**Using LCA to group students pathways throughout the course (weeks 2-13)**

LCA was performed on the dataset (N = 290) consisting of 12 categorical variables, one for each week of the course (week 2 – week 13), representing the cluster each student belonged to in the given week. Variables for weeks 3, 5, 7, 9, 11 and 12 have 4 distinct values, as 4 clusters were detected for those weeks; variables for the other 6 weeks have 5 different values, matching the 5 identified clusters.

To avoid running into local maximum, the model building process was repeated 50 times (using different initial parameter values) for each considered number of classes (3-7). Table 1 shows evaluation metrics. The obtained metric values suggest the model with 6 classes as the best one (lowest values for AIC and BIC).

Table 1. Evaluation metrics for different number of classes

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

nclass AIC BIC LogLike ChiSquare

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

3 8559.303 9029.047 -4151.651 61200378

4 8486.923 9114.473 -4072.461 41745235

5 8463.473 9248.827 -4017.736 59853096

**6 8439.807 9382.967 -3962.904 11169413**

7 8448.175 9549.139 -3924.087 17333040

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

**Results for the solution with 6 classes**

**Interpretation of the classes**

The most probable path for students in **Class 1** (29.7%):

W2: C5(.36)|C4(.32) -> W3: C4(.4)|C1(.3)|C3(.3) -> W4: C3(.46)|C2(.25)|C5(.24) -> W5: C2(.55)|C4(.33) -> W6: C1(.45)|C2(.30) -> W7: C2(.59)|C4(.29) -> W8: C3(.41)|C1(.39) -> W9: C4(.58)|C2(.39) -> W10: C1(.68)|C4(.22) -> W11: C2(.82) -> W12: C4(.48)|C3(.45) -> W13: C1(.53)|C4(.20)

The most probable path for students in **Class 2** (20.3%):

W2: C2(.44)|C3(.28) -> W3: C3(.51)|C2(.43) -> W4: C2(.63)|C1(.37) -> W5: C1(.72) -> W6: C2(.34)|C4(.25)|C3(.20) -> W7: C1(.75) -> W8: C5(.52)|C2(.33) -> W9: C3(.77) -> W10: C3(.64)|C4(.21) -> W11: C1(.84) -> W12: C1(.83) -> W13: C2(.31)|C3(.26)|C4(.24)

The most probable path for students in **Class 3** (19%):

W2: C5(.35)|C1(.35) -> W3: C1(.54)|C4(.28) -> W4: C4(.34)|C2(.28)|C5(.20) -> W5: C4(.72)|C2(.22) -> W6: C2(.58) -> W7: C4(.57)|C3(.26) -> W8: C4(.60)|C3(.23) -> W9: C2(.78) -> W10: C2(.66)|C4(.26) -> W11: C3(.48)|C2(.40) -> W12: C2(.55)|C4(.38) -> W13: C1(.66)

The most probable path for students in **Class 4** (16%):

