Statistical Models for Market Segmentation

L. Camilleri and M. Green

Centre for Applied Statistics, Lancaster University, Lancaster LA1 4YF. UK.
l.camilleri@lancaster.ac.uk
m.green@lancaster.ac.uk
Student Oral Presentation

1 Introduction

It is an essential element of market research that customer preferences are considered and the heterogeneity of these preferences is recognized. By segmenting the market into homogeneous clusters the preferences of customers is addressed. Latent class methodology for conjoint analysis, proposed by Green (2000), is one of the several conjoint segmentation procedures that overcome the limitations of aggregate analysis and priori segmentation. This approach proposes the proportional odds model as a proper statistical model for ordinal categorical data in which the item attributes are included in the linear predictor. The likelihood is maximized through the EM algorithm. This paper considers two extensions of this methodology that incorporate individual characteristics into the models.

Keywords: Proportional Odds Model; Latent Class Model; EM algorithm; Conjoint Analysis; Segmentation

2 A General Model

Individuals are presented with several items with different characteristics and each item has to be rated on an ordinal scale. The observation y_{nj} is a rating response to the jth item elicited by the nth respondent. In the first extension individual characteristics together with item attributes are included in the same linear predictor.

$$P\left(y_{nj} = r|\boldsymbol{\alpha}, \boldsymbol{\beta}\right) = F\left[\alpha_r + \eta_{nj}\right] - F\left[\alpha_{r-1} + \eta_{nj}\right]$$

$\boldsymbol{\beta}$ is a vector of regression parameters and $\boldsymbol{\alpha}$ is a vector of cut-point parameters. The linear predictor $\eta_{nj} = \eta\left(\mathbf{x}_j, \mathbf{z}_n\right)$ includes item attribute covariates, $\mathbf{x}_j$, individual covariates, $\mathbf{z}_n$ and interaction terms. In market research η_{nj} is referred to as the "worth". The choice of $F(.)$ considered is the extreme value distribution leading to the complementary log-log link. The proportional odds model assumes that all respondents act in a similar way in their choice behaviour and that it treats all respondents as homogeneous. For the segmentation procedure a latent class model with K segments is considered.

$$P\left(\mathbf{y}_{nj} = \mathbf{r}|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\pi}\right) = \sum_{k=1}^{K} \pi_k . P\left(\mathbf{y}_{nj} = \mathbf{r}|\boldsymbol{\alpha}, \boldsymbol{\beta}_k\right)$$

where π_k is the proportion of respondents in the kth segment and the parameters within the segments are estimated at the same time that the segments are uncovered.

In the second approach the item attributes are included in the proportional odds model. This is the same model proposed by Green (2000)

$$P\left(y_j = r|\boldsymbol{\alpha}, \boldsymbol{\beta}\right) = F\left(\alpha_r + \mathbf{x}_j'\boldsymbol{\beta}\right) - F\left(\alpha_{r-1} + \mathbf{x}_j'\boldsymbol{\beta}\right)$$

The individual covariates are included in a mixture model through a classifying function π_{nk}. The choice of parameterization for π_{nk} corresponds to a multinomial logit probability model.

$$\pi_{nk} = \frac{\exp\left(\mathbf{z}'_n \gamma_k\right)}{\sum_{k=1}^{K} \exp\left(\mathbf{z}'_n \gamma_k\right)}$$

A mixed multiplicative model blends this multinomial logit model containing individual covariates with the proportional odds model containing item attributes covariates.

$$P\left(\mathbf{y}_{nj} = \mathbf{r} | \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\pi}\right) = \sum_{k=1}^{K} \pi_{nk}.P\left(\mathbf{y}_{nj} = \mathbf{r} | \boldsymbol{\alpha}, \boldsymbol{\beta}_k\right)$$

3 Implementation

In this work we concentrate on the more general second approach. The model is fitted using the EM algorithm and is implemented as a set of GLIM macros. The responses are converted to zero/one indicators that allow the use of the Poisson Likelihood in the model fit. The proportional odds model being a non-linear model can be accommodated using the OWN model facilities. The EM algorithm for fitting latent class models, proposed by Dempster et al (1977), is equivalent to iterative fitting of a weighted GLM with posterior probabilities recalculated at each iteration.

