despite decades of development by many statisticians, there are still many issues in their construction and use. This talk will present some ongoing work in the area, work being done jointly with Elias Moreno, Gonzalo Garcia-Donato and Luis.

4. Camerlenghi, Federico

   Title: Nested Nonparametric Processes

   Abstract: Modelling heterogeneity across subgroups of a specific population is a challenging problem in Bayesian nonparametrics. From a mathematical standpoint, heterogeneity is modelled by assuming that data are partially exchangeable, namely observations are divided into different, though related, groups and exchangeability holds true within the same group. The definition and investigation of Bayesian nonparametric priors to accommodate for partially exchangeable data is an active line of research, and several classes of priors have been proposed in the last decade. In this talk we focus on nested nonparametric processes, which allow a two-layered clustering both at the group and observational level. The nested Dirichlet process (NDP) is the archetype of a nested prior, but unfortunately it degenerates to the fully exchangeable case when there are ties across groups at the observational level. To overcome this drawback, we introduce and investigate novel classes of nested processes. In particular, we concentrate on the nested Common Atoms Model (CAM). The proposed CAM does not suffer from the degeneracy issue of the NDP and it is amenable to scalable posterior inference through the use of computationally efficient algorithms. We further discuss how to extend the proposed modelling framework to handle discrete measurements, and we conduct posterior inference on a real microbiome dataset. We finally hint some other proposals with the aim to improve the mathematical tractability and flexibility of the CAM.

5. Castellanos, María Eugenia

   Title: Model Selection Quantification in Presence of Missing Data

   Abstract: We consider the problem of model uncertainty quantification (MUQ) in the context of missing values in covariates. Within a Bayesian framework, it is easy to quantify uncertainty over models from a probabilistic view. The situation of missing data in explanatory variables is widespread in real applied problems. Usually, it is handled either by removing all rows with some unobserved value or using imputation techniques. However, the literature on MUQ is very scarce and has not been properly documented which are the implications of actual standard practices and if there is any better alternative. We approach the resulting problem by revisiting the fundamentals of Bayes factors (BFs), focusing on the situation with missing values in explanatory variables in linear regression models. Using a formal approach to handling missing values with a given imputation model, we can determine the prior predictive marginals and thus the BFs. In this context, we explore how to define priors over the model-specific parameters since standard priors, such as g-priors, are no longer valid due to their dependence on the design matrix which is regarded as random because of the missing data.

6. Drovanti, Chris

   Title: Bayesian Synthetic Likelihood – Asymptotics and Misspecification

   Abstract: Bayesian synthetic likelihood (BSL) is now a commonly used statistical method for likelihood-free Bayesian inference, which are applicable when the complex stochastic model under consideration can be efficiently simulated but lacks a tractable likelihood function. The method constructs an approximate likelihood by taking a vector of informative summary statistic as being multivariate normal, with the unknown mean and covariance estimated by simulation. In this talk I will describe the poor behaviour of BSL under model misspecification, and how it can be extended to produce robust inferences and help identify the source of misspecification. I will also discuss some asymptotic results for BSL, which for large sample sizes, relate to the form of the BSL posterior, its computational efficiency and adjustments when a misspecified form is assumed for the covariance. This is joint work with David Frazier, David Nott and Robert Kohn.

7. Durante, Daniele

   Title: The Role of Skew-Symmetric Distributions in Bayesian Inference: Conjugacy, Scalable Approximations and Asymptotics

