The Journal of Choice Modelling 21 (2016) 10–14

Contents lists available at ScienceDirect

The Journal of Choice Modelling

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

On determining priors for the generation of efficient stated choice experimental designs

Michiel C.J. Bliemer, Andrew T. Collins *

Institute of Transport and Logistics Studies, The University of Sydney Business School, The University of Sydney, NSW 2006, Australia

ARTICLE INFO

Article history:
Received 7 October 2015
Received in revised form
10 March 2016
Accepted 16 March 2016
Available online 6 April 2016

Keywords:
Stated choice
Experimental design
Efficient designs
Bayesian priors

ABSTRACT

Bayesian priors are required in order to generate efficient and robust experimental designs for stated choice surveys. Although such priors are commonly obtained through a pilot study, in this paper we provide a simple alternative in which the analyst depends only on their own expert judgement and possibly on parameter estimates obtained from the literature. The process consists of ranking attribute levels, balancing choice tasks to obtain trade-offs, and setting probabilities in sample choice tasks to establish scale.

1. Introduction

Consider designing an experiment for estimating a typical multinomial logit model. We assume that the utility function for alternative j in choice task s for respondent n is given by

$$U_{nsj} = V_{nsj} + \epsilon_{nsj}, \text{ with } V_{nsj} = \sum_{k=1}^{K} \beta_k x_{nsjk}, \text{ for all } n = 1, \ldots, N; \ s = 1, \ldots, S; \ j = 1, \ldots, J, \tag{1}$$

where V_{nsj} is the systematic utility function given by a function that is linear in the unknown parameters $\boldsymbol{\beta} = [\beta_k]$ and linear in the attributes k defined by attribute levels x_{nsjk}, and unobserved random components ε_{nsj} are identically and independently extreme value type I distributed with variance $\frac{1}{6}\pi^2\lambda^{-2}$, where λ is a positive scale parameter. In this paper we assume that all parameters are generic across all alternatives, i.e. β_k is present in, and does not vary across, the utility functions for all J alternatives.

Let y_{nsj} be the choice indicator that equals one if respondent n chooses alternative j in choice task s, and zero otherwise. The probability that the respondent chooses alternative j is given by the conditional logit model (McFadden, 1974):

$$P_{nsj} = \Pr(y_{nsj} = 1) = \frac{\exp(\lambda V_{nsj})}{\sum_{i=1}^{J} \exp(\lambda V_{nsi})}. \tag{2}$$

Due to identification issues, it is not possible to estimate λ and $\boldsymbol{\beta}$ separately, hence it is only possible to estimate scaled parameters $\boldsymbol{\beta}^* = \lambda\boldsymbol{\beta}$. In the special case of choice tasks with two alternatives (i.e., $J = 2$), the probability of choosing

* Corresponding author.
E-mail addresses: michiel.bliemer@sydney.edu.au (M.C.J. Bliemer), andrew.collins@sydney.edu.au (A.T. Collins).

http://dx.doi.org/10.1016/j.jocm.2016.03.001
1755-5345/

alternative j instead of i can be written as

$$P_{nsj} = \frac{1}{1 + \exp\big(\lambda(V_{nsi} - V_{nsj})\big)}. \tag{3}$$

In order to estimate unknown scaled parameters $\boldsymbol{\beta}^*$, the analyst can conduct a stated choice survey by creating hypothetical choice tasks defined by specific combinations of profiles. A profile describes the attribute levels of a single alternative j, which is defined by the vector of attribute levels $\mathbf{x}_{nsj} = (x_{nsj1}, \ldots, x_{nsjK})$. Hence, each choice task s for respondent n is described by a combination of profiles, $(\mathbf{x}_{ns1}, \ldots, \mathbf{x}_{nsJ})$.

In order to increase the reliability of the parameter estimates or reduce the necessary sample size N, analysts often resort to efficient designs that maximise Fisher information, which is defined by $\mathbf{I} = \mathbf{Z}'\mathbf{Z}$, where elements in the $NSJ \times K$ matrix $\mathbf{Z} = [z_{nsjk}]$ are given by (Huber and Zwerina, 1996):

$$z_{nsjk} = \left(x_{nsjk} - \sum_{i=1}^{J} P_{nsi} x_{nsi} \right) \sqrt{P_{nsj}}. \tag{4}$$

Since the Fisher information matrix depends on the probabilities, it can only be determined once the scaled parameters $\boldsymbol{\beta}^*$ are known, hence the analyst has to rely on (scaled) priors $\tilde{\boldsymbol{\beta}}^* = [\tilde{\beta}_k^*]$ that are best guesses for these unknown parameters, where $\tilde{\boldsymbol{\beta}}^* = \tilde{\lambda}\tilde{\boldsymbol{\beta}}$. While in estimation one does not distinguish between parameters and scale, for determining priors this distinction will turn out to be useful. Since there is a lot of uncertainty about such prior values, it is common to use random prior distributions (commonly referred to as Bayesian priors, see Sándor and Wedel (2002)), typically following a normal distribution, i.e.,

$$\tilde{\beta}_k^* \sim \mathbf{N}(\tilde{\lambda}\mu_k, \tilde{\lambda}\sigma_k), \tag{5}$$

where μ_k and σ_k are the corresponding mean and standard deviation. It is good practice to determine these prior distributions through a pilot study by using an initial experimental design (e.g., random, orthogonal, or with maximised Fisher information assuming $\tilde{\boldsymbol{\beta}} = \mathbf{0}$), with data collected from a small number of initial respondents. After estimating the parameters of the model based on choices in this pilot study, one can set μ_k equal to the respective parameter estimate and set σ_k (denoting the uncertainty) equal to the standard error.

Although pilot studies are strongly recommended for obtaining feedback on the experiment and information on priors, not everyone conducts such pilot studies, for various reasons (e.g., time or budget constraints). Instead, some researchers obtain priors from previous studies described in the literature, or base priors on expert judgement. While 'good' priors can significantly increase the amount of information contained in the data through smart choice tasks with appropriate trade-offs across the attributes, 'bad' priors (often due to inappropriate scaling) may have a detrimental effect on the efficiency of the design (as shown in Section 3). In this paper we provide a quick and easy technique for determining reasonable priors without conducting a pilot study. These priors can also be used for generating an efficient experimental design for a pilot study.

2. Quick and easy priors

In the following subsections we describe the steps for the analyst to obtain priors without conducting a pilot study. Sándor and Wedel (2002) also describe a method for calculating such priors, however they make several restrictive assumptions and consider only effects coded variables. In this paper we use an air travel choice example to illustrate the procedure. In this example, the analyst would like to investigate long haul travel choices (Sydney to London) in which the travel choices consist of different flights identified by three attributes, namely return airfare, duration (from Sydney departure to London arrival, inclusive of flight and stopover time), and the in-flight entertainment system (from now simply referred to as 'Entertainment').

Table 1
Attribute levels and ranking.

Attribute	Levels (from least to most preferred)
Airfare	\$3000, \$1400
Duration	32, 22 h
Entertainment	Shared screen (shared), Personal screen with limited movie selection (personal), Personal screen with video on demand (VOD)