4 Application

To illustrate the methodology a conjoint study of 186 customers was conducted to investigate consumer car preferences. Five factors were identified as being key determinant attributes in the car market. The car attributes were brand, price and the number of doors and the individual characteristics were gender and age. The study compared 4 different price values, 4 brands and whether the car had 3 or 5 doors. We utilized a full profile method of collecting respondent evaluations. The design chosen had two blocks of 16 cards each. The respondents were handed a set of 16 cards to compare with random assignment to block. The rating responses had seven categories where 1 corresponds to "worst" and 7 to "best". The terms for the effect of price in the linear predictor was assumed to be quadratic. This relationship allows a dual role for price, the negative cost deterrent effect and a positive effect due to perceived quality.

Models with four segments were chosen on the basis that they comply with what we would expect as regards to human behaviour. The Mixture model price profiles in figure 1 show the expected worth of each brand in the four fitted segments. Segment 1 represents consumers who have a moderate brand preference and are not strongly influenced by price. Respondents in segment 2 exhibit a strong reliance on price as a signal of quality but who hardly discriminate between the brands. People in segment 3 are differentiating between the brands and are assessing the price of the product as a monetary constraint in choosing it. Respondents in segment 4 have a strong brand preference and applying an "ideal price" as a signal that buying at very low prices could result in too low quality but see no bargain in buying at high prices.Figure 2 shows the fitted model for segment membership probability as a function of age and gender. Segment 3 differs from the other segments mainly that it consists of a younger age group which is a cautious cost driven but brand selective group. Segment 1 differs from the other segments mainly that it consists of more females than males. The final paper will include the price profiles of the Latent Class model and a comparison of the segmentation by the two models will be discussed.

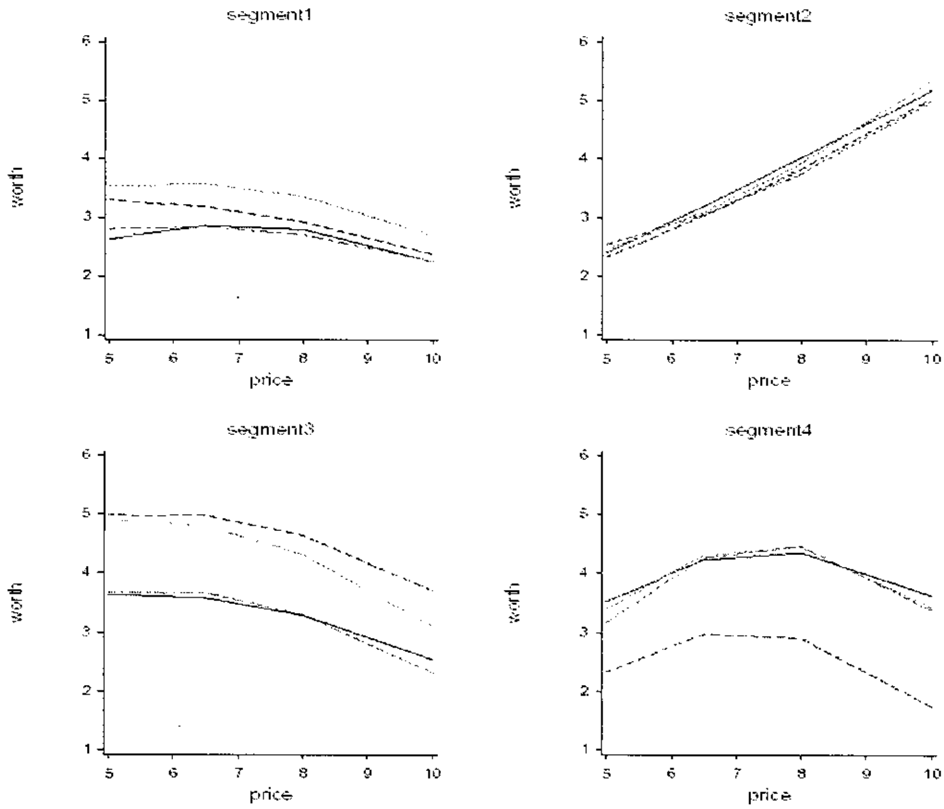

Figure 1:

