4OR-Q J Oper Res manuscript No.
(will be inserted by the editor)

Heuristic and metaheuristic algorithms for the generation of optimal experimental designs

Daniel Palhazi Cuervo

Received: date / Accepted: date

This is a summary of the author's Ph.D. thesis, supervised by Peter Goos and Kenneth Sörensen, and defended on June 15th, 2015 at the Universiteit Antwerpen, Belgium. The thesis is written in English and is available from the author upon request at daniel.palhazicuervo@uantwerpen.be. The topic of this Ph.D. lies in the intersection between *operations research* and *statistics*: it proposes new algorithmic techniques for the generation of optimal experimental designs.

The purpose of an experiment is to identify the influence that a set of experimental variables has on the process under study. By systematically manipulating the settings of these variables, it is possible to quantify how and to which extent they affect one or more response variables that measure the process's behaviour. The design of an experiment mainly consists in determining the number of experimental runs, the settings of the experimental variables in each run, and the sequence in which the runs need to be executed. This should be done with the purpose of maximizing the amount of information produced by the experiment. Several standard experimental designs have been proposed to achieve this goal. Although these designs have very good properties, they cannot always be applied to the complex experimental scenarios found in practice. A better strategy is to generate a custom design that is specifically tailored to the characteristics of the process. This approach is called *optimal design of experiments* (ODOE) and its goal is to find the best possible design that can be carried out for the experimental scenario at hand. To this end, the ODOE approach treats the generation of a design as an optimization problem and makes use of different optimization algorithms to solve it.

The benefits of using algorithmic techniques for the ODOE have been extensively documented in the literature. This approach, however, has been criticized by some members of the statistical community and is not yet considered a routine practice. One of their main arguments is that the designs generated by algorithmic methods do not always match the quality of the standard experimental designs traditionally used. More importantly, for many scenarios where a standard design is known to be the best alternative, the existing algorithmic techniques fail to generate such a design. Many statisticians are therefore reluctant to opt for the ODOE approach and remain faithful to standard designs. This dissertation addresses this kind of criticism levelled against the ODOE approach: it proposes new and more efficient algorithmic techniques for the generation of optimal experimental designs. These techniques are based on the family of optimization algorithms known as *metaheuristics*. As shown by an extensive set of computational experiments, the proposed algorithms are able to generate designs with better quality and in shorter execution times than other algorithmic methods. Additionally, it is also shown how the flexibility of these algorithms can be leveraged in order to generate new designs that, in many cases, have better properties than the standard designs.

This dissertation is divided into two parts, each dealing with different types of experiments. The first part focuses on optimal designs for industrial experiments. These kinds of experiments are widely used for the development and improvement of products and processes. In this part, we describe a new selection strategy for the coordinate-exchange algorithm (CEA), one of the most important algorithms for the generation of optimal experimental designs. This improved version of the CEA is then integrated into an iterated local search (ILS) algorithm. The ILS algorithm is able to generate better designs for large-scale *screening* experiments than the restart strategy used by the original CEA. In this part of the dissertation, we also describe a new algorithmic framework for the generation of optimal designs for two-stratum experiments (i.e., *blocked* and *split-plot* experiments) with flexible constraints. This framework generates designs for two-stratum experiments in which the number of groups and the number of observations per group are not fixed, but limited only by upper bounds. The framework is based on a two-level local search approach, one for each stratum in the experiment. The lower optimization level applies an extension of the CEA while the higher optimization level applies a strategy similar to that of the variable neighborhood search. The results show that this additional flexibility in the design generation process can significantly improve the quality of the designs.

The second part of the dissertation focuses on optimal designs for stated choice experiments. These experiments are widely used to study how people make choices and to identify the elements that drive people's preferences. In this part, we propose an integrated algorithm for the generation of optimal designs for stated choice experiments with partial profiles. This algorithm optimizes the set of constant attributes and the levels of the varying attributes simultaneously. The results show that the designs produced by the integrated algorithm outperform those obtained by existing algorithms, and match the optimal designs that have been analytically derived for a number of benchmark experiments. We also extend the integrated algorithm to generate designs for stated choice experiments for estimating models with interaction effects. We show that the performance and the flexibility provided by this algorithm are major arguments in favour of its use.