ARTICLE IN PRESS

Journal of Choice Modelling xx (xxxx) xxxx–xxxx

Contents lists available at ScienceDirect

Journal of Choice Modelling

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

Asymmetric triangular mixing densities for mixed logit models

Thijs Dekker

Institute for Transport Studies, University of Leeds, 36-40 University Road, LS2 9JT Leeds, UK

ARTICLE INFO

Keywords:
Mixed logit
Mixing density
Asymmetry
Triangular distribution

ABSTRACT

A novel method is proposed to estimate random parameter logit models using the asymmetric triangular distribution to describe unobserved preference heterogeneity in the population of interest. The asymmetric triangular mixing density has the potential to overcome behavioural limitations associated with the most frequently applied mixing densities like the normal and log-normal distribution. With only three parameters it remains parsimonious whilst its bounded support can easily be brought in line with behavioural intuitions. The triangular mixing density is not associated with an incredibly large upper (or lower) bound and it can accommodate varying degrees of skewness in unobserved preference heterogeneity. The proposed estimation procedure is based on the principle of mixture densities and circumvents additional simulation chatter arising when applying the inverse cumulative density function method to generate draws from the mixing density.

1. Introduction

The mixed logit model (MIXL), also referred to as the random parameters logit (RPL) model, represents one of the most popular econometric models to analyse discrete choice type data. Advantages of the MIXL model include (i) the ability to model heterogeneity in the patterns of choices across respondents:[1] (ii) non-constant error variances across alternatives via a relaxation of the independently and identically distributed error terms assumption; and (iii) the potential accommodation of correlation in choices across repeated choice observations by the same respondent (e.g. Hensher and Greene, 2003; Scarpa et al., 2005). Given its multi-functionality, Keane and Wasi (2013) acknowledge the MIXL model hosts an infinite number of alternative model specifications varying in the number and selection of alternative mixing densities.

The current paper adds the asymmetric triangular density to the set of potential mixing densities available to the analyst. By being able to control for skewness in the distribution of preferences over the population of interest, the asymmetric triangular density is more flexible than its symmetric counterpart. Its bounded support at both ends of the distribution makes it a particularly attractive density relative to the more frequently used (log-)normal density. The asymmetric triangular density thereby answers the call of Hensher and Greene (2003) for the implementation of simple, but flexible distributional forms complying with behavioural expectations. So far this call mainly resulted in the adoption of the (constrained) symmetric triangular distribution (e.g. Brouwer et al., 2010; Hensher and Greene, 2003).

The focus of this paper is on the development of a maximum simulated likelihood (MSL) estimation method for the asymmetric triangular density. The proposed estimation method is based on the principle of mixing densities and recognizes that any triangular

E-mail address: t.dekker@leeds.ac.uk.

[1] Whilst it is common to interpret the random parameter coefficients as representing purely preference heterogeneity, the confoundment between scale and preference parameters in most discrete choice models implies that any modelled heterogeneity should more correctly be interpreted as representing a mixture of both preference and scale or error heterogeneity (Hess and Rose, 2012).

http://dx.doi.org/10.1016/j.jocm.2016.09.006
Received 24 September 2015; Received in revised form 19 August 2016; Accepted 21 September 2016
Available online xxxx
1755-5345/

density can be described by means of two one-sided triangular densities with a common mode. A Bayesian estimation procedure was already developed by Dekker and Rose (2011).

The structure of the paper is as follows. Section 2 introduces the MIXL model and defines the triangular density. Section 3 then develops the MSL estimation framework. Section 4 presents a Monte Carlo simulation and Section 5 concludes the paper.

2. Model structure

2.1. The random coefficients multinomial logit model

Suppose individual n is presented with J alternatives in choice task $t = 1,\dots,T$. The Random Utility Maximisation model postulates that the individual selects the alternative with the highest level of utility, i.e. $y_{nt} = i$ when $U_{nit} > U_{njt},\ \forall j \neq i \in J$. Utility U_{nit} is decomposed in a structural part V_{nit} and a stochastic part ϵ_{nit}, where $U_{nit} = V_{nit} + \epsilon_{nit}$. After assuming that ϵ_{nit} follows a type-I extreme value distribution the choice probability for alternative i can be described by:

$$P_{nit} = \frac{\exp(V_{nit})}{\sum_{j=1}^{J}\exp(V_{njt})} \tag{1}$$

V_{nit} is characterised by a linear utility function $V_{nit} = X_{nit}\beta_n$. Let X_{nit} represent a set of exogenous variables and β_n defines the vector of marginal utility parameters. The subscript in β_n denotes marginal utility that may vary across respondents. In most applications, an insufficient number of observations per respondent is available to estimate individual specific utility parameters. Hence, random coefficients are used to capture the heterogeneity in β_n across the population of interest. Let $f(\beta_n|\Omega)$ denote a mixing density function describing the distribution of marginal utility over the population of interest, where Ω is the vector of associated hyper-parameters. The expected choice probability of observing the sequence of choices y_n can then be described by the individual specific likelihood L_n:

$$L_n = \int_{\beta_n}\prod_{t=1}^{T}\frac{\exp(X_{nit}\beta_n)}{\sum_{j=1}^{J}\exp(X_{njt}\beta_n)}f(\beta_n|\Omega)d\beta_n \tag{2}$$

2.2. The triangular distribution

In this paper, $f(\beta_n|\Omega)$ is described by a triangular density. The density is a function of only three hyper-parameters being respectively the lower-bound a, the upper-bound b and the mode c. These three hyper-parameters define the density function:

$$f(\beta_n|a,b,c) = \frac{2(\beta_n - a)}{(b-a)(c-a)} \text{ for } a \le \beta_n \le c \qquad \frac{2(b-\beta_n)}{(b-a)(b-c)} \text{ for } c \le \beta_n \le b \tag{3}$$

When $\beta_n < a$ or $\beta_n > b$ the density $f(\cdot)$ will be zero. In short, the triangular distribution qualifies as a mixing density that is simple but flexible in shape and easily complies with behavioural expectations. Namely, the flexible mode of the triangular distribution allows for both positively- and negatively-skewed distributions, but also symmetry by setting $(c-a) = (b-c)$.[2] The support of the distribution can be constrained by fixing either the lower- or the upper-bound or both. Accordingly, the triangular distribution can accommodate non-negative (or non-positive) marginal utilities without inducing a fat upper-tail.

3. Maximum simulated likelihood estimation

3.1. The inverse cdf problem

The principles of MSL require that the simulated density can be obtained by means of rescaling and relocating a standard shape of the underlying distribution. For example, a normal distribution can be simulated by taking draws from a standard normal distribution, which is subsequently relocated by the estimated mean and rescaled by the estimated standard deviation. For the asymmetric triangular density, draws from $f(\beta_n|\Omega)$ can be generated using an inverse cumulative density function (cdf) transformation approach for a given a, b, c (see (4)), where U_n^r represents a draw r from the standard uniform distribution defined over $[0, 1]$ for individual n.

$$\beta_n^r = a + \sqrt{U_n^r(b-a)(c-a)} \text{ for } U_n^r < \frac{c-a}{b-a} \qquad b - \sqrt{(1-U_n^r)(b-a)(b-c)} \text{ for } U_n^r \ge \frac{c-a}{b-a} \tag{4}$$

The described inverse cdf approach, however, introduces additional chatter in the simulation. During each optimization iteration the values for a, b, and c adjust, implying that the number of draws assigned to the first (and second) part of (4) change. The

[2] Note that by drawing a straight line from the density at the mode to the zero density at the bounds the share of the population is decreasing at a constant rate when moving away from the mode.