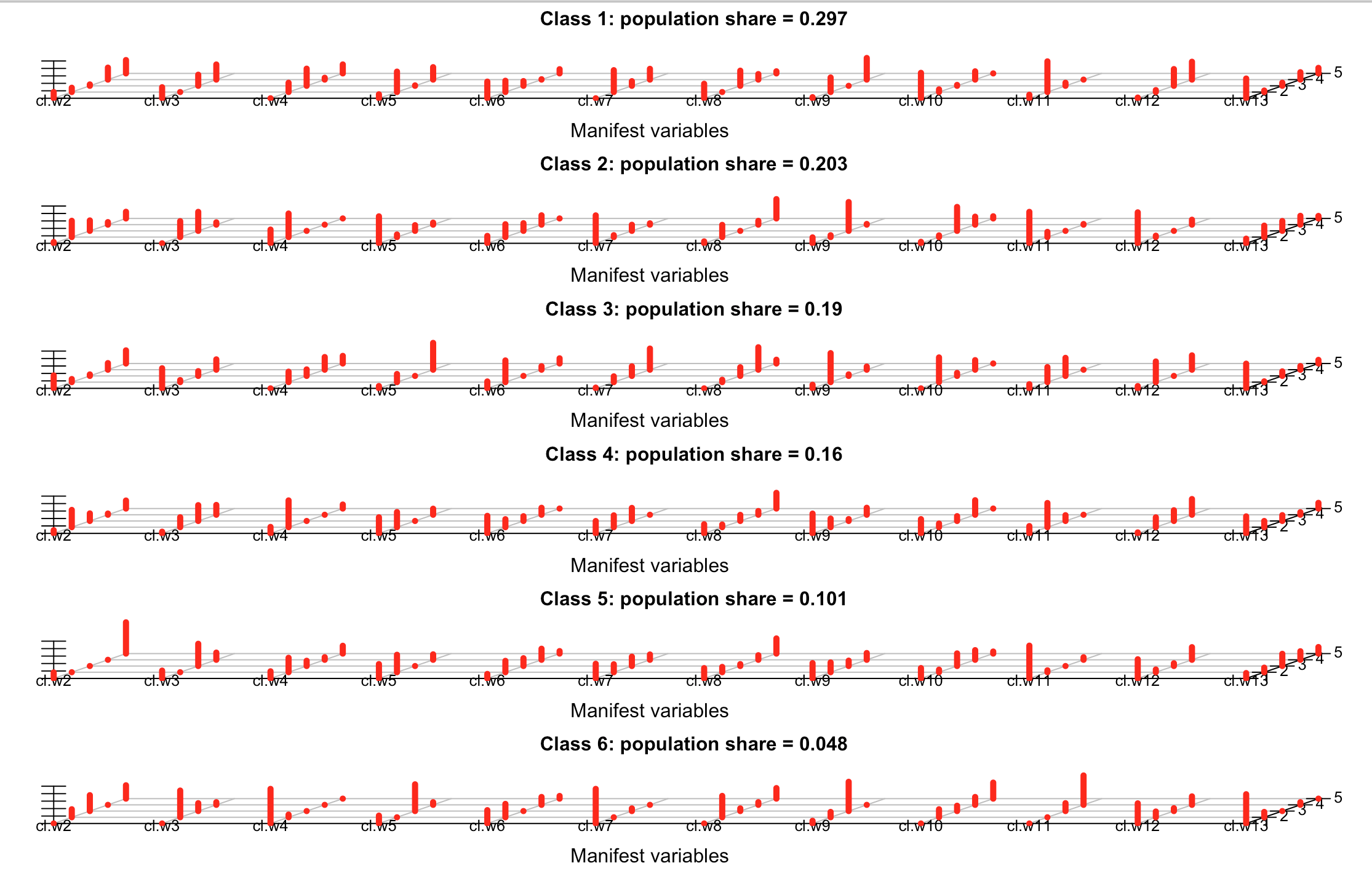
W2: C2(.46)|C5(.22)|C3(.20) -> W3: C3(.42)|C2(.27)|C4(.26) -> W4: C2(.72) -> W5: C1(.43)|C2(.40) -> W6: C1(.47)|C2(.21) -> W7: C3(.35)|C1(.33)|C2(.32) -> W8: C5(.42)|C1(.24) -> W9: C1(.53)|C2(.23) -> W10: C4(.39)|C1(.38) -> W11: C2(.65)|C1(.20) -> W12: C4(.42)|C3(.29)|C2(.26) -> W13: C1(.46)

The most probable path for students in **Class 5** (10.1%)

W2: C5(.84) ->W3: C3(.60)|C1(.21) -> W4: C2(.39)|C5(.21) -> W5: C2(.47)|C1(.38) -> W6: C2(.30)|C4(.30)|C3(.21) -> W7: C1(.39)|C3(.25)|C2(.21) -> W8: C5(.41)|C1(.28) -> W9: C1(.41)|C2(.26) -> W10: C3(.34)|C1(.28)|C4(.26) -> W11: C1(.88) -> W12: C1(.51)|C4(.34) -> W13: C3(.33)|C4(.23)

The most probable path for students in **Class 6** (4.8%)

W2: C3(.43)|C5(.36)|C2(.22) -> W3: C2(.72)|C3(.21) -> W4: C1(.93) -> W5: C3(.72)|C1(.21) -> W6: C1(.36)|C2(.36)|C4(.21) -> W7: C1(.93) -> W8: C2(.57)|C5(.28) -> W9: C3(.78) -> W10: C5(.43)|C2(.22)|C4(.21) -> W11: C4(.79)|C3(.21) -> W12: C1(.54)|C4(.23) -> W13: C1(.78)

