Journal of the Korean Statistical Society 42 (2013) 105–116

Contents lists available at SciVerse ScienceDirect

Journal of the Korean Statistical Society

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

Sequential estimation for covariate-adjusted response-adaptive designs

Yuan-chin Ivan Chang [a], Eunsik Park [b,*]

[a] Institute of Statistical Science, Academia Sinica, Taipei 11529, Taiwan

[b] Department of Statistics, Chonnam National University, Gwangju 500-757, Republic of Korea

ARTICLE INFO

Article history:
Received 19 December 2011
Accepted 2 June 2012
Available online 23 June 2012

AMS 2000 subject classifications:
primary 62L12
secondary 62L10

Keywords:
Covariate-adjustment
Logistic regression
Response-adaptive design
Sequential estimation
Stopping time
Targeted drug
Utility function

ABSTRACT

In clinical trials, a covariate-adjusted response-adaptive (CARA) design allows a subject newly entering a trial a better chance of being allocated to a superior treatment regimen based on cumulative information from previous subjects, and adjusts the allocation according to individual covariate information. Since this design allocates subjects sequentially, it is natural to apply a sequential method for estimating the treatment effect in order to make the data analysis more efficient. In this paper, we study the sequential estimation of treatment effect for a general CARA design. A stopping criterion is proposed such that the estimates satisfy a prescribed precision when the sampling is stopped. The properties of estimates and stopping time are obtained under the proposed stopping rule. In addition, we show that the asymptotic properties of the allocation function, under the proposed stopping rule, are the same as those obtained in the non-sequential/fixed sample size counterpart. We then illustrate the performance of the proposed procedure with some simulation results using logistic models. The properties, such as the coverage probability of treatment effect, correct allocation proportion and average sample size, for diverse combinations of initial sample sizes and tuning parameters in the utility function are discussed.

1. Introduction

From an ethical viewpoint, it is desirable to minimize the number of subjects allocated to inferior treatments in the course of a clinical trial without jeopardizing the generation of useful and meaningful statistical inferences. The response adaptive (RA) design in clinical trials (Hu & Rosenberger, 2006; Zelen & Wei, 1995) is dedicated to this purpose. The advantage of an RA design is that the information collected from subjects previously entering the trial can be used to adjust the allocation probability so that a newly entering subject can have a better chance of being allocated to a superior treatment. Because of the sequential characteristic in this process, sequential statistical methods should be used in order to efficiently analyze these kinds of data sets. Since data collected in this manner are no longer independent, sequential methods that rely on assumption of independent observations are not valid. Moreover, due to innovation in genomic technologies and the nature of developing targeted drugs (Simon & Maitournam, 2005), it is natural to incorporate the information available on individual covariates that have a strong influence on responses to a model, since they may be associated with the efficacy of treatments. Hence, the existence of an interaction between treatment and covariate becomes a reasonable presumption as far as, for example, a targeted drug is concerned. Traditionally, we apply an RA design by assuming there is no treatment-covariate interaction effect. However, when there is an interaction between covariates and treatments, a method that uses an RA design will make incorrect treatment allocation. This is especially the case when a targeted drug or other adaptive treatment

* Corresponding author. Tel.: +82 10 8623-0332; fax: +82 62 530 3449.
E-mail address: espark02@gmail.com (E. Park).

1226-3192/$ – see front matter
doi:10.1016/j.jkss.2012.06.001

strategy is being used. Thus, it is reasonable to assume that a CARA design should perform better than an RA design in terms of correct allocation proportions when this interaction situation should not be ignored. However, little work has been done on CARA designs. In addition to the ethical considerations, this is a further good reason for considering a CARA design. Further discussion about the properties of RA and CARA designs can be found in, Bandyopadhyay, Biswas, and Bhattacharya (2007), Bandyopadhyay and De (2009) and Hu and Rosenberger (2006) and so on.

Although the sequential characteristics of RA and CARA designs are clear, and the sequential sampling method, which allows the sample size to be determined based on the observed information, is known to be an adequate choice for making efficient and valid statistical inference, most discussions in the literature to date have been limited to the asymptotic properties of different designs. Even when the idea of a stopping rule has been adopted, there has still been very little discussion of estimation under those stopping criteria. Zhang and Hu (2009) and Bandyopadhyay and De (2009) are two typical examples. The former study presents theoretical results on asymptotics of CARA designs, and the latter study conducted only large-scale simulation studies to compare the properties of their designs and to provide information regarding suitable sample sizes for their designs. In another example, Moler, Plo, and Miguel (2006) treated the allocation ruled by an urn model as a Robbins–Monro scheme, but the property of the stopping rule was still ignored. In addition, Thall and Wathen (2005) compared the CARA design to the balanced randomization design, however, the same stopping rule based on the balanced randomized design was applied to both designs, which is inappropriate as indicated in their paper. The sequential method is a natural choice for a CARA design based clinical trial (Hu & Rosenberger, 2006); however, it is rare to find literature regarding the application of stopping rules for the sequential estimation procedure based on CARA designs. In this paper, a sequential procedure is proposed for estimating treatment effect under a general CARA design. Our goal is to estimate the treatment effects, with the minimum sample size, such that the estimates satisfy a prescribed precision, and subjects can be allocated to the superior treatment without interfering with the quality and efficiency of estimation of treatment effects. The asymptotic properties of sequential estimates are obtained under this general CARA design. In addition, we also show that the allocation rule, under the proposed stopping criterion, maintains the same asymptotic properties as those obtained in its non-sequential counterpart. In our numerical study, for illustration purposes, we adopt the method of Bandyopadhyay et al. (2007) and use a utility function to balance the ethical consideration and the efficiency of the estimate for treatment allocation. We then modify the utility function to vary the tuning parameters sequentially, depending on the precision of the estimate at every allocation stage, such that subjects are allocated to a "more adequate" treatment.