   Abstract: A broad class of regression models that routinely appear in several fields of application can be expressed as partially or fully discretized Gaussian linear regressions. Besides incorporating the classical Gaussian response setting, this class crucially encompasses probit, multinomial probit and tobit models, among others, and further includes key extensions of such formulations to multivariate, non-linear and dynamic contexts. The relevance of these representations, which often act also as building-blocks of more complex constructions, has motivated decades of active research within the Bayesian field. A main reason for this constant interest is that, unlike for the Gaussian response setting, the posterior distributions induced by these models do not apparently belong to a known and tractable class, under the commonly-assumed Gaussian priors. This has led to the development of several alternative solutions for posterior inference relying either on sampling-based methods or deterministic approximations, that often experience scalability, mixing and accuracy issues, especially in high dimension. In this talk, I will review, unify and extend recent advances in Bayesian inference and computation for such a class of models, proving that unified skew-normal (SUN) distributions (which include Gaussians as a special case) are conjugate to the general form of the likelihood induced by these formulations. This result opens new avenues for improved posterior inference, under a broad class of widely-implemented models, via novel closed-form expressions, tractable Monte Carlo methods based on independent and identically distributed samples from the exact SUN posterior, and more accurate and scalable approximations from variational Bayes and expectation-propagation. These results will be further extended, in asymptotic regimes, to the whole class of Bayesian generalized linear models via novel limiting approximations relying on skew-symmetric distributions.

8. Forbes, Florence

   Title: Automatic Learning of Functional Summary Statistics for Approximate Bayesian Computation

   Abstract: Choosing informative summary statistics is a key and challenging task for successful inference via ABC algorithms. An important line of research towards automatic learning of summary statistics has started in the seminal paper of Fearnhead and Prangle [2012]. Their semi-automatic ABC approach targets approximations of the posterior mean as an optimal summary statistic under a quadratic loss. In this work, we propose to go beyond summary statistics as point estimators and consider functional summary statistics. Approximations of the full posterior distributions are used as such functional summaries. The parametric framework used to provide these approximations is a family of Gaussian mixtures. The whole procedure can be seen as an extension of the semi-automatic ABC framework and can also be used as an alternative to sample-based ABC approaches. The resulting ABC quasi-posterior distribution is shown to converge to the true one, under standard conditions. Performance is illustrated on both synthetic and real data sets, where it is shown that our approach is particularly useful when the posterior is multimodal. This is joint work with Hien Nguyen, TrungTin Nguyen and Julyan Arbel.

9. Forte, Anabel

   Title: Screening the Discrepancy Function of a Computer Model

   Abstract: Variable selection has been always a very important problem helping us to better understand and improve the process of interest. Much effort has been put in studying model selection-based variable selection methods in different scenarios such as linear or generalized linear models. But what happens in the world of mathematical models? How can we select variables when we depart from statistical models? This is a fundamental question since mathematical models require a careful study of the uncertainty behind them which should include the study of potentially related covariates. In this sense we can find inert covariates which are included in the mathematical model but has no effect in the final outcome, also covariates related with the studied process but not considered in the model or even variables related with the process and included in the model but in an incorrect way. In this work we consider the bias function usually used to assess the gap between the model and realty as a tool to perform this complex but necessary variable selection process. In particular, we use a bias function based on the one proposed by Linkletter et al (2006) which parameters regulate the importance of the covariates in the bias. Then we implement a model selection process using models which differ in the prior choice for those parameters and compute the corresponding Bayes factors using an efficient bridge sampling procedure. This is a joint work with Pierre Barbillion and Rui Paulo.

10. Grazian, Clara

    Title: Finding Structures in Observations: Consistent(?) Clustering Analysis

Abstract: Clustering is an important task in almost every area of knowledge: medicine and epidemiology, genomics, environmental science, economics, visual sciences, among others. Methodologies to perform inference on the number of clusters have often been proved to be inconsistent and introducing a dependence structure among the clusters implies additional difficulties in the estimation process. In a Bayesian setting, clustering in the situation where the number of clusters is unknown is often performed by using Dirichlet process priors or finite mixture models. However, the posterior distributions on the number of groups have been recently proved to be inconsistent. This talk aims at reviewing the Bayesian approaches available to perform via mixture models and give some new insights.

11. Gutiérrez-Peña, Eduardo

    Title: Model Selection as Point Estimation

    Abstract: Model selection can be regarded as a point estimation problem over a suitably defined extended class of models. We explore this idea in an objective Bayes context. In the M-closed view, under certain assumptions, estimating the 'parameter' that describes the true model leads to various forms of Bayes factors, but the idea can also give rise to other interesting procedures. On the other hand, in the M-completed view, the resulting procedure is akin to finding a 'maximum likelihood estimator'. Together with a weighted likelihood bootstrap scheme, this latter procedure can also be used to determine appropriate weights for Bayesian model averaging.

