Journal of Mathematical Psychology 82 (2018) 56–72

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

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

A hierarchical Bayesian approach for examining heterogeneity in choice decisions

Sunghoon Kim [a,*], Wayne S. DeSarbo [b], Duncan K.H. Fong [b]

[a] Department of Marketing, W.P. Carey School of Business, Arizona State University, Tempe, AZ, 85258, United States
[b] Department of Marketing, Smeal College of Business, Pennsylvania State University, University Park, PA, 16802, United States

HIGHLIGHTS

- A new hierarchical Bayes multivariate probit mixture model incorporating variable selection.
- Heterogeneity of the important features that drive consumer choices.
- A consumer psychology application involving the consideration of Sports Utility vehicles.
- Favorable methodological comparisons with a variety of alternative benchmark models.

ARTICLE INFO

Article history:
Received 9 April 2017
Received in revised form 22 October 2017

Keywords:
Finite mixtures
Variable Selection
Bayesian multivariate probit models
Consumer psychology
Choice heterogeneity

ABSTRACT

There is a vast behavioral decision theory literature that suggests different individuals may utilize and/or weigh different attributes of an object to form the basis of their opinions, attitudes, choices, and/or evaluations of such stimuli. This heterogeneity of information utilization and importance can be due to several different factors such as differing goals, level of expertise, contextual factors, knowledge accessibility, time pressure, involvement, mood states, task complexity, communication or influence of relevant others, etc. This phenomenon is particularly pertinent to the evaluation of stimuli involving large numbers of underlying attributes or features. We propose a new hierarchical Bayesian multivariate probit mixture model with variable selection accommodating such forms of choice heterogeneity. Based on a Monte Carlo simulation study, we demonstrate that the proposed model can successfully recover true parameters in a robust manner. Next, we provide a consumer psychology application involving consideration to buy choices for intended consumers of large Sports Utility Vehicles. The application illustrates that the proposed model outperforms several comparison benchmark choice models with respect to face validity and choice predictive validation performance.

1. Introduction

There is a plethora of behavioral decision theory research in psychology that suggests that individuals may utilize and/or weigh different attributes of stimuli to form their opinions, attitudes, and/or evaluations (e.g., preference or choice). As recently suggested by Park, Rajagopel, Dillon, and Chaiy (2017), this heterogeneity of information utilization, importance, and process can be due to several different factors including individuals' differing goals (Payne, Bettman, & Johnson, 1993), individual differences involving prior knowledge and/or level of expertise (Alba & Hutchinson, 1987), contextual factors (Hutchinson & Alba, 1991), knowledge accessibility (Feldman & Lynch, 1988), time pressure (Wright & Weitz, 1977), level of involvement (Petty & Cacioppo, 2012), mood states (Isen, 1993), task complexity (DeSarbo, Fong, Liechty, & Coupland, 2005), etc. In consumer psychology, many researchers such as Payne, Bettman, and Luce (1996) have summarized several well documented non-compensatory decision heuristics (e.g., elimination by aspects, lexicographic rules, satisficing strategies, etc.) that permit consumers to simplify their complicated decision tasks by selectively processing limited attribute/feature information especially in consumptive situations where large numbers of attributes or features are involved. Related to the choice consideration-set heuristics, advances in recent methods (e.g., Bayesian inferences, machine-learning, etc.) help to understand a variety of heuristics including applications in modern complex products with large number of features (Hauser, 2014). Kamakura, Kim, and Lee (1996) identify both preference and structural heterogeneity in choice which purports potential heterogeneity in both the parameters of the choice model being

* Corresponding author.
E-mail addresses: skim348@asu.edu (S. Kim), wsd6@psu.edu (W.S. DeSarbo), i2v@psu.edu (D.K.H. Fong).

https://doi.org/10.1016/j.jmp.2017.11.002

estimated, as well as potential heterogeneity in the choice process undertaken. Okada and Lee (2016) and Park, DeSarbo, and Liechty (2008) have examined both forms of such heterogeneity in various types of multidimensional scaling models.

Behavioral research concerning individuals' goals has found that decision makers with accuracy goals undertake more extensive processing of attribute information than decision makers with effort minimization goals (Payne et al., 1993); thus, individuals with the goal of effort minimization are likely to utilize fewer attributes than individuals with the goal of accuracy in decision-making. Similarly, individuals with justification goals have been shown to focus on attributes that help them justify their final choices (Kunda, 1990). Hence, the type of goal can be a very important determinant of the number and type of attributes utilized during decision making and choice.

In a similar vein, the existent research on expertise has shown that experts possess greater and more detailed knowledge structures about categories than novices (Johnson & Mervis, 1997), are able to recall more dimensions/attributes about alternatives than novices (Vicente & Wang, 1998), and are able to make more accurate attribute-benefits linkages than novices (Dellaert & Stremersch, 2005). This suggests that when decision making is memory-based, experts will utilize more attributes than novices, and are more likely to focus on important and relevant attributes; likewise, novices are likely to rely on more salient and prototypical attributes/dimensions (Alba & Hutchinson, 1987). Cowley and Mitchell (2003) found that novices who were exposed to information about a consumer product in the context of a specific use situation were less able to recall the product in the context of a different use situation; thus, their usage of product information was related to use situation. However, experts were better able to reorganize product information in memory and recall product information that was appropriate for new use situations (Hutchinson & Alba, 1991; Miller & Ginter, 1979). This also suggests the possibility of different choice processes being employed to make choice decisions.