The rest of this paper is organized as follows: A sequential estimation procedure for treatment effect is proposed in Section 2. Simulation results are applied to logistic models using a modified allocation rule (Bandyopadhyay et al., 2007) in Section 3. We then conclude with discussion in Section 4. Proofs of theorems are given in the Appendix.

2. Sequential estimation for CARA designs

Let $N_{m,k}$ be the number of subjects assigned to treatment k during the first m assignments and $\mathbf{N}_m = (N_{m,1}, \ldots, N_{m,K})$. Suppose that $\{Y_{m,k}, m = 1, 2, \ldots, k = 1, \ldots, K\}$ denotes responses of the m-th subject to the k-th treatment and $\mathbf{Y}_m = (Y_{m,1}, \ldots, Y_{m,K})$. Let $\boldsymbol{\xi}_m$ be the covariates of the m-th subject. Suppose that $\mathbf{X}_1, \mathbf{X}_2, \ldots$ is the sequence of random treatment assignments and $\mathbf{X}_m = (X_{m,1}, \ldots, X_{m,K})$, $X_{m,k} \in \{0, 1\}$, denotes assignment of treatment k to the m-th subject. Then $X_{m,k} = 1$ for some k and $\sum_{k=1}^{K} X_{m,k} = 1$. That is, each subject is allocated to one treatment only. Hence, it follows that the response of subject m to the treatment k, $Y_{m,k}$, is observed only if $X_{m,k} = 1$. (Note that this implies that $\mathbf{N}_m = \sum_{i=1}^{m} X_i$.)

Define $\mathcal{X}_m = \sigma(\mathbf{X}_1, \ldots, \mathbf{X}_m)$, $\mathcal{Y}_m = \sigma(\mathbf{Y}_1, \ldots, \mathbf{Y}_m)$, and $\mathcal{Z}_m = \sigma(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_m)$, $\boldsymbol{\xi}_i \in \mathbb{R}^p$, be the corresponding σ-fields. Let $\mathcal{F}_m = \sigma(\mathcal{X}_m, \mathcal{Y}_m, \mathcal{Z}_m)$, then a general CARA design is defined as

$$\boldsymbol{\psi}_m = E[\mathbf{X}_m | \mathcal{F}_{m-1}, \boldsymbol{\xi}_m] = E[\mathbf{X}_m | \mathcal{X}_{m-1}, \mathcal{Y}_{m-1}, \mathcal{Z}_m],$$

where $\boldsymbol{\psi}_m$ is actually a vector of randomization probabilities for treatments $1, \ldots, K$. Suppose that, for each $m \geq 1$, the responses and covariate vector satisfy

$$E[Y_{m,k} | \boldsymbol{\xi}] = \mu_k(\theta_k, \boldsymbol{\xi}), \tag{1}$$

where $\mu_k(\cdot, \cdot)$ are known functions, V_k denotes the covariance matrix based on Eq. (1) and $\theta_k \in R^p$ for $k = 1, \ldots, K$. The asymptotic properties of the estimate of $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$ and allocation function under such a general CARA design has been discussed in Zhang, Hu, Cheung, and Chan (2007). The estimation of θ is the primary goal in a clinical trial. Thus, it will be beneficial if treatment effects can be estimated with a certain accuracy using a minimum required sample size whilst simultaneously still retaining the good allocation properties. Since, in a CARA design, the design at the current stage depends on the past history, sequential analysis is the statistical tool of choice. Here, a sequential estimation procedure is proposed for constructing a confidence set for θ with a prescribed accuracy, and we show that the asymptotic properties of allocation function remain the same as their non-sequential counterparts under such a sequential sampling strategy.

2.1. Sequential estimation of treatment effects

Suppose no prior information about the effects of treatments is available. In order to estimate the treatment effects, at the beginning, we need to assign $m_0 (> 0)$ subjects to each treatment using restricted randomization. Hence, when we allocate