12. Held, Leonhard

    Title: Reverse-Bayes Methods for the Analysis of Replication Studies

    Abstract: I will report on recent work that builds on an approach to inference put forward over 70 years ago by IJ Good, by reversing the conventional prior-likelihood-posterior ("forward") use of Bayes' theorem. Such Reverse-Bayes analysis allows priors to be deduced from the likelihood by requiring that the posterior achieve a specified level of credibility. I will summarize the technical underpinning of this approach, and show how it opens up new approaches to common inferential challenges. I will show how Reverse-Bayes methods constitute a formal framework to challenge and substantiate scientific findings, particularly suitable to analyze replication studies.

13. Holmes, Chris

    Title: Predictive Inference: a Path Towards Objectivity?

    Abstract: De Finetti promoted the importance of predictive models for observables as the basis for Bayesian inference. The assumption of exchangeability, implying aspects of symmetry in the predictive model, motivates the usual likelihood-prior construction and with it the traditional learning approach involving a prior to posterior update using Bayes' rule. We discuss an alternative approach, treating Bayesian inference as a missing data problem for observables not yet obtained from the population needed to estimate a parameter precisely or make a decision correctly. This motivates the direct use of predictive models for inference, relaxing exchangeability to start modelling from the data in hand (with or without a prior). Martingales play a key role in the construction. Predictive models have widely accepted, established, empirical, criteria for construction and evaluation. This provides a useful route towards Bayesian objectivity that complements the traditional approach based on exchangeability and the pursuit of objective priors. This is joint work with Stephen Walker and Edwin Fong.

14. Liang, Feng

Title: Automatic Relevance Determination with Statistical Guarantees

Abstract: Automatic Relevance Determination (ARD) is a popular Bayesian procedure for variable pruning. The relevance of each variable is encoded in a hyper-parameter in the ARD prior, which is then determined automatically through the data. Despite its empirical success, little is known about the theoretical properties of the ARD procedure. In this paper, we study the ARD procedure in the context of high-dimensional linear regression under sparsity constraints. We propose to estimate the relevance parameters using a variational approach, which approximates the posterior distribution by independent Gaussian distributions, one for each regression coefficient. We show that for some coefficients the corresponding Gaussian distribution will degenerate to a point mass at zero; that is, some variables will be automatically filtered out by the ARD procedure. We further establish convergence results, in terms of parameter estimation and variable selection, for the resultant variational solution.

15. Liseo, Brunero

Title: An Extension of the Unified Skew-Normal Family of Distributions and Applications to Bayesian Binary Regression

Abstract: We consider the general problem of Bayesian binary regression with a large number of covariates. We introduce a new class of distributions, the Perturbed Unified Skew Normal (PSUN), which generalizes the SUN class and show that it is conjugate to any binary regression model, such that the link function can be expressed as a scale mixture of Gaussian densities. We discuss in detail the *probit* and *logistic* cases. We show that, when the number of covariates p is larger than the sample size n, it is possible to produce an exact posterior simulation both in the probit and logit frameworks. The approach also works in the $n > p$ case, where a Gibbs sampler algorithm is necessary. This is a joint work with Paolo Onorati**,** Ph.D. Student, Sapienza Università di Roma.

16. Mena, Ramsés

Title: Exchangeable Random Measures and Stick-Breaking Priors

Abstract: We will present the general class of stick-breaking processes with exchangeable length variables, which generalize well-known Bayesian nonparametric priors in an unexplored direction. Apart from giving existence and support conditions for this new class, we explore several cases of interest, in particular being able to modulate the stochastic ordering of the weights and recovering Dirichlet and Geometric priors as extreme cases. A general formula for the distribution of the latent allocation variables is derived and an MCMC algorithm is proposed for density estimation purposes.

