

October 23, 2007

Computer Lab for LCA in Mplus

Goals for today's lab

1. To run a 2, 3 and 4 class latent class model on the parental health related norms data.
2. To learn how to interpret the output of an LCA

-
1. Download the data from the class data page <http://www.biostat.umn.edu/~melanie/PH7435/DATA/index.html> into a folder on your PC. The dataset is listed under "Parent health related norms" and is called eatdata.dat (actually the file attached to link is called "allvarsnomissing.dat"). Also download the Mplus program called "parentformplusdemo.inp".
 2. Open Mplus and read in the program "parentformplusdemo.inp"

```
file is "c:\mplusatroot\mexamples\multiplecauseobesity\allvarsnomissing.dat";

variable: names are
ID GENDER RACE tloverwt numcohor
momh1ch momfit momdiet mom_myfd mom_nypa
mom_mywt dadh1ch dadfit daddiet dad_myfd
dad_nypa dad_mywt veghome vegdin junkhome
juichome milkmeal chiphome chochome pophome
YRFASTED YRLITTLE YRPILLS YRVOMIT YRLANA YRDIURET
YRSUBST YRSKIP YRSKOKED SES_3 tloverwt
white black hisp asian seshigh seemid;

usevariables are
momh1ch momfit momdiet mom_myfd mom_nypa mom_mywt;
idvariable = id;

categorical are
momh1ch momfit momdiet mom_myfd mom_nypa
mom_mywt;

classes=cph(2) ;

analysis: type = mixture;
starts = 300 20;
stiterations = 30;

savedata: file is "c:\mplusatroot\mexamples\multiplecauseobesity\
lca4estimates.dat";
save = cprobabilities;
```

3. Change the "file is" path so that it points to the location on your machine where you have put the allvarsnomissing.dat data. Also change the path in the "savedata" command line so that it points to a directory that exists on your machine.

4. There are many variables in this dataset but we will focus only on the `momhlth`, `momfit`, `momdiet`, `mom_myfd`, `mom_mypa`, and `mom_mywt`. These variables contain the yes/no responses of the adolescent girls to the following 6 questions:

My mother...

- | | |
|---|---------------------------|
| a. Cares about eating healthy food | (<code>momhlth</code>) |
| b. Cares about staying fit and exercising | (<code>momfit</code>) |
| c. Diets to lose weight or keep from gaining weight | (<code>momdiet</code>) |
| d. Encourages me to eat healthy foods | (<code>mom_myfd</code>) |
| e. Encourages me to be physically active | (<code>mom_mypa</code>) |
| f. Encourages me to diet to control my weight | (<code>mom_mywt</code>) |
5. Run the Mplus program. The Mplus program performs a latent class analysis on these 6 dichotomous variables and fits 2 underlying classes. The lines of code that tell it to do a latent class analysis are the “`classes=cph(2)`” command (which specifies the name of the latent class “`cph`” and how many categories it should have “2”) and the “`analysis: type = mixture`” command which tells Mplus to fit a mixture model (a general form of a latent class model).
 6. There are two main types of parameters of interest to examine from the output...1. the overall class membership probabilities, and 2. the conditional probability of a “yes” response to a question given membership in a particular class. Find these parameters in the output and try to interpret them. Note, Mplus labels the observed “0 = no”, “1 = yes” responses as Category 1 and 2 respectively. Also note, Mplus outputs overall class probabilities “based on the model” as well as “posterior probabilities” as well as the “most likely latent class membership”. You’ll note the first two are the same, these are the estimated population percents. The output labeled “most likely class membership” gives the percents based on partitioning the particular observed sample, where each individual is chosen to be in one and only one class (will examine this again in part 8).
 7. Rerun the model for 3 and then 4 latent classes. Interpret the different models, that is, interpret how the models split the girls up into different groups. What percent are in each group and how would you describe the different groups.
 8. After running the 4 latent classes model, scan through the outputted dataset “`lca4estimates.dat`”. The description of the columns of this dataset can be found at the end of the Mplus Output under “Savedata Information”. This dataset contains the “`cprobabilities`” which are the posterior probabilities of class membership and also it designates the most likely class membership. Indicate at least one pattern of observed responses that end up having a high probability of being in class 1,2,3 and 4 (there will be a different observed pattern for each of the different classes).