****

**INTERPRETATION BASED ON THE CODING SCHEME**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **CLASS 1** | **CLASS 2** | **CLASS 3** | **CLASS 4** | **CLASS 5** | **CLASS 6** |
| **WEEK 2** | C1|B1 | D|E | C1|A | D|C1|E | C1 | E|C1|D |
| **WEEK 3** | B2|A|D | D|E | A|B2 | D|E|B2 | D|A | E|D |
| **WEEK 4** | B1|C2|B2 | C2|E | A|C2|B2 | C2 | C2|B2 | E |
| **WEEK 5** | B1|A | D | A|B1 | D|B1 | B1|D | E|D |
| **WEEK 6** | C2|C1 | C1|D|E | C1 | C2|C1 | C1|D|E | C2|C1|D |
| **WEEK 7** | B1|C1 | E | C1|D | D|E|B1 | E|D|B1 | E |
| **WEEK 8** | B2|B1 | D|E | A|B2 | D|B1 | D|B1 | E|D |
| **WEEK 9** | B1|A | E | A | D|A | D|A | E |
| **WEEK 10** | B1|B2 | D|B2 | A|B2 | B2|B1 | D|B1|B2 | E|A|B2 |
| **WEEK 11** | B1 | D | F|B1 | B1|D | D | E|F |
| **WEEK 12** | B2|B1 | E | F|B2 | B2|B1|F | E|B1 | E|B2 |
| **WEEK 13** | A|C1 | E|D|C1 | A | A | D|C1 | A |

**The abbreviations used in the class-comparisons table**:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **EQT** | **EXC\*** | **VEQ** | **VID** | **ORG** | **DBOARD** | **GDEN** |
| A | No events | Sporadic low number of events | No events | No events | No events | No events | Zero |
| B1 | No event | Numerous attempts. Almost all incorrect | No events | No events | No events | No events | Negligible |
| B2 | No events | Numerous attempts. Mostly incorrect | No events | Slight or no interaction | No events | No events | Low or negligible |
| C1 | Moderate interaction | Moderate attempts. Correct and incorrect balance | Very low engagement | Low engagement | Very high before midterm | No events | Medium |
| C2 | Moderate interaction | Moderate attempts. Correct and incorrect balance | Low engagement | High engagement | No events | No events | Medium |
| D | Large to very large engagement | Numerous attempts. Mostly correct | Moderate engagement | Moderate engagement | No events | Only in two weeks | High |
| E | Very large engagement | High number of attempts. Correct and incorrect balance | Very high engagement | Very high engagement | Very high in week 2 | High engagement | High |
| F | No events | High numbers. Mostly correct | No events | No events | No events | No events | Zero |

**Comparison of LCA classes based on the students’ exam scores**

Since data about students’ exam scores are not normally distributed, non-parametric tests were performed: Kruskal-Wallis test followed by Mann-Whitney U test for pair-wise comparison.

**Midterm exam score**

Descriptive statistics

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

class N median Q1 Q3

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

1 87 12.0 9.00 15

2 60 16.0 13.00 17

3 56 13.0 10.00 15

4 45 15.0 12.00 18

5 28 15.0 13.00 18

6 14 13.5 12.25 16

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

Pairwise comparisons with the FDR correction

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

\ c1 c2 Z p effect.size significant

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

1 1 2 -4.6430 0.000002 0.3829 YES

3 1 4 -4.1430 0.000024 0.3606 YES

4 1 5 -3.6363 0.000204 0.3391 YES

6 2 3 3.2229 0.001138 0.2992 YES

10 3 4 -2.7403 0.005813 0.2727 YES

11 3 5 -2.6487 0.007590 0.2890 YES

5 1 6 -2.1215 0.033066 0.2111 NO

2 1 3 -1.3843 0.167100 0.1158 NO

12 3 6 -1.2902 0.200612 0.1542 NO

9 2 6 0.8882 0.380404 0.1033 NO

15 5 6 0.8324 0.413437 0.1284 NO

14 4 6 0.5819 0.567666 0.0758 NO

13 4 5 -0.3762 0.711797 0.0440 NO

7 2 4 0.3512 0.727802 0.0343 NO

8 2 5 -0.0405 0.969667 0.0043 NO

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

Significant differences with respect to the midterm exam score are detected between the following pairs of classes: 1-2, 1-4, 1-5, 2-3, 3-4, and 3-5.