17. Pérez, Maria-Eglée

Title: Building Bridges: Bayesian Approaches for Increasing Reproducibility in Null Hypothesis Significance Testing

Abstract: In recent years there has been a wide discussion on the validity of methods for Null Hypothesis Significance Testing (NHST). This talk will present joint work with Luis Raúl Pericchi, aimed to use Bayesian ideas for providing alternatives to the $p = 0.05$ significance level via adaptive significance levels (Pérez and Pericchi, 2014). A refinement of adaptive significance levels for linear models (Vélez, Pérez and Pericchi, 2022) involves not only the sample size, but also the correlation structure via the design matrix. An alternative approach is the calculation of minimum Bayes factors as the lower bound presented in Sellke et al (2001). We propose an adjustment of the minimum Bayes factor with the information to approximate it to an exact Bayes factor, including cases when p is a pseudo p-value as defined by Casella and Berger (2001).

18. Rockova, Veronika

Title: Adversarial Bayesian Simulation

Abstract: In the absence of explicit or tractable likelihoods, Bayesians often resort to approximate Bayesian computation (ABC) techniques for inference. Our work bridges ABC with deep neural implicit samplers based on generative adversarial networks (GANs) and adversarial variational Bayes. Both ABC and GANs are simulation techniques that compare aspects of observed and fake data targeting posteriors and likelihoods, respectively. We develop a Bayesian GAN (B-GAN) sampler that directly targets the posterior, not the likelihood. B-GAN is driven by a deterministic mapping trained on the ABC reference table using conditional GANs. Once the mapping has been learned, iid posterior samples are obtained by filtering noise at negligible additional cost. We propose two post-processing local refinements using (1) data-driven proposals with importance reweighing, and (2) variational Bayes. We support our findings with frequentist-Bayesian results, showing that the typical total variation distance between the true and approximate posteriors converges to zero for certain neural network generators and discriminators. Our findings on simulated data show highly competitive performance relative to some of the most recent likelihood-free posterior simulators.

19. Rodriguez, Abel

Title: Default Priors and Robust Estimation for Generalized Linear Models

Abstract: As a Master's student in the late 90s, Luis Pericchi had a big influence in my training and in how I think about statistical inference problems. This talk highlights how that influence endures and carries on to some of my own students. The first part of the talk discusses some of our recent work on developing default priors for model comparison in the context of generalized linear models. The proposed methodology relies on training samples to calibrate improper priors. While this type of approach goes back at least to the 70s and 80s, my first exposure to it was through the Intrinsic Bayes Factors introduced in Berger and Pericchi (1996). The second part of the talk focuses on the development of robust priors for estimation, and their role in the context of variable selection. More specifically, we discuss the specification of continuous shrinkage priors with adaptive tail coefficients using Polya Trees. Some of the original inspiration for this work comes from my early exposure to Pericchi and Walley (1991) and, much more recently, to Perez, Pericchi and Ramirez (2017). The work presented in this talk is a collaboration with my Ph.D. student Anupreet Porwal.

20. Rossell, David

Title: Improper Models for Data Analysis

Abstract: A long-standing debate in Statistics and Machine Learning is whether one should use a model-based approach, or a framework defined by minimizing a loss function. We argue that both can be viewed as part of a common framework where loss functions may define an improper model, i.e. the data is described by probabilities that integrate to infinity. We discuss how to interpret such a framework via relative probabilities (which are well-defined, even for improper models) and how to embed it within general Bayes. As a particular instance, we define a loss function based on the Hyvarinen score that can be interpreted as giving optimal infinitessimal relative probabilities. We provide posterior consistency and model selection results, including for challenging loss function hyper-parameters (e.g. the learning rate in generalised Bayes / Gibbs posteriors / PAC Bayes). As examples we consider robust regression and non-parametric density estimation where popular loss functions define improper models for the data and hence cannot be dealt with using standard model selection tools. These examples illustrate advantages in robustness-efficiency trade-offs and provide a Bayesian implementation for kernel density estimation, opening a new avenue for Bayesian non-parametrics.

