Computational Statistics and Data Analysis xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Robustness of classical and optimal designs to missing observations

Byran J. Smucker[a,*], Willis Jensen[b], Zichen Wu[a], Bo Wang[a]

[a] *Department of Statistics, Miami University, Oxford, OH, United States*
[b] *W.L. Gore & Associates, Flagstaff, AZ, United States*

ARTICLE INFO

Article history:
Received 18 February 2016
Received in revised form 29 November 2016
Accepted 1 December 2016
Available online xxxx

Keywords:
Central composite designs
Factorial designs
Missing data
D-efficiency
I-efficiency

ABSTRACT

Missing observations are not uncommon in real-world experiments. Consequently, the robustness of an experimental design to one or more missing runs is an important characteristic of the design. Results of an evaluation of the robustness of classical and optimal designs to missing observations are presented, and optimal designs fare relatively well in terms of robustness compared to classical designs. Additionally, a modified version of an existing robustness criterion is used to construct designs that are robust to missing observations.

1. Introduction

When designing experiments, there is a tension between robustness and optimality. Optimality suggests a single, best design while robustness implies quality under a broad range of conditions. Robustness is crucial in the real world of experimentation, as noted generally by Wendelberger (2010). There is also a growing acknowledgment that multiple criteria should be used when constructing optimal designs (e.g. Lu et al., 2011; Gilmour and Trinca, 2012). In the face of many different kinds of variability and uncertainty, it is important for designs that are optimal for estimation or prediction to be robust in other ways. Since there is generally not a single design that is optimal in all aspects, an acceptable compromise is to consider appropriate tradeoffs between key characteristics. Though Box and Draper (1975) gave a list of fourteen attributes that generally relate to a design's robustness and practicality, one aspect not specifically mentioned is robustness to missing values. Box and Draper do mention outliers or "wild observations", and to the extent that these observations can be omitted from the analysis in a principled way they would fall under the purview of our work. However, we discuss in particular the situation where no values are obtained for a particular run. The primary goal of our work is to consider some common classical and optimal designs in terms of their robustness to missing observations. Secondarily, we adapt an existing criterion to construct some new designs that are missing-robust.

Several robust-to-missing-observations criteria have been proposed in the literature. Box and Draper (1975) pointed out the connection between the diagonal values of the hat matrix and a design's robustness and applied their related criterion to central composite designs. Siddiqi (2011) applied a similar criterion to the smaller response surface designs of Draper and Lin (1990). Herzberg and Andrews (1976) considered the probability that a design will not estimate the desired model,

* Correspondence to: 311 Upham Hall, 100 Bishop Circle, Oxford, OH 45056, United States. Fax: +1 513 529 0989.
E-mail address: smuckerb@miamioh.edu (B.J. Smucker).

http://dx.doi.org/10.1016/j.csda.2016.12.001
0167-9473/© 2016 Published by Elsevier B.V.

and Andrews and Herzberg (1979) suggested maximizing the expected value of the determinant of the information matrix under possible missing observations. Akhtar and Prescott (1986) developed a criterion that minimizes the maximum loss due to missing observations and applied it to the evaluation and generation of central composite designs, and Ahmad and Gilmour (2010) used this measure to study the robustness of so-called subset designs (Gilmour, 2006). Herzberg et al. (1987) proposed equi-information designs, which retain equal information when up to two design points are missing. Imhof et al. (2002) presented results based on a different maximin criterion assuming continuous designs.

Missing from the literature is (1) a systematic comparison of classical and optimal designs in terms of their robustness to missing observations; and (2) a general algorithmic approach to generating designs that are robust to missing observations. We note that Hackl (1995) did generate robust designs for small response surface experiments, though they have not been formally published. Furthermore, a referee pointed us to a recent paper (da Silva et al., 2016) that incorporates a leverage-based missing-robustness criterion used within a larger compromise criterion that balances pure error and lack-of-fit degrees of freedom (see Gilmour and Trinca, 2012).

The rest of the paper unfolds as follows. We first motivate our work by explaining why missing runs occur in practical situations. We then outline several evaluative measures of missing-robustness. Next we demonstrate the impact of missing runs on some standard designs, including fractional factorial and D-optimal designs for screening experiments, and central composite and I-optimal designs for prediction/optimization. Included in these comparisons are some new designs constructed using a version of a robustness criterion from the literature. We conclude with a discussion and some thoughts on future research.

2. Motivation

A common solution in the statistical literature for dealing with missing observations is imputation. Imputation methods fill in the missing values based on the other data points, to allow for model fitting. Though imputation methods are common for large observational datasets, we believe these methods are problematic in an experimental design situation because of the relatively small number of runs. Thus, we do not consider this a viable solution to many experimental, missing-observation problems.

Obviously, if possible, the first priority is to redo the runs to fill in the missing observations. However, if an experimental run fails during an experiment, it is sometimes difficult or impossible to redo. Consider, as an example, a manufacturing process experiment that uses actual production equipment. Laboratory or pilot plant experiments may be useful in the early phases of experimentation to suggest acceptable factor ranges and a rough idea of the optimal factor region. Ultimately a final experiment using the production equipment will be necessary to determine the best operating conditions for that equipment, because the results from the laboratory or pilot plant are not typically directly generalizable due to varying conditions in the respective facilities. In order to maximize manufacturing throughput and minimize production costs, it is not unusual to have production equipment running constantly. It becomes very expensive – though necessary – to pause production to execute a series of experimental runs on that equipment. As soon as the experimental runs are completed, production resumes and the experimental runs are tested and analyzed. If a run was inadvertently missed, there will be strong resistance to shutting down production again for a single run. In this environment, testing errors made on the experimental runs could render one or more of the runs unusable. These missing values are generally not tied to one of the experimental factors of interest and can thus be treated as missing at random.

Missing runs can also occur in new product development. With proper planning, an adequate amount of custom-made raw material can be obtained from a supplier for use in an experiment. In some situations, however, the material is limited because of cost or scarcity. In this case, a failed run would require additional material which is likely to be costly and time-inefficient to obtain. A probable outcome is that the run is simply lost to the experiment.

Complex processes can also cause missing runs. In today's global and complex manufacturing environment, it is not unusual to have processes whose steps span multiple, possibly far-flung, locations. In these situations, it may be difficult or impossible to do a full experiment across all process steps involving all factors so experiments will be performed on localized portions of the process. However, in order to obtain the final measurements on the experimental runs it may be necessary to continue processing them for multiple steps beyond those within the scope of the experimental factors of interest. Inadvertent events in downstream processing may result in a loss of some of the experimental runs and it is difficult to redo the runs without a significant time lag or change in processing conditions.

As mentioned previously, an assumption throughout is that any missing observations are missing at random. There are certainly many cases for which the assumption does not hold—for instance, runs near the edge of the experimental region may have a higher probability of yielding unusable or no information. This can occur with a poor selection of factor levels and can be prevented with more thorough exploration of factors ranges prior to formal experimentation. We concentrate on those cases where the missing observations are independent of the factor levels used.

Industrial experience suggests that the number of missing observations is generally low ($<20\%$). If the percentage of missing observations is large, it is clear that little information may be obtained from the experimental results and it is likely that more fundamental issues need to be addressed before attempting to execute the experiment. Here, we focus on situations in which an observation is missing with a low probability. This is important to note because the evaluations we perform implicitly or explicitly make this low-probability assumption.

Please cite this article in press as: Smucker, B.J., et al., Robustness of classical and optimal designs to missing observations. Computational Statistics and Data Analysis (2017), http://dx.doi.org/10.1016/j.csda.2016.12.001