Apart from goals and expertise/knowledge, contextual variables such as time pressure, mood state, and involvement can also affect attribute/feature information selection and utilization. For example, under moderate time pressure, individuals are likely to process each stimulus alternative separately; while under severe time pressure, individuals have been shown to switch to select a few important attributes and evaluate alternatives based on this restricted set of dimensions (Payne et al., 1996). Houston and Sherman (1995) found that the starting alternative in a choice process determined the type of attribute that received greater weight during choice. Thus, features shared by the choice alternatives were canceled and greater weight was placed on the unique features of the alternative that was the starting point for comparison. Since the starting alternative in a choice set is likely to be different for different individuals, different attributes would emerge as unique versus common resulting in different attributes being weighed differently during the choice process. Involvement with the stimulus category or the decision has also been shown to have a significant effect on various decision-making processes and information processing. Individuals with higher levels of involvement with the decision or the object have been shown to pay greater attention to the decision-making task, and process more information than individuals with lower levels of involvement. Highly involved decision makers have also been shown to focus on relevant aspects of the choice task as compared to subjects with low involvement levels who tend to focus on peripheral aspects of the choice task (Petty & Cacioppo, 2012; Petty, Cacioppo, & Goldman, 1981; Sujan, 1985). Finally, research on positive affect has shown that people in a positive mood are cognitively more flexible than people in negative or neutral moods, and have been shown to be able to utilize more attributes and broader dimensions during decision making (e.g., Isen, 1993; Isen, Daubman, & Nowicki, 1987).

Thus, there can be substantial heterogeneity in the way different subjects derive their preferences, choices, and/or decisions in terms of the different types of attribute information focused upon, as well as the choice rules employed. This heterogeneity becomes particularly relevant when dealing with their evaluations of stimuli defined on many attributes or features and with subjects with differing level of familiarity or expertise with the study stimuli. Let's now explore how this heterogeneity in information utilization manifests itself in a typical choice setting. Consider a binary choice case where each subject must choose between selecting or not-selecting a certain stimulus (in our consumer psychology application, we use intended consumers' decisions to consider buying or not buying brands in a designated product class) described by different combinations of P known attributes or features; and, also assume each subject repeat this choice for M different alternatives. Let $\underline{C}_i = (C_{i1}, \ldots, C_{iM})^T$ denote the M choices made by the ith subject which depends upon stimulus features via the following individual-level generalized linear model:

$$\underline{C}_i = g\left(\underline{X}_i \underline{\beta}_i + \underline{\epsilon}_i\right), i = 1, \ldots, n,$$

where $\underline{\epsilon}_i$ denotes the error term and $\underline{X}_i$ denotes the corresponding $M \times P$ attribute matrix. Because the M choices made by each subject are likely to be interrelated, we propose a multivariate probit model framework for this setting. Regarding the multivariate setting, recent empirical findings in psychology demonstrate that previous choice tasks can affect the current choice, indicating inter-dependence across multiple choices (Leong & Hensher, 2012). We employ a finite mixture formulation to parsimoniously reflect subject response (choice) heterogeneity (see Rossi, Allenby, & McCulloch, 2005 for a review of alternative approaches utilizing hierarchical Bayesian formulations of this heterogeneous choice model; see also Wedel et al., 1999 for developments in Marketing, Bhat, 2017 in engineering, Li & Ansari, 2014 in economics, Yang, 2005 in transportation, etc.). As to be developed shortly, we also provide for simultaneous variable selection per derived cluster.

Although finite mixture based (multivariate) choice models have been previously formulated (see Arminger, Clogg, & Cheng, 2000; Bontemps & Toussile, 2013; Wedel & Kamakura, 2000) and applied to several types of applications (e.g., in revealed choice conjoint analysis), there are several potentially problematic issues to resolve. One major challenge concerns the increasing number of stimulus attributes/features encountered in many applications (i.e., large P). For example, many manufacturers are currently adding more and more features into consumer products such as smart phones, digital cameras, flat screen TV's, automobiles, laptop computers, tablets, etc. This trend restricts the use of some traditional methods such as conjoint analysis or revealed preference analysis where it is recommended not to utilize more than six or seven attributes or features (Green & Srinivasan, 1978) to collect such preferences or choices. In responding to this new challenge, researchers in choice modeling have proposed some recent alternatives including hybrid techniques using self-explication (Johnson, 1987) methods that rely on subjects in the experiment to reveal explicitly what the important features are to them. These alternative approaches, however, do not often reflect real-life choice scenarios and may therefore provide results that are incomplete and/or difficult to interpret. Given a long list of attributes/features, subjects usually make their choice decisions only based on a subset of important attributes/features due to convenience, cost of thinking, or lack of expertise about some attributes/features, etc. (Gilbride, Allenby, & Brazell, 2006), in contrast to the assumption held in traditional choice methods that subjects consider every attribute/feature before manifesting their