Journal of Statistical Planning and Inference 170 (2016) 144–157

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

A-optimal and A-efficient designs for discrete choice experiments

Fangfang Sun[a,*], Angela Dean[b]

[a] Department of Management Science and Engineering, Harbin Institute of Technology, 92 West Dazhi Street, Harbin, 150001, China
[b] Department of Statistics, The Ohio State University, 1958 Neil Avenue, Columbus, OH 43210, USA

ARTICLE INFO

Article history:
Received 11 May 2015
Accepted 16 September 2015
Available online 22 October 2015

MSC:
62K05
62K10
62K15

Keywords:
A-optimal
Efficient designs
Local optimality
Multinomial logit model
Robustness
Variance bound

ABSTRACT

Discrete choice experiments are widely used in fields such as marketing, planning, transportation, and medical care to obtain information on consumer preferences. In such experiments, choice sets consisting of two or more profiles are presented to subjects, where a profile consists of a set of attributes (as a list or picture) which describe the product or process. Subjects are asked to select their most preferred profile from each choice set, and the importance of the attributes can be deduced from the choices made.

This paper investigates locally A-optimal designs for estimating main effects of the attributes, together with their interactions, under the multinomial logit model. Lower bounds are derived for the average variance of any set of orthonormal contrasts of interest. A new approach is proposed for generating locally A-optimal or A-efficient designs. It is shown through examples that the new construction method enables highly efficient designs to be constructed without a complete search.

1. Introduction

To investigate a target population's preference on, for example, marketing products, transportation or medical plans, *discrete choice* experiments (also known as stated choice or forced choice experiments) have frequently been used. In many such studies, a fixed set of n choice sets is presented to each subject, where every choice set contains two or more *profiles*. Each profile consists of *a attributes* which describe the product, and the subject is asked to select his or her most preferred profile from each choice set. Street and Burgess (2007) give a range of such examples from the medical field; Liu et al. (2009) give an example of credit card choice with the three attributes APR, issuing bank, and rewards; and Li et al. (2013) give an example of choice of hearing aids with attributes ear position, quality and price.

The outcomes of such studies are discrete and various models have been proposed, (see, for example, Train, 2003; Ruan et al., 2008; Bliemer and Rose, 2010; Fiebig et al., 2010; Goos et al., 2010; Bush et al., 2010; Crabbe et al., 2013; Lancsar et al., 2013; Großmann and Schwabe, 2015). The multinomial logit model (MNL model) has frequently been used, and is the model assumed in this paper. Since it is a non-linear model, the variance–covariance matrix of the maximum likelihood estimator of the unknown parameters is a function of those same parameters and, consequently, locally optimal designs or, alternatively, Bayesian designs need to be considered. We take the former approach in this paper.

---

[*] Corresponding author.
E-mail addresses: fangfang@hit.edu.cn (F. Sun), dean.9@osu.edu (A. Dean).

http://dx.doi.org/10.1016/j.jspi.2015.09.007
0378-3758/© 2015 Elsevier B.V. All rights reserved.

One popular approach for designing a choice experiment is to obtain the optimal design under the assumption that all profiles in the design are equally attractive; this is sometimes called the *utility neutral* case. In this setting, several authors have provided results for locally D-optimal designs as well as a few results for A-optimal designs, (see, for example, Street and Burgess, 2007, 2012). In general, in the utility neutral case, Graßhoff and Schwabe (2008) show via a linearization of the MNL model that optimality results under the corresponding linear model are relevant.

In the unequal attractiveness (unequal utility) setting, work on locally optimal designs includes that of Huber and Zwerina (1996) who constructed A- and D-efficient "utility balanced" designs through swapping and relabeling attribute levels within choice sets. Graßhoff and Schwabe (2008) showed that optimal designs depend strongly on the values of the unknown parameters. They found locally D-optimal designs over almost the entire parameter space for choice sets of two profiles and two binary attributes. Due to the complexity of their method, it is almost impossible to achieve solutions for more complicated settings (see the discussion in Graßhoff and Schwabe, 2008). A Bayesian approach, where a prior distribution for the unknown parameters is assumed, has been used for unequal utilities under the MNL model by, for example, Sandor and Wedel (2001); Kessels et al. (2006, 2008, 2009).