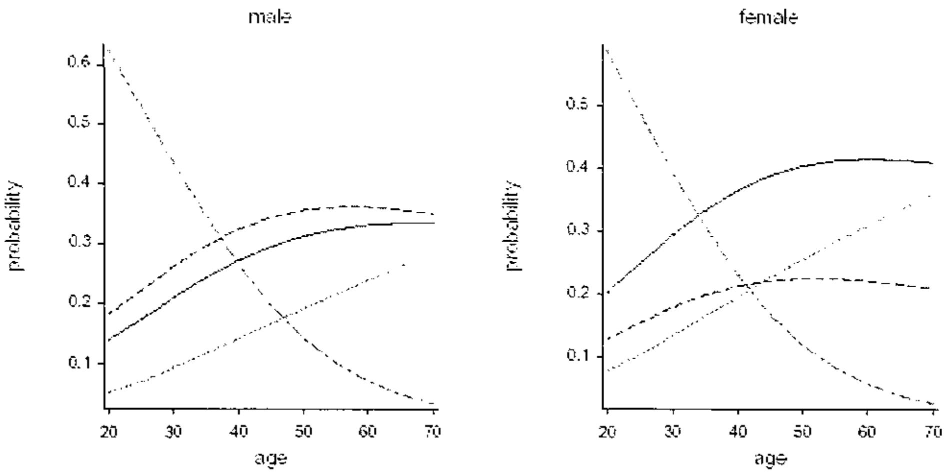

Figure 2:

5 Predicting preferences

Comparing the deviances of the two models is inadequate because the models are not nested. Standard diagnostic tools to check for outliers, influential data points and other model misspecifications cannot be used because the proportional odds model is a non linear and a non standard GLM. So a further task was included in the study in which each person was presented with four choice cards to choose the item that he preferred most. For the extreme value distribution it is possible to derive the probability of preference from the predicted worth. The expected frequencies can hence be estimated by using the following result

$$P\left(\text{preference for } j^{th} \text{ item}\right) = \frac{\exp\left(W_j\right)}{\exp\left(W_1\right) + \ldots + \exp\left(W_4\right)}$$

	Expected Frequency			
	Brand			
	A	B	C	D
Seg 1	24.7	7.43	3.51	16.4
Seg 2	7.05	1.71	0.99	19.3
Seg 3	7.23	32.8	10.9	11.1
Seg 4	24.4	12.7	4.11	1.80
Total	63.3	54.6	19.5	48.5

	Observed Frequency			
	Brand			
	. A .	. B .	.C .	. D .
Seg 1	22	10	4	16
Seg 2	7	4	3	15
Seg 3	9	29	11	13
Seg 4	23	9	7	4
Total	61	52	25	48

The two tables show the observed and expected frequencies for the Mixture model. Visual comparison shows that the model is picking up the main features of individual preferences quite well. Similar results for the Latent Class model will be given in the final paper and the results of the two models will be compared. A larger data set will be used in the final presentation.

References

[1]Dempster,A.,Laird,N. and Rubin,D. (1977) Maximum likelihood from incomplete data via the EM algorithm, *J.R.Stat Soc B.* **51**, 127-138.

[2]DeSarbo, W.S.,Wedel,M.,Vriens.M. and Ramaswamy, V (1992), Latent class metric conjoint analysis, *Marketing Letters*. **3**, 273-288.

[3]Green, M. (2000), Statistical Models for Conjoint Analysis, proceedings of the 15th International Workshop on Statistical Modelling, Bilbao. 216-222.