Preference Mapping With Incomplete Blocks: A Review

Graham Cleaver
Unilever R&D
Vlaardingen, Netherlands

Sensometrics Conference
July 25th - 28th 2010
Overview

• Background
• Design issues
• Sensometrics 2004: Workshop
• Examples of different analysis techniques in the context of incomplete block designs
  • External analysis: PrefMax, Latent Class
  • Internal analysis: MDPref, CLIP, PrefScal, LSA
• Summary

Acknowledgments

• Richard Popper: Sensometrics Workshop Summary
• Pascal Schlich: PrefMax, CLIP
• Frank Busing: PrefScal simulations
• Danny Ennis: LSA slides
Background: Challenge

Preference Mapping Objectives:
Systematic coverage of relevant sensory space
Robust models and understanding of drivers of liking
Large no. of products – typically 12-16

Modelling Objective
At level of individual consumers
Maximise no. products / respondent

Practical Constraint
Need to avoid sensory fatigue
Minimise no. products / respondent

Pragmatic Solution:
Split products to be tested over more than one day and session
Not ideal (cost, consistency over time) – need to consider alternatives

Incomplete Block Designs
Design Issues

- Thickness
- Sweetness
- Smoothness
- Rate of melt
- Firmness
- Sourness
- Bitterness
Design Considerations

Concerns

• Estimated individual ideal depends on particular set of products assessed
• Segmentation may be driven by incomplete patterns
• Single very influential product could dominate segmentation

Design aspects are critical

General good practice
Incomplete designs balanced for order and carry-over effects

Exploit product structure
e.g. Block designs for factorial and fractional factorials

Ref:
Wakeling, I.N. & MacFie H.J.H. Designing consumer trials balanced for first and higher orders of carryover effect when only a subset of k samples from t may be tested
Food Quality and Preference 6 (1995) 299-308

7 factors each at 2 levels
$2^7 = 128$ possible products

Fractional factorial design
$1/4 = 32$ products

Balanced block design
8 products / consumer
Design Opportunities

- Total no. of products = 12
  - Incomplete block design
  - 6 products / respondent

- Total no. of products = 48
  - Incomplete block design
  - 12 products / respondent
Data Analysis Workshop

Consumer Segmentation & Key Drivers Analysis

7th Sensometrics Meeting
July 28-30, 2004
Davis, CA
USA

Data supplied by:
CFIFL / INRA (Pascal Schlich)

Review of findings by:
Richard Popper
## Study Description

<table>
<thead>
<tr>
<th></th>
<th>Study Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tomato Varieties</td>
<td>17</td>
</tr>
<tr>
<td>Sensory Panel</td>
<td>14 panelists</td>
</tr>
<tr>
<td>Sensory Attributes</td>
<td>11</td>
</tr>
<tr>
<td>Physical/Chemical Analyses</td>
<td>15</td>
</tr>
<tr>
<td>Consumers</td>
<td>N=379 tasted 10 of 17 varieties</td>
</tr>
<tr>
<td>Hedonic Rating</td>
<td>Overall liking</td>
</tr>
<tr>
<td>Reason for Preference</td>
<td>Preference between most &amp; least liked, with reason for preference checklist</td>
</tr>
<tr>
<td>Appearance Liking</td>
<td>7 varieties ranked for appearance liking</td>
</tr>
<tr>
<td>Usage and Attitudes</td>
<td>17 questions</td>
</tr>
</tbody>
</table>
Method Comparisons

- Segmentation technique
  - liking alone
  - use external variables (e.g. sensory)
- Treatment of missing values
  - accept missing values
  - impute missing values
- Data pre-treatment
  - liking data normalized
  - data reduction technique for sensory
- Selection of number of clusters
  - judgment
  - statistical criterion
- Type of selection of drivers?
  - linear only
  - quadratic drivers included
## Method Comparisons

<table>
<thead>
<tr>
<th></th>
<th>Ledauphin</th>
<th>Lengard</th>
<th>Lundahl</th>
<th>Cleaver</th>
<th>Meullenet</th>
<th>Schlich</th>
<th>Tang</th>
<th>Zalila</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Segmentation based on</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liking alone</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Liking w respect to external variables</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Impute missing values?</strong></td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Liking data pre-treatment</strong></td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Selection of number of clusters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Judgment</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical criterion</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Include quadratic drivers?</strong></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>
**PrefMaX Method**

For each pair of consumer/attribute:
1. Fit a quadratic regression of hedonic scores on attribute means
2. Define optimal intensity (XOPT)
3. Store all XOPT into a *consumer x attribute* matrix
4. XOPT matrix is the input of subsequent analyses
5. In these analyses, weight each XOPT by the $R^2$ from the corresponding quadratic regression

**Optimal Sensory Recipes by Consumer Segment**

<table>
<thead>
<tr>
<th>Attribute</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acidic</td>
<td>0</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Bitter</td>
<td>0</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Salty</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sweet</td>
<td>0</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Umami</td>
<td>+</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*Each white star is the ideal point of a consumer*

*Each violet star is a product projected onto the map as a supplementary point using its attribute mean intensities*

2001 Tomato PrefMaX

S1 « Flavor & Firm segment » (49%)
S2 « Flavor & Soft segment » (27%)
S3 « Firm segment » (13%)
S4 « Soft segment » (11%)

379 Consumers – 17 Tomato varieties – 13 Attributes

• Does not have pre-requisite for complete data
• Potential to work well with incomplete data: models at underlying segment level not individuals