In this paper, we focus on locally A-optimal designs for discrete choice experiments under the MNL model. The model and information matrix are described in Section 2. Section 3 provides lower bounds for the sum (and hence average) of the variances of any set of orthonormal contrasts of interest, and conditions are given under which the bound can be attained and local A-optimality deduced. A specific form of the bound is given for the utility neutral case which can be used to provide a measure of efficiency. For the case when these bounds cannot be attained or cannot be verified, a linearization of the model (cf. Graßhoff et al., 2013) allows a construction method to be proposed in Section 4 that uses the choice sets that make the greatest "contributions" to (i.e. provide the most information on) the contrasts of interest. Using these ideas, a new sequential approach for finding A-efficient designs quickly without conducting a complete search is illustrated in Section 5 and it is shown, through an example, that the proposed method can identify designs with high A-efficiency for most of the parameter space. Finally, the same example is used in Section 6 to investigate the efficiency of a utility neutral design to alternative local values of the utilities.

As in Street and Burgess (2004); Street et al. (2005), and Graßhoff and Schwabe (2008), and others, we assume that subjects choose one profile from each of n choice sets, all subjects evaluate the same group of choice sets, and that individual subject effects are not included in the model. For simplicity, we limit our discussions to designs with the following properties: (i) every profile has a 2-level attributes; (ii) each of the $v = 2^a$ profiles is included in the experiment; (iii) every choice set is of size m; and (iv) a choice set is never shown more than once to a subject. Experiments involving attributes at more than 2 levels will be discussed in future work as will designs which do not include all profiles, (also see Sun, 2012). The latter requires careful definition of the contrasts of interest.

2. Model and information matrix

Let the v profiles in an experiment be labeled $P_1, P_2, \ldots, P_v$, where $P_\ell = (\ell_1, \ldots, \ell_a)$ and ℓ_u is the level of the uth attribute in profile P_ℓ $(1 \le \ell_u \le L_u)$, and L_u is the total number of levels of the uth attribute in the experiment. Suppose that n distinct choice sets, each consisting of $m(\le v)$ profiles, are included in the design. For $i = 1, \ldots, \binom{v}{m}$, we may write choice set C_i as $\{P_\ell : \ell \in S(i)\}$ where $S(i)$ contains the subscripts of the m profiles which are in choice set C_i. For example, $S(1) = \{1, 3, 4\}$ indicates that choice set C_1 is of size $m = 3$ and consists of profiles (P_1, P_3, P_4).

Following the development of the multinomial logit (MNL) model as in Train (2003), Chapter 3, and Street and Burgess (2007), Chapter 3, we assume that, when presented with a set of profiles, subjects choose the profile that has the *maximum utility* to them. Suppose $U_{\ell\alpha}$ is the utility assigned by subject α, to profile P_ℓ, $\ell = 1, 2, \ldots, v$, and that subject α chooses profile P_j from choice set $C_i = \{P_\ell : \ell \in S(i)\}$ if

$$U_{j\alpha} > U_{\ell\alpha}, \quad \forall \ell \ne j \in S(i).$$

Let $U_{j\alpha} = \gamma_{j\alpha} + \epsilon_{j\alpha}$, where $\gamma_{j\alpha}$ is the systematic component of the utility of profile P_j that can be captured and measured for subject α, and $\epsilon_{j\alpha}$ is the unobserved random component. Since, in this paper, every subject receives the same set of choice sets and subject effects are not included in the model, the optimal design is the same for each subject and we may drop the subscript α. If the ϵ_j $(j = 1, \ldots, v)$ are independently distributed with identical extreme value type 1 (Gumbel) distributions, then the MNL model results (cf. Train, 2003, Section 3.1). Under this model, the probability, p_j, of profile P_j being chosen from choice set $C_i = \{P_\ell : \ell \in S(i)\}$ is

$$p_j = P(U_j > U_\ell, \ \forall \ell \ne j \in S(i)) = \frac{e^{\gamma_j}}{\sum\limits_{\ell \in S(i)} e^{\gamma_\ell}}. \tag{2.1}$$

To ensure identifiability, a normalizing constraint is imposed on the v systematic components so that $\sum_{k=1}^{v} \gamma_k = 0$, where v is the total number of profiles in the entire design. If all profiles are equally attractive, that is $\gamma_1 = \cdots = \gamma_v$ $(= 0)$, then all profiles in choice set $C_i = (P_{i_1}, \ldots, P_{i_m})$ have an equal chance of being chosen with $p_{i_1} = \cdots = p_{i_m} = 1/m$.

Let $\omega_i = 1/n$ if choice set i is in the design, and $\omega_i = 0$ if it is not in the design. Then, as in Street and Burgess (2007), the information matrix $I(\boldsymbol{\gamma})$ (information per choice set) for estimating $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_v)$ for a design with n distinct choice sets