BaSIC
Bayesian Sensory Model Integrated with Characteristics: A Stochastic MDS Model for Sensory Analysis

in4mation insights, LLC

Introduction

Food technologists and perfumers seek to understand how to combine dozens of ingredients and how the blend of these will affect consumers' taste perceptions and product preferences. Most companies producing foods, beverages, fragrances, and personal care products rely on sensory testing in their product development efforts. Sensory testing generally makes use of a trained expert panel that judges the sensory qualities of the products and on chemometric research which is used to identify the main chemical ingredients. These two areas are assumed to activate the full sensory experience.

Increasingly, sensory research has expanded beyond expert panels and chemical analyses to include the analysis of judgments of products by *consumers* in terms of their preference or liking. When exposed to a product, consumers are believed to form their judgments by weighting together product dimensions that are generally unobservable to the sensory analyst (Coombs, 1964). The task of the analyst is to derive these unobserved sensory dimensions from the data collected and to assess the relative locations of the people and products on the dimensions.

Multidimensional Scaling (MDS) models were developed to analyze such data and are called *unfolding models*. Internal unfolding models are used to simultaneously extract person and product coordinates on the unobserved sensory dimensions from consumer preferences, without any prior knowledge of what these dimensions are. In addition, while product coordinates are estimated from consumer preferences, external information on the products – the sensory and chemometric analyses – are used to aid in the interpretation of the dimensions.

Early MDS models were limited in that their formulation was deterministic; they did not take into account the noise or uncertainty in the data. In the late 1990s, stochastic models were developed with several advantages over deterministic MDS methods:

> *"Maximum likelihood fitting with realistic distributional assumptions for error enables researchers to test hypotheses about dimensionality and other aspects of the complex structures being fit. Researchers can also determine confidence regions for parameter estimates and utilize the rest of the statistical armamentarium necessary to move MDS from the domain of purely exploratory methodology to that of the more confirmatory methodology needed for a maturing marketing science. [...] Ideally, maximum likelihood, with appropriate distributional assumptions, would be used for fitting model(s) to data, making available all the confirmatory statistical tools associated with that approach."* Carroll and Green (1997)

An important extension of deterministic MDS models that was made possible through the stochastic formulation is the extraction of preference segments at the same time that the product dimensions were determined. The *STUNMIX* model of Wedel and DeSarbo (1996) subsumed most of the prior MDS methods and allowed for a wide range of preference tasks, the simultaneous fitting of the chemical and sensory data to the preferences, the identification of consumer preference segments and the profiling of them with demographics.

While stochastic MDS models have had broad appeal, they do have difficulty in accommodating individual-level, rather than segment-level, ideal points and profiling these with consumer demographics. Moreover, they suffer from the limitation of being based on the assumption of the independence of the preference judgments. In particular, two-stage methods that explain product and respondent locations using external descriptor variables have well-documented problems of propagation of estimation errors, which result in a reduced reliability and questionable validity of the estimates.

The BaSIC model

Our solution uses Bayesian estimation methods and is called BaSIC (**Ba**yesian **S**ensory Model **I**ntegrated with **C**haracteristics). The Bayesian statistical framework provides a unique and attractive way of estimating MDS models that addresses many of the challenges of analyzing sensory preference data. In particular, initial uncertainty about the location of people and products on the landscape map is formally represented in the model and that uncertainty is updated with the information that is contained in the data, resulting in a data-based set of estimates. Additionally, while statistical inference for traditional MDS models only applies in large samples, the Bayesian approach facilitates inference in small or large samples alike.

The major advantage of the Bayesian approach is that this approach can estimate complex statistical models which conform to theories of psychological, economic, sensory, and marketing data. Models with many sub-components, including unobserved dimensions, missing data, individual-level parameters, and hierarchical data structures are possible within the Bayesian framework. This integrated estimation procedure eliminates the problems associated with two-stage estimation procedures popular in typical *post hoc* profiling of MDS solutions.

Thus, *BaSIC (Bayesian Sensory Model Integrated with Characteristics)* offers in one integrated framework:

- An unfolding MDS model for various types of consumer preference data;
- A mixture model to estimate the number of, and size of, consumer preference segments;
- The distribution of consumer ideal points within these segments; and,
- Using a hierarchical formulation,
  - Product locations are profiled with sensometric and chemometric variables; and,
  - Respondent-level ideal points are profiled with demographic, behavioral, attitudinal, and other individual-level descriptors.

Importantly, the simultaneous estimation technique employed by BaSIC prevents the propagation of estimation errors common in *post hoc* methods that are often used for landscape analyses and allows for measures of the statistical reliability of all estimates to be easily computed.

Advantages of BaSIC

BaSIC offers the following advantages over previous MDS approaches for sensory analysis.

1. BaSIC employs a stochastic formulation and thus enables statistical testing of various solutions and the assessment of the uncertainty of all estimates.

2. BaSIC accommodates a host of other MDS models as special cases. We can tailor BaSIC to the sensory study in question by including or excluding specific model components. In addition, the fit of a particular model to the data can be easily tested by zeroing out specific components.

3. The Bayesian estimation procedure allows the model with all of its constituent components to be estimated simultaneously, rather than stacking a series of statistical procedures on top of each other. Competitive models use the output of one procedure as the input to the next, which has the drawback of propagation of errors, unreliability and bias in parameter estimates, and unreliability of standard errors of estimates and statistical tests. Such biases are bound to affect decisions that are made based on MDS maps estimated with standard procedures.

4. In BaSIC, the specification of the effect of the sensory panel and chemometric information and the effect of consumer demographics is done in a way that conforms to the hierarchical nature of their effects.

5. BaSIC enables the analyst to address issues of limited data, which occur frequently in sensory tests which use reduced experimental designs to save time or costs. When data are limited, Bayesian estimation borrows strength from other products and people, rendering product- and respondent-specific estimates more stable. In addition, when data are entirely missing, they can be easily imputed as part of the estimation algorithm (rather than using casewise or listwise deletion or mean substitution). In addition, effects of the order in which products are being tested in the experiments can be included in the model and partialled out of the MDS solutions.

6. Finally, BaSIC lends itself to supporting optimal decisions. In sensory analysis, the analyst may want to decide the location of a new product on the map and the required sensometric and chemometric profile to achieve that product. Alternatively, if the analyst wants to consider product modifications that bring an existing product closer to the ideal-point of a segment of respondents or to respondents with a specific demographic profile, this can be accommodated in BaSIC.

References

Carroll, J. D. and P.E. Green (1997). Psychometric Methods in Marketing Research: Part II, Multidimensional Scaling. *Journal of Marketing Research*, 34 (May), 193-204.

Coombs, C.H. (1964). *A Theory of Data*. New York: Wiley.

Wedel, M., and W.S. DeSarbo (1996). An exponential family multidimensional scaling mixture methodology. *Journal of Business and Economic Statistics* 14 (4), 447-459.