1 Class

Model Class 1

100 %

2 Classes

Model Class 1

100 - P1 %

Model Class 2

P1 %

3 Classes

Model Class 1

P1 %

Model Class 2

P2 %

Model Class 3

100 - P1 - P2 %

Simplicity Complexity

Under-Fitting

Over-Fitting

CAIC

External: Latent Class Regression Simultaneous Modelling & Segmentation

No. Of Latent Classes

External: Latent Class Regression Simultaneous Modelling & Segmentation

No. Of Latent Classes
External: Latent Class Regression
Example: Tomato Data Sensometrics (2004)

- 17 Tomato varieties
- Each consumer rated 10 / 17

379 Consumers

Latent Class Regression
Extended Model

‘Random Scoring’

Class 1
49%

Class 2
31%

Class 3
20%

$R^2 = 0.97$

$R^2 = 0.98$
Internal: MDPref

- Ideal vectors – suitable for where ideal regions are towards the outside
- Does have pre-requisite for complete data
**MDPref: Monte-Carlo Simulation**

### Missing value imputation
- Expectation Minimisation (Beale & Little)
- Row-Column Substitution (Krzanowski)
- Proc PRINQUAL (SAS)
- MISTRESS Algorithm (van Buuren)
- Mean substitution

### Factors varied
- No. of subjects: 50 200
- No. of stimuli: 10 30
- Dimensionality of pref space: 2D 4D
- Level of noise in data: SD=1 SD=2.0
- Proportion incomplete data: 5 35 65%

### Outcome
- Simple mean substitution as good as other techniques
- Level of noise was most influential factor
- Product positions stable with incomplete data
- Level of incompleteness:
  - 5% : All techniques gave good results
  - 35% : Results may be questionable
  - 65% : No technique gave good results

**Ref:**

Cluster Analysis
Usually hierarchical, applied to raw liking scores for each product to cluster respondents

Conventionally, requires complete data for each respondent

Some techniques allow missing values
eg PROC FASTCLUS (SAS)

Has not been evaluated systematically in context of incomplete block preference mapping

May be more suited to randomly distributed missing values, rather than (high) proportion of missing values for each line of data
Internal: Clustering (CLIP)

• **CLustering of Incomplete Preferences**
  - Define measure of similarity between respondents based on scores for products
  - MDS to create plot of respondents
  - Cluster analysis of respondents

• Three data sets used as a basis for simulation where data was systematically removed

• Results / Recommendation
  - Low noise data: Half the samples should be tasted by each assessor
  - Noisy data: Two-thirds of the samples are required.

Internal: PrefScal

• Ideal point unfolding
• With optimal scaling of liking scores
• Can incorporate external information – ‘restricted unfolding’
• Does not require complete data

Refs


• Busing, F.M.T.A., Heiser, W.J., Cleaver, G.J. ‘Restricted unfolding: Preference analysis with optimal transformations of preferences and attributes’ Food Quality and Preference 2010 Vol21 (1) p82-92
PrefScal: Simulation study

**Method**
- Real and simulated data with varying levels of completeness
- Comparison solutions based on incomplete vs complete data

**Criteria**
- Tucker’s congruence coefficient ($\Phi$)
- Kendall’s rank order correlation ($\tau_b$)

**Outcome**
Charts with guidance on proportion inclusion required, in relation to:
- (a) No. of products
- (b) No. of respondents
- (c) Level of variation in data

- Small no. of products (10- products): Include complete data for each respondent
- Larger studies (15+ products and 40+ respondents): Up to 50% can be missing and still give comparable results

**Ref:**
LSA first “unfolds” liking and creates a space relevant to consumer acceptability (6 products, 44 consumers)

- The closer a consumer is to a product, the more he/she likes it
- Contours indicate consumer densities and facilitate the visualization of potential segmentation

Ref

- IFPrograms (Institute For Perception)
Landscape Segmentation Analysis®

**Background (2/2)**

- Descriptive data is then added by regressing the attributes on the map using the relationship between the original scale data and the values predicted by projecting each product on the map’s attribute.
- Some attributes can be fit on the map and are drivers of liking.
- Others can’t and are less relevant to consumer acceptability.
LSA results with complete block and unbalanced incomplete block arrangements

- 200 consumers, 10 cookies
- Degree of incompleteness
  - Complete block: all 200 consumers evaluate all 10 products
  - Unbalanced incomplete block: 20% of the data randomly removed
    - Should be a worse case scenario than a balanced incomplete block

![Bar chart showing the number of products evaluated by a consumer]
LSA results with complete block and unbalanced incomplete block arrangements

- Solutions are almost identical re products and both show two segments
- Drivers of liking highly similar (next slide)
LSA results with complete block and unbalanced incomplete block arrangements

Complete block solution

Unbalanced incomplete block solution

Driver of Liking®
Cross-Methdoology Evaluation

Actual Data
Validated Models
Varying Product categories / Cultures / Size / etc

Simulated Data
Known Models
Systematic variation in Size / Noise / Dimensionality / etc

Creation Of Subsets Of Varying Completeness
Systematic removal

PrefMax  LCR  MDPref  CLIP  LSA  PrefScal

Comparison vs Validated / True Model
Based on common performance criteria
Summary

• Some analysis preference mapping techniques have a pre-requisite for complete data for each respondent, e.g. MDPref. Most do not.

• There are many examples of application of preference mapping to incomplete data and evaluations of the impact of different levels of incompleteness.

• Scope for systematic evaluation across methodologies based on common criteria.

• Recommendation may depend on absolute number of products per respondent, rather than proportion

• Design aspects are critical and deserve further attention