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

Q1 Bayesian crossover designs for generalized linear models

Q2 Satya Prakash Singh, Siuli Mukhopadhyay *

Department of Mathematics, Indian Institute of Technology Bombay, Mumbai 400 076, India

ARTICLE INFO

Article history:
Received 7 December 2015
Received in revised form 15 April 2016
Accepted 8 June 2016
Available online xxxx

Keywords:
Bayesian designs
Count data
Efficiency
Gamma response
Generalized estimating equations
Logistic regression

ABSTRACT

This article discusses optimal Bayesian crossover designs for generalized linear models. Crossover trials with t treatments and p periods, for $t <= p$, are considered. The designs proposed in this paper minimize the log determinant of the variance of the estimated treatment effects over all possible allocation of the n subjects to the treatment sequences. It is assumed that the p observations from each subject are mutually correlated while the observations from different subjects are uncorrelated. Since main interest is in estimating the treatment effects, the subject effect is assumed to be nuisance, and generalized estimating equations are used to estimate the marginal means. To address the issue of parameter dependence a Bayesian approach is employed. Prior distributions are assumed on the model parameters which are then incorporated into the D_A-optimal design criterion by integrating it over the prior distribution. Three case studies, one with binary outcomes in a 4×4 crossover trial, second one based on count data for a 2×2 trial and a third one with Gamma responses in a 3×2 crossover trial are used to illustrate the proposed method. The effect of the choice of prior distributions on the designs is also studied. A general equivalence theorem is stated to verify the optimality of designs obtained.

1. Introduction

In this article we introduce Bayesian optimal crossover designs for generalized linear models (GLMs). Crossover trials with t treatments and p periods, for $t \leq p$ are considered. The designs selected minimize the log determinant of the variance–covariance matrix of the treatment effects, over all possible allocation of the n subjects to the treatment sequences. Due to the dependence of the variance matrix on the model parameters a Bayesian approach is proposed.

Crossover designs were originally developed to be used in agricultural sciences (Cochran, 1939). Later, these repeated measurement designs were found to be useful in many other fields, such as pharmaceutical and clinical trials, bioequivalence and biological studies. Optimal crossover designs for normal response have been studied by many researchers, namely Hedayat and Afsarinejad (1975, 1978), Cheng and Wu (1980), Laska et al. (1983), Laska and Meisner (1985), Stufken (1991), Carrire and Reinsel (1993), Kushner (1997, 1998) and Carriere and Huang (2000). For a detailed review of crossover designs, we would like to refer to the paper by Bose and Dey (2013) and books by Bose and Dey (2009), Senn (2002) and Jones and Kenward (2014).

Most of the available literature on optimal crossover designs (as discussed above) mainly focuses on normal responses. However, in biological studies, very often we find responses that are non-normal (Layard and Arvesen, 1978; Forster, 1992) and have to be modeled using a generalized linear model (GLM). While methods for analyzing GLM data arising from crossover trials are available in Senn (2002) and Jones and Kenward (2014), the question of designing such studies for

* Corresponding author.
E-mail address: siuli@math.iitb.ac.in (S. Mukhopadhyay).

http://dx.doi.org/10.1016/j.csda.2016.06.002
0167-9473/

GLMs in an optimal manner does not seem to have been much explored in the statistical literature. Waterhouse et al. (2006) studied optimal 2×2 crossover trial for binary data in some special cases, like the carryover effect is proportional to the direct treatment effect and no period effects are considered. Adaptive crossover designs restricted to two period two treatment binary data useful in clinical trials have also been investigated by Bandyopadhyay et al. (2009).

In this article, we study optimal Bayesian crossover designs for GLMs. Three case studies based on non-normal responses are used to illustrate the proposed methodology. Generalized estimating equations of Liang and Zeger (1986) are used to estimate the marginal means. The correlation between observations within subjects are modeled using a "working correlation structure", which is assumed to be compound symmetric or auto regressive in nature. Since the main interest is in estimating the treatment effects, the subject effects are taken as nuisance parameters. As in all GLM designs, the variance of the treatment effect estimator depends on the model parameters. To address the issue of the parameter dependence and obtain robust designs we propose the Bayesian approach to design selection. Bayesian designs have been a popular choice whenever the variance–covariance matrix depends on the model parameters, for some references see Chaloner and Larntz (1989), Dette and Sperlich (1994), Woods and Van de Ven (2011) and Mylona et al. (2014). In our approach, a prior distribution is assumed on the model parameters, which is then incorporated into an appropriate objective function (variance of the treatment contrast) by integrating and averaging over the prior distribution. Similar to our Bayesian design criterion, an average criterion called A-criterion has been used before for crossover designs for normal responses by Kempton et al. (2001), Baily and Kunert (2006), Zheng (2013) and Li et al. (2015).

2. Case studies

For illustration purpose we consider three case studies based on crossover trials involving binary, count and Gamma responses.

2.1. A four periods four treatments binary response crossover trial

The first case study presented here is from a trial based on the four-period, four treatment Williams design. It has been reported in Kenward and Jones (1992). The four treatments are denoted by A, B, C and D. Eighty subjects are randomly assigned to the four treatment sequences $\{ABCD, BDAC, CADB, DCBA\}$, with about twenty subjects allocated to each treatment sequence. The response is a binary outcome taking values 1 and 0 based on patient relief and no relief, respectively.

The research question which arises from the above case study is why did the experimenter select the 4 treatment sequences $\{ABCD, BDAC, CADB, DCBA\}$ forming a Williams design (Williams, 1949). Is this the best possible selection of treatment sequences? The book by Bose and Dey (2009, page 40) shows that for normal response crossover models, for the 4 treatment and 4 periods case, Williams design is the optimal design. But can we be sure that the same design applies to a binary response crossover framework as well? Does the selected design change if the correlation structure between observations change say, from equicorrelated to auto regressive structure?

2.2. Two periods two treatments Poisson response crossover trial

This study is based on an example described in Layard and Arvesen (1978). Two drugs, standard drug A and an innovation drug B, is administered for controlling angina in 20 patients. It is known that the innovative drug B is no worse than the standard drug A. For a given patient, number of angina attacks on weekly basis is assumed to follow a Poisson distribution (Layard and Arvesen, 1978). Number of attacks for each patient of consecutive two weeks are recorded. Treatment sequences considered are $\{AB, BA\}$ and 10 patients are assigned to each of the treatment sequences. This is a 2-treatments 2-periods crossover trial.

As in case study I, the question arises why does the experimenter choose the design $\{AB, BA\}$. Is this the best or most efficient design under the repeated measures setup when responses follow a Poisson distribution?

2.3. Three periods two treatments Gamma response trial

The length of hospital stay is an important measure of the success of hospital activity, costs incurred by patients and the treatment administered to a patient. However, its empirical distribution is often right skewed and a Gamma distribution with a log link has been seen to be a good fit (Faddy et al., 2009). In this case study we consider a crossover trial where two treatments are applied over three periods and length of hospital stay, assumed to having a Gamma distribution, is the primary end point.

As in the earlier two case studies, we investigate the best design for a two treatment three periods design with a gamma response.

Please cite this article in press as: Singh, S.P., Mukhopadhyay, S., Bayesian crossover designs for generalized linear models. Computational Statistics and Data Analysis (2016), http://dx.doi.org/10.1016/j.csda.2016.06.002