Customer Heterogeneity in Purchasing Habit of Variety Seeking Based on Hierarchical Bayesian Model

University of Tsukuba
Kondo, Fumiyo N.; Kuroda, Teppei
Date: August 13, 2008
Place: Technische University of Dortmund
Agenda

1. Research Objective and Background
2. Analyzed Data
3. Analyzed model
   a mixture normal-multinomial logit model in a hierarchical Bayesian framework
4. Result 1 (latent class VS hierarchical Bayesian)
5. Result 2 (Bawa model Vs proposed model)
6. Summary and Future Research Topics
A product choice behavior is called as “inertia” if a customer chooses the same product as the previously purchased and “variety seeking” if it is a different product from the previous one. (Givon (1984), Lattin et al. (1985))

These kinds of behaviors are frequently observed in the product category of “low involvement” (Dick and Basu (1994), Peter and Olson (1999)).
Consumers tend to purchase a “low involvement” product such as beverage or cake based solely on experience, inertia, or atmosphere. In addition to “inertia” or “variety seeking”, Bawa (1990) proposed a model for segmentation purposes. It has an additional segment of “hybrid” customer, of which purchasing tendency changes from “inertia” to “variety seeking” or vice versa.
Illustration of purchase history by customer type

- Inertia: AAAAAAAAA
- Variety seeking: ABCDCFGAFE
- Hybrid: AAABBBBCC
Research Objective

1. To express product choice behavior in terms of Inertia / Variety Seeking toward product attribute by customer.
2. To explore effective marketing strategy.
3. To compare results with those by Latent class model.

- a mixture normal-multinomial logit model in a hierarchical Bayesian framework
Analyzed Data

Analyzed store:

5 super market stores around Tokyo

Analysis period: 2000.1.1～2001.5.31

Analysis subcategory:

Japanese tea • Chinese tea

① extract 7000 customers by random sampling from all of 13238 panels.
Analyzed Data
< latent class model vs hierarchical Bayesian model >

② screening

A. exclude simultaneous purchase opportunities
B. include customers who purchased once or more in 3 periods (2000.1.1~6.30; 7.1~12.31; 2001.1.1~5.31)
C. include customers with 24 times or more purchases (only heavy users)
D. exclude customers with once or less brand switching
E. exclude customers with 3 times or less purchases on hold-out samples (in the third period)
Multinomial Logit Model (MNL)

$U_{ijt}$: utility of product $j$ for customer $i$ in period $t$

$v_{ijt}$: fixed utility

$\varepsilon_{ijt}$: random utility (double exponential distribution)

$X_{ijt}$: explanatory variable of product $j$ for customer $i$ in period $t$

$\beta_i$: parameter for customer $i$

\[
U_{ijt} = v_{ijt} + \varepsilon_{ijt} \quad v_{ijt} = X_{ijt} \beta_i
\]
Explanatory Variable

**Inertia / Variety seeking**

- repeat purchasing times $r$ of a brand and $r^2$
  
  (Bawa(1990,1995), Sakamaki(2005))

  let the latest brand switching time as period $s$

  $$r_{ij} = \sum_{t=s}^{t-1} y_{itj}$$

  $$Z = -\frac{\exp(purchasing\ interval - a)}{1 + \exp(purchasing\ interval - a)} + 1$$

- $r \times Z$ and $(r^2) \times Z$

**Promotion variable** (Seetharamann et al(1998), Kawabata(2004))

- discount rate; displays; flyers for each subcategories of Japanese or Chinese tea

- Constant term
Explanatory Variable

<repeat purchasing times r & r^2>

\[ \nu_{ijt} = \beta_{i1} r_{ijt} + \beta_{i2} r_{ijt}^2 \]

\( \nu_{ijt} \): fixed utility of inertia / variety

see king for customer \( i \) in period \( t \) brand \( j \)

\( r_{ijt} \): repeat purchasing times for customer \( i \) in period \( t \) brand \( j \)

\( r_{ijt}^2 \): the second power of \( r_{ijt} \)

\( \beta_{i1}, \beta_{i2} \): parameters
Explanatory Variable

\[ Z = -\frac{\exp(purchasing\ interval - a)}{1 + \exp(purchasing\ interval - a)} + 1 \]
Latent class model

\(\pi_s\): probability of segment \(s\)

\(p_{it}(j | \alpha_s)\) : choice probability of product \(j\) belonging segment \(s\)

\[
p_{it}(j | \pi, \beta) = \sum_{s=1}^{S} p_{it}(j | \beta_s) \pi_s
\]

where \(\sum_{s=1}^{S} \pi_s = 1\) \((\pi_s \geq 0, \forall s = 1, \cdots, S)\),

\[
\pi = [\pi_1, \cdots, \pi_s], \beta = [\beta_1, \cdots, \beta_s]
\]
A mixture normal-multinomial logit model in a hierarchical Bayesian framework (Rossi et al. (2005))

\[
y_{ijt} \sim MNL(P_{it}(X_{ijt}, \beta_i)) \quad \text{(MNL: multinomial logit model)}
\]

\[
\beta_i \sim N(\mu_{ind_i}, \Sigma_{ind_i})
\]

\[
\mu_k \sim N(\mu, \Sigma_k \otimes a_{\mu}^{-1})
\]

\[
\Sigma_k \sim IW(v, V)
\]

\[
P_{it}(X_{ijt}, \beta_i): \text{choice probability of product } j \text{ for customer } i \text{ in period } t
\]

\[
X_{ijt}: \text{explanatory variable of product } j \text{ for customer } i \text{ in period } t
\]

\[
\beta_i: \text{parameters for customer } i
\]

\[
\text{ind}_i \sim \text{Multinomial}_K(pvec)
\]

\[
pvec \sim \text{Dirichelet}(\alpha)
\]
Parameter Distribution Estimation Methods & Information Criterion

