Multilevel Models Project

Working paper

24 June 1997

Non parametric 2-level models

Aitkin (1996) describes a 2-level nonparametric model where, instead of a continuous (Normal) distribution at level 2 for the intercept random effects we assume a finite number of clusters and maximise the resulting likelihood. We can write the likelihood as follows

$$L(\beta,\alpha_1,....,\alpha_K,p_1,....p_K \mid y)=\prod_i^n\sum_{k=1}^K p_k f(y\mid\beta,\alpha_k) \qquad (1)$$

where p_k is the mass or probability associated with the *k*-th cluster and α_k is the effect for the *k*-th cluster, k=1,….,K.

We can write the linear predictor for the *k*-th cluster as

$$\eta = X\beta + \alpha_k \qquad (2)$$

where α_k is the coefficient of the dummy variable defining the *k*-th cluster.

The likelihood equations to be solved are given by

$$\sum_i\sum_k w_{ik}s_{ik}(\beta^*)=0, \quad w_{ik}=\frac{p_k f_{ik}}{\sum_l p_l f_{il}}$$
$$\beta^*=\begin{pmatrix}\beta\\ \alpha\end{pmatrix}, \quad \alpha^T=(\alpha_1,...,\alpha_K) \qquad (3)$$

The term $s_{ik}(\beta^*)$ is the usual derivative wrt to β^*. We therefore proceed as in a standard model but with the additional weights for the level 1 units w_{ik}. We incorporate the weights by multiplying the random part explanatory variables by $w_{ik}^{-1/2}$. Note that the procedure can be used for complex level 1 variance models. This applies whether we are fitting Normal linear models or generalised linear models. The weights need to be updated at each iteration from (3), using the appropriate density function (e.g. Normal or binomial). Starting from equal weights for each cluster (there will be a dummy variable of weights for each cluster) we estimate (2) and the probabilities given by

$$\hat{p}_k=\sum_i\frac{w_{ik}}{n} \qquad (4)$$

The number of clusters has to be chosen but can be varied. It is not clear how this can be extended to more than 2 levels although an extension to a random coefficient model does seem possible.

For example, for a Normal model at level 1, at iteration m we have

$$f_{ik}^{(m)} \propto \exp\left(-\frac{(y_{ik}-\eta_{ik})^2}{2\sigma_e^2}\right)$$

and for a binary response

$$f_{ik}^{(m)} = \frac{e^{\eta}}{(1+e^{\eta})^2}$$

Starting values are needed and these can be obtained from an OLS analysis for the parameters. At convergence we obtain the mass point probabilities $\hat{p}_k$ and corresponding level 1 unit weights., with the linear predictor given by (2).

Harvey Goldstein

6-Mar-97