RESEARCH SUPPORT
UNIVERSITY OF MICHIGAN BUSINESS SCHOOL

APRIL 1997

A NOTE ON LATENT SEGMENTATION MODELS

WORKING PAPER #9712-08

BY
STEPHEN WALKER
IMPERIAL COLLEGE, LONDON, UK

AND

PAUL DAMIEN
UNIVERSITY OF MICHIGAN BUSINESS SCHOOL

A Note On Latent Segmentation Models

Stephen Walker[1]
and
Paul Damien[2]

[1] Department of Mathematics, Imperial College, 180 Queen's Gate, London SW7 2BZ.

[2] Business School, University of Michigan, Ann Arbor, 48109-1234, USA.

Summary

Developing polytomous models of choice which could enable in identifying market segments has been the focus of attention in many papers in recent literature (Ramaswamy et al., 1996 and the references therein).

In this note, we show that latent segmentation models of the type developed by Ramaswamy et al. is *always* nonidentifiable, and hence the parameters of interest are nonestimable.

The joint segmentation model of Ramaswamy et al. (1996) considers data from consumers on two sets of categorical variables that serve as distinct bases for segmentation and are interdependent (specifically, benefits sought and services used). It is apparent that the data collected provides no information concerning the allocation of consumers to particular segments (indexed by (i, j)). Therefore, it is readily apparent that it is impossible to make inference (i.e., estimate) the parameter θ_{ij}, which is the joint probability of a consumer belonging to segment i and segment j. Consequently, the model they construct is nonidentifiable.

To illustrate this point we consider an analagous model: let observations from N individuals be categorised from the finite set $\{1, ..., m\}$, and let $x_{nk} = 1$ if the nth individual is in category k and zero otherwise, and let π_k be the probability that an individual is in category k. Now let θ_i represent the probability of a consumer being in segment i and π_{ik} the probability of an individual in segment i being in category k. Obviously we have $\sum_k \pi_{ik} = 1$, $\sum_i \theta_i = 1$ and $\sum_k \pi_k = 1$. The likelihood is given by

$$\prod_n \left(\sum_i \theta_i \left[\prod_{k=1}^{m} \pi_{ik}^{x_{nk}} \right] \right).$$

This likelihood expression is essentially a special version of the type of likelihood appearing in equation (2) of Ramaswamy et al. (1996). Clearly our likelihood can be written as

$$\prod_n \left(\sum_i \theta_i \pi_{ik_n} \right),$$

where k_n is the category of the nth individual. This becomes

$$\prod_{k=1}^{m} \left(\left[\sum_i \theta_i \pi_{ik} \right]^{n_k} \right),$$

where n_k is the number of individuals located in the kth category. Finally, our likelihood becomes

$$\prod_{k=1}^{m} \pi_k^{n_k},$$

which is precisely what we should have started with in the first place. It is obvious the θ_i and π_{ik} are nonestimable. An exact argument to the one above can now be applied to the likelihood appearing in Ramaswamy et al. (1996, equation (2)). The details are now given.

Let L denote equation (2) from Ramaswamy et al. (1996), i.e. the likelihood function. We can write this likelihood in a different way:

$$L = \prod_{n=1}^{N} \sum_{ij} \theta_{ij} \left(\prod_{k=1}^{K} \gamma_{ikl_{kn}} \prod_{m=1}^{M} \delta_{jml_{mn}} \right),$$

where, for example, ikl_{kn} denotes that the nth individual is in the lth category, $(l = 1, ..., L_k)$, for the kth variable, $(k = 1, ..., K)$, in segment i; and so, $\gamma_{ikl_{kn}}$ denotes the probability of this event. Let $\mathcal{C}$ denote the set of possible categories, and n_c denote the number of individuals who are in $c \in \mathcal{C}$. Then, we have

$$L = \prod_{c \in \mathcal{C}} \left(\sum_{ij} \theta_{ij} \pi_{ijc} \right)^{n_c},$$

where π_{ijc} is the probability of being in category c conditional on being in segments i and j. This last expression is identical to

$$\prod_{c \in \mathcal{C}} \pi_c^{n_c},$$

which, again, is the *correct* likelihood: so, clearly, this implies that θ_{ij} is nonidentifiable.

The models discussed above should not be confused with *continuous* mixture models:

$$f(y; \theta, \phi) = \sum_s \theta_s f(y; \phi_s),$$

where $\sum_s \theta_s = 1$. Each (θ, ϕ) defines a *unique* density and hence such models are identifiable (in general). The *discrete* categorical model is given by:

$$p(y = k; \theta, \pi) = \sum_s \theta_s p(y = k|s).$$

Let $\pi_{sk} = p(y = k|s)$. Clearly $\sum_s \theta_s \pi_{sk} = \tilde{\pi}_k$, where $\tilde{\pi}_k = p(y = k; \tilde{\pi})$. Note, then, that for all (θ, π),

$$p(y = k; \theta, \pi) = p(y = k; \tilde{\pi}).$$

Hence, the model is nonidentifiable for θ; i.e., many (θ, π), in fact an infinite number, lead to the same probability model for the data.

Acknowledgement

We thank Venkat Ramaswamy for drawing our attention to the literature on latent class models, and for valuable discussions on this topic.

REFERENCE

Ramaswamy,V., Chatterjee, R. and Cohen, S.H. (1996). Joint segmentation on distinct interdependent bases with categorical data. *Journal of Marketing Research*, 33, 337-350.