Some Considerations on Latent Segmentation Procedures

Considerazioni sulle tecniche di segmentazione nei modelli SEM

Giuseppe Boari, Gabriele Cantaluppi
Dipartimento di Scienze statistiche, Università Cattolica del S.Cuore, Milano
e-mail: giuseppe.boari@unicatt.it

Keywords: latent segmentation, SEM, PLS

1. The concept of market segmentation

Market segmentation is a key element in marketing analysis and planning, which enables to recognize the unobserved heterogeneity of customers. The recent research offers the marketers new techniques for segmentation, mainly related to the availability of the so-called segmentation basis, i.e. a set of variables or characteristics (observable or not) used to assign potential customers to homogeneous groups.

Moreover, in the last decades, great attention has been devoted to the measurement and analysis of Customer Satisfaction (CS), operated by means of the structural equation models (SEM). Later on, we shall make reference to this context.

Following Wedel and Kamakura (1998), we may classify the segmentation methods in accordance with the categories listed in the following table.

Table 1: Segmentation approach

<table>
<thead>
<tr>
<th>A priori (based on:)</th>
<th>Demographic variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Psychographics, Life style, Perceptions,…</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Post Hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finite mixture models</td>
</tr>
<tr>
<td>Latent class models</td>
</tr>
<tr>
<td>Clustering methods</td>
</tr>
</tbody>
</table>

Observe that post-hoc methods are also called response-based procedures, and that, within this framework, a series of interesting solutions was presented in literature. In Vinzi et al. (2007) three different post-hoc methods are discussed and their performance is compared by means of a simulation study. They refer, in particular, to the following PLS algorithm based procedures: FIMIX-PLS, PLS-TPM and REBUS-PLS.

They all take into consideration the manifest variables, as well as the latent ones, whose scores are previously generated adapting a common SEM model to the whole sample of data. Moreover, in particular, FIMIX-PLS assumes that observations come from a finite mixture of normal populations (see, also, Hahn et al., 2002), while the others adopt proper measures of distance from units to the local models.

A latent segmentation procedure is also presented in Boari, Cantaluppi (2007), which employs the estimated score of the latent variables in order to perform a hierarchical cluster analysis of the respondent units. A common aspect of the four previously mentioned procedures is that they all start from the estimation of a global path model; then, afterwards, they tackle the problem of detecting the unobserved heterogeneity and de-
fining the cluster of data. We want to outline that heterogeneity is ascribable essentially to different covariance structures among the groups (and therefore different inner relationships in the local models) and/or to different mean levels of the latent variables.

2. A priori segmentation approach

Typically, heterogeneity should be addressed by assuming that subjects can be assigned to segments a priori, on the basis of proper variables (observable or not), like for example, demographical and psychographic ones, or much better, variables pertaining Personality, Perceptions or Intentions, characterizing the respondents.

With reference to the application of SEM in the measurement of Customer Satisfaction, it was recently reported in Grace (2005) that CS is highly related to the so-called notion of Consumer Disposition toward Satisfaction (CDS). The author derived the following five item scale to measure it (properly purified and validated):

- Usually I am pleased with what I buy
- More often than not, I am a satisfied consumer
- I generally find the goods and services I buy don’t live up my expectations (*)
- Quite often I am dissatisfied with my purchases of good and services (*)
- Overall, I am usually satisfied with the purchases I make

where (*) indicates reverse scoring. Assuming that CDS well defines the existing differences among customers, we propose to adopt its scores in order to perform a priori segmentation. Different path models, of the identical structure, can be adapted locally. Then, an evaluation of the correct identification of the latent groups may be performed by testing the following hypotheses: homogeneity of covariance matrices between clusters (Box’s M test) and/or equality of mean levels of the latent variables (MANOVA test). A specific application to a real case has been planned for a forthcoming CS survey: the questionnaire has been defined, while data collection is still in progress.

References


