



# Discrete choice models based on random walks



Gerardo Berbeglia

Melbourne Business School, University of Melbourne, Australia

## ARTICLE INFO

### Article history:

Received 13 May 2015

Received in revised form

20 January 2016

Accepted 20 January 2016

Available online 29 January 2016

### Keywords:

Discrete choice models

Markov chain

Random utility

Revenue management

Stochastic preferences

Random walk

## ABSTRACT

We show that a large class of discrete choice models which contain the Markov chain model introduced by Blanchet, Gallego, and Goyal (2013) belong to the class of discrete choice models based on random utility.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Discrete choice models (DCM) have been widely used to model the choices individuals make when they are offered a set of alternatives (or options). DCM have played an important role in several research areas such as psychology, economics, marketing, and, more recently, in the field of revenue management.

In the standard discrete choice model setting, there is a universe of alternatives or product types  $\mathcal{C} = \{1, \dots, N\}$ . Individuals are then exposed to offer sets  $S \subseteq \mathcal{C}$ , and they must select one element from  $S$  or nothing at all. Each subset  $S$  offered to individuals is known as the choice set. In this setting, a discrete choice model is characterized by a function  $\mathcal{P} : \mathcal{C} \times 2^{\mathcal{C}} \rightarrow \mathbb{R}$  such that for all  $i \in \mathcal{C} \cup \{0\}$  and all  $S \subseteq \mathcal{C}$ ,  $\mathcal{P}(i, S)$  denotes the probability of selecting element  $i$  given that the offer set is  $S$ . The function  $\mathcal{P}$  is often called the system of choice probabilities that characterizes a specific discrete choice model.

The practical advantage of using a particular discrete choice model depends on two characteristics. The first one is the degree upon which it is able to model the choice behaviour of individuals in different settings. The better a discrete choice model can approximate complex choice behaviour, the *richer* is said to be. The second one is by the degrees of freedom it has. A model with a large number of degrees of freedom would typically require a large historical data-set in order to find the right values of its parameters and it is more prone to over-fitting problems. These

two dimensions are usually in conflict, i.e., discrete choice models that can capture complex choice behaviours generally have a higher number of degrees of freedom. Finding the “right” discrete choice model for a given application involves searching a model with a proper trade-off between the two mentioned dimensions (see, e.g., [11]).

One of the most simple and also most used models in discrete choice is the Multinomial Logit Model (MNL) [15]. Because of its simplicity, the MNL fails to predict complex behaviour and that has motivated the creation of richer models such as the Probit Model [1], the Nested Logit Model [14], the Mixed Multinomial Logit [17] and distance-based models [18] which contain the Mallows model [16] as a special case. All these models belong to an important and large class of discrete choice models based on random utility known as *Random Utility Models* (RUMs) which are defined for completeness in Section 2. Recently, Webb [22] showed that a predominant class of discrete choice models to understand neural decision making, known as *Bounded Accumulation models* [20], are as well a subclass of RUMs.

There exist nevertheless classes of discrete choice models that fall outside the class of RUMs. We briefly mention the most relevant ones. Echenique et al. [8] have considered a discrete choice model in which agents select an alternative following a perception priority order. The authors showed that their model, called *Perception-Adjusted Luce Model* (PALM), is able to explain recent experiments carried out with consumers [7] that cannot be explained by any random utility model. Another choice model, initially proposed to understand sales diversity under monopolistic competition [21,6], is the *Representative Agent Model* (RAM). In a

E-mail address: [g.berbeglia@mbs.edu](mailto:g.berbeglia@mbs.edu).

RAM, the choice among the alternatives is decided by a specifically constructed agent that is a representative of the whole population. Hofbauer and Sandholm [12] proved that when the number of alternatives is at least 4, there is always a RAM which is not a RUM. Natarajan et al. [19] have presented a strict generalization of the RUM, known as *Semi-parametric Choice Model* (SCM), and showed how it can be used to make predictions from a real-life transportation data set that outperform several RUM models. The question of what is the relationship between the classes RAM and SCM was answered only recently by Feng et al. [10]. The authors provide a detailed classification of several discrete choice models and proposed a new class which they called welfare-based choice models. Their main contribution is a proof that the RAM, the SCM, and their new proposed class are essentially the same. (It follows from their result that welfare-based choice models include RUMs as a proper subset.)

Recently, another class of discrete choice models have been proposed by Blanchet et al. [3]. In their model, individual preferences are built using a Markov chain in which states are the alternatives or products. The model prescribes how would an individual decide which alternative to select when she is offered a subset of alternatives  $S$ . Specifically, with probability  $p_i$  an individual has alternative  $i$  as her most favourite. If the alternative  $i$  is not available, with probability  $p_{ij}$  the individual who preferred  $i$  walks to select the alternative  $j$ . If again, the alternative  $j$  is not available, with probability  $p_{jk}$  she walks to alternative  $k$ , etc. Since it is assumed that for all  $i \in \mathcal{C}$ ,  $\sum_{j \in \mathcal{C} \setminus \{i\}} p_{ij} < 1$ , this walking process will end either when the individual walks to the no-choice option (in which case selects nothing) or when she arrives to an alternative from the set  $S$  (and selects it). Blanchet et al. [3] showed how this Markov chain model is an appealing model in the area of revenue management, which consists of a set of methodologies firm use to decide on the availability and/or the price of their products and/or services. First, they showed that the model generalizes the widely used MNL, and second, the assortment problem under this new model can be solved efficiently. The assortment problem under the Markov chain model was recently studied in more depth by Feldman and Topaloglu [9] and Désir et al. [5] where the authors analysed different extensions such as when there is limited inventory, or when there is an upper bound on the size of the assortments that can be shown to individuals.