21. Telesca, Donatello

Title: Functional Partial Membership Models

Abstract: Partial membership models, or mixed membership models, are a flexible unsupervised learning method that allows each observation to belong to multiple clusters. In this talk, we propose a Bayesian partial membership model for functional data. By using the multivariate Karhunen-Loève theorem, we are able to derive a scalable representation of Gaussian processes that maintains data-driven learning of the covariance structure. Within this framework, we establish conditional posterior consistency given a known feature allocation matrix. Compared to previous work on partial membership models, our proposal allows for increased modeling flexibility, with the benefit of a directly interpretable mean and covariance structure. Our work is motivated by studies in functional brain imaging through electroencephalography (EEG) of children with autism spectrum disorder (ASD). In this context, our work formalizes the clinical notion of ``spectrum" in terms of feature membership probabilities.

B) Tutorials

1. Ghoshal, Subhashis

   Tutorial: Bayesian Inference in High-Dimensional Models

   Abstract: Models indexed by parameters of high dimension, sometimes more than the available sample size, are now routinely adopted in various applications. A sensible inference is possible in such a situation only by properly utilizing a lower-dimensional structure. Non-Bayesian procedures for making inferences in the high-dimensional setting, typically by penalizing to induce sparsity in the solution, were developed. Bayesian methods were proposed more recently, where the prior takes care of the sparsity structure. These methods have the natural ability to also automatically quantify the uncertainty of the inference through the posterior distribution. Theoretical studies of Bayesian procedures in high-dimension have been carried out recently. Questions that arise are whether the posterior distribution contracts near the true value of the parameter at the minimax optimal rate, whether the correct lower-dimensional structure is discovered with high posterior probability, and if a credible region has adequate frequentist coverage. In this tutorial, we review the properties of Bayesian and related methods for several high-dimensional models such as many normal means problem, linear regression, generalized linear models, Gaussian and non-Gaussian graphical models, and others. Effective computational approaches will be also discussed.

2. Guindani, Michele

   Title: Bayesian Nonparametrics Methods with applications to the analysis of neuroimaging data

   Abstract: In this tutorial, we introduce a set of widely employed, flexible, Bayesian nonparametric (BNP) models for learning features of observables with minimal information. We will illustrate the use of BNP methods for characterizing distributional heterogeneity in the neurosciences, where time-series data are assumed to be organized in different, but related, units (e.g., neurons and/or regions of interest). For example, we will discuss models for multi-subject analysis that will identify population subgroups characterized by similar brain activity patterns as well as several approaches to characterize how neuronal responses vary in response to different external stimuli in single subjects. This introductory presentation will start from outlining first principles in Bayesian nonparametric modeling and will conclude with a discussion of current and upcoming trends.

3. Papaspiliopoulos, Omiros

   Title: Scalable Bayesian computation for hierarchical models: methods and theory

   Abstract: I will give an overview of methodologies with provable scalability properties for fundamental classes of hierarchical models. I will consider both optimisation-based (such as those for MAP estimation and for variational approximation) and Monte Carlo sampling methods. Within optimization, I will discuss both direct (e.g., based on Laplace approximations and sparse linear algebra) and iterative (such as coordinate-wise ascent) methods. Scalability refers to computational complexity that is linear in the size of the model and the number of data points. For iterative schemes the computational complexity is the product of cost per iteration and the number of iterations to achieve certain error. For variational approximations I will also show results about how the approximation error scales with the size of the model and the number of data points.

4. Steel, Mark

   Title: An Introduction to Bayesian Model Averaging

   Abstract: Model uncertainty is a pervasive problem in all applied fields and cannot be ignored without potentially serious consequences. It is important to realize that simply sticking to a single model (as is the traditional approach) has the dangerous consequence of presenting conclusions with an excess of confidence, since the results do not take into account the host of other possible models that could have been tried. Even if we do acknowledge the existence of other models, the use of model selection techniques and presenting our inference conditional upon the single chosen model typically leads to an underestimation of our uncertainty and can induce important biases. In this tutorial, I focus on Bayesian Model Averaging (BMA) which is a formal, principled and natural way to approach the problem, following the simple rules of probability calculus. We discuss the role of prior assumptions and the implementation of BMA through Markov chain Monte Carlo methods in large problems. The main ideas are introduced in the simple context of variable selection in linear regression models, but I also mention other, more challenging, settings. Some illustrative applications are briefly discussed.