Bull. Aust. Math. Soc. **104** (2021), 171–172
doi:10.1017/S0004972720001124

EXPERIMENTAL DESIGN FOR DEPENDENT DATA

S. G. J. SENARATHNE

(Received 20 August 2020; first published online 9 November 2020)

2020 *Mathematics subject classification*: primary 62K05; secondary 62L05, 62L12.

Keywords and phrases: adaptive design, Laplace approximation, model discrimination, parameter estimation, prediction, sequential Monte Carlo, total entropy.

An experimental design defines the values of controllable variables or covariates with which to collect data. In planning an experiment, such a design can be optimised such that specific analysis goals including parameter estimation, model discrimination and prediction can be efficiently achieved, where 'efficient' means being economic in the number of data points or required experimental resources. In many experiments conducted in pharmacology, agriculture and environmental science, dependent data such as bivariate mixed outcomes, repeated measures data and spatially dependent data are observed. This poses some unique challenges in finding experimental designs for such data. Unfortunately, this has received considerably less attention in the literature when compared to observing univariate and/or independent data from experiments, motivating a need to develop new approaches to find designs for collecting dependent data.

In this thesis, we develop new methods for designing experiments where dependent data are observed. The scope of this work is Bayesian design, that is, designs found on the basis of undertaking a Bayesian analysis of the data. We consider such an approach due to the rigorous handling of uncertainty and availability of important utility functions such as total entropy. However, undertaking Bayesian design typically requires sampling from or approximating a large number of posterior distributions resulting in finding designs being more computationally difficult than inference alone. This motivates the need to develop fast posterior approximations in Bayesian design in general, but especially when dependent data are observed, as typically the models for such data are complex to ensure that the dependence between observations is appropriately modelled. Through addressing this need for dependent data, we advance the field of experimental design and make approaches from this field more applicable to real-world experiments.

Thesis submitted to Queensland University of Technology in February 2020; degree approved on 26 June 2020; supervisors James McGree and Christopher Drovandi.

In the first part of the thesis, we extend current methods in Bayesian adaptive design for dependent data. In particular, the standard sequential Monte Carlo (SMC) algorithm is extended to derive designs for bivariate mixed outcomes [1]. In the next part of the thesis, a new adaptive design algorithm is developed to derive designs more efficiently when compared to the SMC algorithm. In this new algorithm, a Laplace importance sampling based method is proposed to update the posterior distribution of the parameters efficiently and, hence, to reduce computational time required to find adaptive designs [2]. These developments are then extended to experiments where repeated measures data are observed. In general, such data are described within a mixed modelling framework, in which random effects are used to capture the dependency structure in the data, for example, the between-subject variability. For such models, we propose an adaptive design algorithm to derive designs for estimating both population and random effect parameters within a mixed model setting [3].

Finally, the last section of the thesis tackles an important geostatistical design problem in which spatially dependent data are observed. Such design problems are often concerned with accurate spatial prediction under parameter uncertainty. This motivates the need to consider a utility function that minimises the uncertainty from two sources: (1) uncertainty about the parameter values and (2) uncertainty in the predicted response conditional on a parameter value. For this purpose, we propose a novel dual-purpose utility function for parameter estimation and spatial prediction and develop a fast approximation for this utility such that it can be used to design spatial processes where bivariate mixed spatial data are observed. A paper, 'Bayesian design for minimising uncertainty in spatial processes', has been submitted for publication.

References

[1] S. G. J. Senarathne, C. C. Drovandi and J. M. McGree, 'Bayesian sequential design for Copula models', *TEST* **29**(2) (2019), 454–478.

[2] S. G. J. Senarathne, C. C. Drovandi and J. M. McGree, 'A Laplace-based algorithm for Bayesian adaptive design', *Stat. Comput.* **30** (2020), 1183–1208.

[3] S. G. J. Senarathne, A. M. Overstall and J. M. McGree, 'Bayesian adaptive *n*-of-1 trials for estimating population and individual treatment effects', *Stat. Med.*, to appear.

S. G. J. SENARATHNE, School of Mathematical Sciences,
Science and Engineering Faculty,
Queensland University of Technology, Brisbane, Australia
e-mail: jagath.gedara@hdr.qut.edu.au