- **Parameter Distribution Estimation Methods**
  - latent class model: Maximum Log-likelihood
  - hierarchical Bayesian model: MCMC method

- **Information Criterion**
  - AIC (Akaike)
  - BIC (Schwarz)
  - CAIC (Bozdogan)
  - DIC (Spiegelhalter et al., 2002)

The smaller value of information criterion, the better model.
Analysis Result 1

< latent class model: for heavy users of 63 panel >

-Determination of No. of Segments-

<table>
<thead>
<tr>
<th></th>
<th>AIC</th>
<th>BIC</th>
<th>CAIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1segment</td>
<td>3892.91</td>
<td>3988.52</td>
<td>3988.52</td>
</tr>
<tr>
<td>2segment</td>
<td>3910.15</td>
<td>4106.97</td>
<td>4106.99</td>
</tr>
<tr>
<td>3segment</td>
<td>3925.08</td>
<td>4223.13</td>
<td>4223.16</td>
</tr>
</tbody>
</table>

- Hypothesis A (2 segments): VS • Inertia & Hybrid
- Hypothesis B (3 segments): VS • Inertia • Hybrid

For 1 segment, the model was the best with the minimum value for all of Information Criterions
Analysis Result

<comparison of 3 models: for heavy users of 63 panel>

-hit rate & Information Criterion-

<table>
<thead>
<tr>
<th>model</th>
<th>Log-L</th>
<th>DIC</th>
<th>Hit rate1</th>
<th>Hit rate2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latent class model</td>
<td>------</td>
<td>------</td>
<td>0.749</td>
<td>0.624</td>
</tr>
<tr>
<td>H. Bayes model (1 normal dist.)</td>
<td>-958</td>
<td>5425</td>
<td>0.798</td>
<td>0.680</td>
</tr>
<tr>
<td>H. Bayes model (3 normal dist.)</td>
<td>-942</td>
<td>5333</td>
<td>0.811</td>
<td>0.734</td>
</tr>
</tbody>
</table>

- Two hierarchical Bayesian models that can estimate parameters for each customer are better than latent class model in terms of hit rate.

- a mixture normal (3 dist.)-multinomial logit model in a hierarchical Bayesian framework is selected as the best model for all of criteria.
Analyzed Result

<Bawa model vs proposed model: for heavy users of 129 panel> -hit rate & DIC-

<table>
<thead>
<tr>
<th>Model</th>
<th>Log-L</th>
<th>DIC</th>
<th>Likelihood</th>
<th>Hit rate1</th>
<th>Hit rate2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bawa model</td>
<td>-2147</td>
<td>12251</td>
<td>-2210</td>
<td>0.856</td>
<td>0.713</td>
</tr>
<tr>
<td>Model A</td>
<td>-2151</td>
<td>12287</td>
<td>-2227</td>
<td>0.860</td>
<td>0.756</td>
</tr>
<tr>
<td>Model B</td>
<td>-2139</td>
<td>12223</td>
<td>-2206</td>
<td>0.863</td>
<td>0.750</td>
</tr>
<tr>
<td>Model C</td>
<td>-2145</td>
<td>12230</td>
<td>-2210</td>
<td>0.860</td>
<td>0.736</td>
</tr>
</tbody>
</table>

- Bawa model : no purchase interval considered
- Proposed model A : a=10
- Proposed model B : a=15
- Proposed model C : a=20

Proposed model B is the best model than Bawa model in terms of DIC and hit rate1.
Analysis Result

-Response to promotion for Japanese tea-

<table>
<thead>
<tr>
<th>Japanese tea</th>
<th>j-discount</th>
<th>j-display</th>
<th>j-flyers</th>
<th>No. customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inertia</td>
<td>1.55</td>
<td>-0.21</td>
<td>0.13</td>
<td>41</td>
</tr>
<tr>
<td>VS</td>
<td>1.05</td>
<td>0.37</td>
<td>0.34</td>
<td>10</td>
</tr>
<tr>
<td>Hybrid</td>
<td>1.14</td>
<td>-0.49</td>
<td>0.59</td>
<td>26</td>
</tr>
<tr>
<td>Zero-order</td>
<td>3.79</td>
<td>0.08</td>
<td>0.21</td>
<td>52</td>
</tr>
</tbody>
</table>

- Zero-order: high response to discounts
- Inertia • VS • Hybrid: low response to discounts
- A strategy different from usual discounts for the customers of Variety Seekers are necessary!
Summary

- Latent class model
  No valid segmentation was possible.

- Hierarchical Bayesian Models
  - It is possible to estimate parameters for all customers.
  - It is possible to do the optimum promotion for each Hybrid customer.
  - For VS customers, it may be necessary to consider brand choices of previous 2 purchases.
Future Research Topics

- Analysis on data on different shop type with different customer characteristics or on different usage scenes

- To vary the decreasing speed of tendency of Inertia or Variety seeking by customer accompanying with purchasing interval.
Reference

Thank you for patience!