Data Summary, Model Information, and Fit Statistics (EM Algorithm)

There were no missing data in this example, so the number of subjects in the dataset and analysis are the same.

Number of subjects in dataset: 2587

Number of subjects in analysis: 2587

Number of measurement items: 7

Response categories per item: 2 2 2 2 2 2 2

The program echoes back some information so that you can check the number of indicators, response categories, groups, and classes. It also reminds you that you used a random seed to generate a set of random starting values.

Number of groups in the data: 1

Number of latent classes: 4

Rho starting values were randomly generated (seed = 4893).

No parameter restrictions were specified (freely estimated).

The model converged in 108 iterations.

Maximum number of iterations: 5000

Convergence method: maximum absolute deviation (MAD)

Convergence criterion: 0.000001000

The program reminds you that you did not have any parameter restrictions on the model. For this model, the EM algorithm converged in 108 iterations. Finally, it also reminds you the settings for the convergence criteria. Be sure to check the SAS log to make sure your model converged!

=============================================

Fit statistics:

=============================================

Log-likelihood: -6512.18

G-squared: 253.06

AIC: 315.06

BIC: 496.66

CAIC: 527.66

Adjusted BIC: 398.17

Entropy: 0.93

Degrees of freedom: 96

Test for MCAR

Model fit and selection criteria and the degrees of freedom. When models with different numbers of latent classes are compared, this information helps you select the optimal model for interpretation and further analysis.

Log-likelihood: -6385.65

G-squared: 222.25

Degrees of freedom: 514

Parameter Estimates

Because this model has parameters estimated to be on the boundary (e.g, 0.000 or 1.000 for rho parameters), standard errors could not be computed. The “RHO PRIOR = 1” option can be specified to use a small, data-driven prior to enable standard error estimation. See the users’ guide for more information.

(Standard errors could not be computed; please see the log file for details.)

Gamma estimates (class membership probabilities):

Class: 1 2 3 4

0.1423 0.5197 0.2447 0.0932

Rho estimates (item response probabilities):

Response category 1:

Class: 1 2 3 4

LIFETIME : 1.0000 0.1199 1.0000 0.3097

PREV\_YR : 0.9016 0.0000 1.0000 0.0000

PREV\_MO : 0.2611 0.0000 0.7339 0.0000

Estimated prevalences for the latent classes, also known as “latent class membership probabilities” or “gamma parameters.” These are the proportions of the population expected to be members of the classes. For example, 14.23% of the population are expected to be in Class 1. Here, Class 2 is the largest and Class 4 the smallest.

NEXT\_MO : 0.3058 0.0159 0.8749 0.2010

APRV\_TRY : 0.6672 0.1658 1.0000 0.9807

APRV\_OCC : 0.1978 0.0054 0.9975 0.9850

APRV\_REG : 0.0206 0.0019 0.6077 0.4254

Response category 2:

Class: 1 2 3 4

LIFETIME : 0.0000 0.8801 0.0000 0.6903

PREV\_YR : 0.0984 1.0000 0.0000 1.0000

PREV\_MO : 0.7389 1.0000 0.2661 1.0000

NEXT\_MO : 0.6942 0.9841 0.1251 0.7990

APRV\_TRY : 0.3328 0.8342 0.0000 0.0193

APRV\_OCC : 0.8022 0.9946 0.0025 0.0150

APRV\_REG : 0.9794 0.9981 0.3923 0.5746

Each latent class is interpreted based on the pattern of responding, as indicated by the item-response probabilities. Here are some suggested class labels:

Class 1 = Past-year users who approve of trying marijuana. This class has high probabilities of lifetime and past year marijuana use, but low probabilities of past month use and plans to use next month. It also has a high probability of approval of trying marijuana, but low probabilities of approval of occasional and regular use of marijuana.

Class 2 = Non-users who do not approve of marijuana use. This class has low probabilities of marijuana use, and low probabilities of any kind of approval of marijuana use.

Class 3 = Users who approve of regular marijuana use. This class has high probabilities of past month marijuana use and plans to use next month, as well as high probabilities for approval of any kind of marijuana use.

Class 4 = Non-users who approve of occasional marijuana use. This class has low probabilities of marijuana use, but high probabilities of approval of trying and occasional marijuana use.

Notice that our suggested class labels reflect both dimensions of the latent class variable: marijuana use and approval of marijuana use.

Estimated item-response probabilities, also known as “rho parameters.” They are conditional on class membership. They indicate how likely class members are to give particular responses to the items. For example, because response category 1 means “use” for the “LIFETIME” item asking about frequency of use over the participant’s lifetime (see the codebook provided in the LC TC Exercise document), participants in Class 1 have a 100% chance of responding ‘yes’ they had used alcohol in their lifetimes. The rho parameters are used to interpret the classes.