



Sawtooth Software

RESEARCH PAPER SERIES

Achieving Individual-Level Predictions from CBC Data: Comparing ICE and Hierarchical Bayes

Joel Huber,
Duke University
1998

Achieving Individual-Level Predictions from CBC Data: Comparing ICE and Hierarchical Bayes

Copyright 1998 Sawtooth Software, Inc.

This article is adapted from a presentation given at the 1998 Advanced Research Techniques Forum by Joel Huber, Duke University. Joel was supported by significant contributions from Rich Johnson (Sawtooth Software), and Neeraj Arora (Virginia Tech).

Choice-Based Conjoint data have traditionally been analyzed in the aggregate. However, several methods have been developed recently that recognize individual differences among respondents, and permit modeling at the segment or individual level. We'll discuss three of these: Hierarchical Bayes, Latent Class and an extended application of Latent Class called ICE (Individual Choice Analysis).

Why should we worry about heterogeneity in respondent preferences? Marketers know that people are unique. Market simulators based on average preferences can result in incorrect managerial decisions. The distortions are particularly apparent in product line applications where highly similar products are expected to take share from each other. An aggregate logit model is particularly poor at reflecting differential substitution effects because it assumes that each alternative takes from all other alternatives in proportion to their market share. The three models we'll discuss have different ways of avoiding this problem.

Defining the Models

Hierarchical Bayes (HB) methods derive individual part worths by combining information on the distribution of part worths across respondents with the specific choices of the individual. The posterior distribution of the individual is estimated through a computationally-intense method called Gibbs Sampling that produces estimates of each respondent's part worths and standard errors.

Latent Class Analysis (LC) assumes that respondents can be clustered into homogeneous segments, and that differences among segments account adequately for underlying differences among individuals. Segment part worths are estimated so as to maximize the likelihood of the respondent data. The probability is also estimated that each respondent belongs to each segment, so one can also estimate expected individual part worths as probability-weighted combinations of the segment part worths.

ICE is a new product from Sawtooth Software that estimates individual part worths from experimental choice data. ICE also estimates part worths as weighted combinations of segment part worths from LC. However, ICE does not constrain those weights to be positive, thereby allowing for much greater differentiation of individual values from the segments.

Strengths and Weaknesses

HB provides very flexible output. Results can be used to estimate ratios and profits. The researcher may choose among many possible population distributions. HB also reveals uncertainty about each respondent's utilities. The principal difficulties of the approach are that it requires a good deal of expertise to execute properly, current software is not very user-friendly, and run times can be very long.

Latent Class has an elegant theoretical foundation and the segment solutions are often managerially useful. One can later estimate each individual's expected part worths as a weighted combination of the various segments' part worths, where the weights are the probabilities of belonging to each segment. However, since those weights are probabilities and therefore positive, individual part worth estimates don't differ as much from one another as the segment part worths do, so they fail to capture the full richness of individual differences. LC is also vulnerable to local optima, so it is prudent to make many runs from different starting points.

ICE is pragmatic: it estimates individual part worths that best fit individual choices individual choices. It is very fast, taking only a few minutes given a LC solution as a starting point. Its main shortcomings are that it lacks the strong theoretical basis of LC and, like LC, results depend on the choice of segments and the number of segments used.

Which Models Work Better in Practice?

We examined three data sets: a simulation study (synthetic data), a laboratory study (MBAs as respondents), and a field study (actual consumers). We measured performance based on correlations with known utilities, correct prediction of first choices, and accurate prediction of aggregate market shares.

Simulation Study: An artificial data set was generated based on known individual part worths and conforming to the typical assumptions of multivariate normal distribution, with response errors having the extreme value distribution. As expected (given that the data set was produced according to its assumptions) HB performed best. Using five LC segments as a basis, ICE did nearly (90%) as well as HB. Latent Class did least well: constraining the weights hurt individual predictions.

Laboratory Study: MBA students completed 30 customized choice tasks, and twelve additional holdout choice tasks. ICE and HB performed equally well in terms of predicting individual holdout choices, and ICE performed slightly better in predicting holdout choice shares. Latent Class was less able to predict individual choices than the other two methods, but performed relatively well at predicting aggregate choice shares (though not as consistently as ICE and HB.)

Field Study: Three-hundred fifty consumers were interviewed at a mall intercept. Respondents completed 18 choice tasks, with an additional 9 holdout choices. ICE and HB performed equally well in predicting individual holdout choices and aggregate choice shares. Latent Class did not perform as well on either measure.

Conclusions:

The good news is that it is possible to generate reliable individual part worths from choice data and that these values can be used in choice simulators, just as in classic conjoint analysis.

Choice models that recognize heterogeneity through individual-level analysis strongly outperform aggregate-level models in terms of predicting consumer choice.

The important result is that although HB is more theoretically elegant than ICE, our experience suggests that both methods work equally well in practice. Latent Class, for its part, does a poor job of predicting individual choices unless its weights are allowed to be negative, as they are with ICE.