### 1.1. Our results

Blanchet et al. [3], Feldman and Topaloglu [9] and Désir et al. [5] left open a fundamental question about the Markov chain model: how rich is this class? More specifically, how does this class compares to RUMs? In their analysis of discrete choice models, Feng et al. [10] argued that the Markov chain model is not related to other discrete choice models they studied such as the RUMs, SCM and RAM, and they omitted the Markov model from their analysis and classification. (To the best of our knowledge, the Markov chain model is the only discrete choice model studied in the literature whose connection with the RUM remained open.) In this paper, we prove that every Markov chain model belongs to the class of RUMs. Thus, despite their appealing features, Markov chain models are not able to replicate every choice behaviour that can be modelled by other discrete choice models such as the PALM and the RAM (or equivalently, the SCM or the welfare-based models). On the positive side, it follows from our result that the three performance guarantees for the simple and well-known revenue-ordered assortment heuristic for the assortment problem that hold for all RUMs [2] are valid as well for every Markov chain model.

A natural way to enrich the class of Markov models is by providing more ‘memory’ to the individual along her walk over the

alternatives. Suppose now that the probability that an individual who is currently in alternative  $i$  would walk to the alternative  $j$  depends not only on  $i$  and  $j$ , but also on the alternative that was visited right before alternative  $i$ . In other words, the individual now considers the previous two alternatives (instead of only the last one) in order to probabilistically walk to the following alternative. Is it possible that this model extension becomes richer than RUMs? If not, what about the class of models where individuals remember the last three, or the last  $k$  alternatives with  $k \in \mathbb{N}$ ? We answer this question negatively. Specifically, we introduce, in Section 3, a new class of discrete choice models called *Discrete choice Models Based on Random Walks* where the walking probabilities depend on the whole sequence of the alternatives previously visited. We then prove that for every such a model, there always exists a random utility model that is equivalent to it. In other words, regardless on how much memory the individuals are endowed with about the different alternatives that they walked through (this included arbitrarily large memory as sequences’s length are unbounded), the model will fail to explain behaviours that cannot be explained by RUMs.

## 2. Preliminaries

In a Random Utility Model, each alternative  $i$  (including the no-choice option) has associated a random real variable (utility)  $u_i$ . These  $N + 1$  variables are jointly distributed over  $\mathbb{R}^{N+1}$ , with a certain probability measure  $\mathbb{P}$  with  $\mathbb{P}(U_i = U_j) = 0$  for all  $i, j \in \mathcal{C} \cup \{0\}$ ,  $i \neq j$ . Then, the probability of selecting alternative  $x \in S \cup \{0\}$  with  $S \subseteq \{1, 2, \dots, N\}$  is equal to the probability that alternative  $x$  has the highest utility among those in  $S \cup \{0\}$ . The system of choice probabilities is then characterized as follows:

$$\mathcal{P}(x, S) = \mathbb{P}(U_x = \max\{U_c : c \in (S \cup \{0\})\}).$$

An alternative way to think about random utility models is using what is known as *stochastic preference*. Given  $k \in \mathbb{N}$ , let  $S_k$  denote the symmetric group which is the set of all the permutations of the elements in  $\{0, 1, 2, \dots, k - 1\}$ . A stochastic preference model is described by a probability distribution  $Pr$  over the set  $S_{N+1}$  (i.e. a probability distribution over all the rankings of the elements in  $\mathcal{C} \cup \{0\}$ ). Given two elements  $i, j \in \mathcal{C} \cup \{0\}$  and a permutation  $\preceq \in S_{N+1}$ , we say that  $i \preceq j$  whenever  $i$  appears before  $j$  in  $\preceq$  or when  $i = j$ . A system of choice probabilities  $\mathcal{P}$  characterizes a stochastic preference  $Pr$  if

$$\mathcal{P}(x, S) = Pr(\preceq \in S_{N+1} : x \preceq y, \text{ for all } y \in (S \cup \{0\})) \quad (1)$$

for all  $S \subseteq \mathcal{C}$  and  $x \in S \cup \{0\}$ .

Thus  $\mathcal{P}(x, S)$  represents the probability of selecting a ranking from the probability space  $Pr$  in which element  $x$  appears before the no-choice option 0, and any other element in  $S$ .

The equivalence of stochastic preference models and random utility models is expressed in the following theorem.

**Theorem 1.** *A system of choice probabilities  $\mathcal{P}$  characterizes a stochastic preference model if and only if it characterizes a RUM.*

For a proof of this result see, e.g., [4,13].

## 3. Discrete choice models based on random walks

Let  $\Omega$  denote the set of all finite sequences of elements in  $\mathcal{C}$  (including the empty sequence  $s = \emptyset$ ). A discrete choice model based on random walks can be defined based on a series of probability distribution functions  $P_\alpha : \mathcal{C} \cup \{0\} \rightarrow [0, 1]$ , one for each  $\alpha \in \Omega$ . We begin by describing how a random walk is constructed given the PDF’s  $P_\alpha$  and a choice set  $S$ .

Download English Version:

<https://daneshyari.com/en/article/1142223>

Download Persian Version:

<https://daneshyari.com/article/1142223>

[Daneshyari.com](https://daneshyari.com)