Class counts for between level latent class variable

May 8, 2020

In Twolevel Mixture analysis with a between level latent class variable, the most likely latent class variable is computed for each cluster as the latent class for which the posterior probability (i.e. the probability that the cluster belongs to a given class conditional on the observed data and the estimated model) is the largest. The most likely latent class variable can be obtained using the command

SAVEDATA: FILE=1.DAT; SAVE=CPROB;

Since the latent class variable is a between level variable, the most likely latent class variable is also a between level variable, i.e., it is the same for all the observations in the cluster.

Currently, the total number of observations with most likely latent class variable in a particular class, which is not the same as the distribution of the most likely latent class variable, can be found in the output under the heading

CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES BASED ON THEIR MOST LIKELY LATENT CLASS MEMBERSHIP

These quantities are computed based on all the observations in a particular class (not based on the number of clusters in a particular class), which in certain situations is not what is needed. One can obtain the most likely latent class variable distribution in terms of clusters that are classified in a particular class by computing it directly using the 1.DAT file in a simple Excel spread sheet. A similar approach can be used to also obtain the posterior

probability averages reported in this section

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES BASED ON ESTIMATED POSTERIOR PROBABILITIES

which currently are also based on observations within the clusters rather than the clusters. Note that if the size of the clusters is the same, the cluster level results (in terms of distribution) would be identical to those for the observation level results.