**Final exam score**

Descriptive statistics

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

class N median Q1 Q3

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

1 87 14.0 11.00 18.00

2 60 21.5 17.00 32.25

3 56 16.0 13.00 22.25

4 45 21.0 15.00 27.00

5 28 24.0 13.75 31.00

6 14 15.5 11.75 24.25

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

Pairwise comparisons with the FDR correction

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

\ c1 c2 Z p effect.size significant

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

1 1 2 -5.8301 0.000000 0.4809 YES

3 1 4 -4.5020 0.000004 0.3918 YES

4 1 5 -3.7620 0.000118 0.3508 YES

6 2 3 3.6812 0.000189 0.3418 YES

11 3 5 -2.3657 0.017467 0.2581 NO

2 1 3 -2.2865 0.021945 0.1912 NO

9 2 6 2.2323 0.024672 0.2595 NO

10 3 4 -2.2203 0.026065 0.2209 NO

15 5 6 1.5764 0.117349 0.2432 NO

14 4 6 1.4811 0.141075 0.1928 NO

5 1 6 -1.4035 0.162846 0.1397 NO

7 2 4 1.1767 0.241097 0.1148 NO

13 4 5 -0.5849 0.562859 0.0685 NO

8 2 5 0.4350 0.667024 0.0464 NO

12 3 6 0.0000 1.000000 0.0000 NO

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

To conclude, significant differences with respect to the final exam score are detected between the following pairs of classes: 1-2, 1-4, 1-5, and 2-3

Raw output obtained from poLCA

Conditional item response (column) probabilities, by outcome variable, for each class (row)

$cl.w2

Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)

class 1: 0.1670 0.1154 0.0378 0.3215 0.3583

class 2: 0.0343 0.4373 0.2803 0.0685 0.1796

class 3: 0.3469 0.0830 0.0401 0.1776 0.3523

class 4: 0.0859 0.4632 0.2031 0.0325 0.2153

class 5: 0.1611 0.0000 0.0000 0.0000 0.8389

class 6: 0.0000 0.2154 0.4262 0.0000 0.3584

$cl.w3

Pr(1) Pr(2) Pr(3) Pr(4)

class 1: 0.2956 0.0000 0.2985 0.4059

class 2: 0.0000 0.4270 0.5104 0.0626

class 3: 0.5392 0.0552 0.1281 0.2775

class 4: 0.0459 0.2667 0.4243 0.2632

class 5: 0.2105 0.0000 0.5966 0.1930

class 6: 0.0000 0.7169 0.2117 0.0713

$cl.w4

Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)

class 1: 0.0000 0.2492 0.4605 0.0498 0.2406

class 2: 0.3681 0.6319 0.0000 0.0000 0.0000

class 3: 0.0000 0.2809 0.1748 0.3399 0.2043

class 4: 0.1740 0.7181 0.0000 0.0000 0.1079

class 5: 0.1916 0.3874 0.1410 0.0702 0.2098

class 6: 0.9281 0.0719 0.0000 0.0000 0.0000

$cl.w5

Pr(1) Pr(2) Pr(3) Pr(4)

class 1: 0.1062 0.5480 0.0119 0.3339

class 2: 0.7231 0.0694 0.1530 0.0545

class 3: 0.0623 0.2182 0.0000 0.7195

class 4: 0.4322 0.4057 0.0000 0.1621

class 5: 0.3851 0.4686 0.0000 0.1464

class 6: 0.2115 0.0000 0.7172 0.0713

$cl.w6

Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)

class 1: 0.4465 0.3032 0.1268 0.0106 0.1130

class 2: 0.1988 0.3448 0.2057 0.2507 0.0000

class 3: 0.1854 0.5819 0.0000 0.0937 0.1390

class 4: 0.4709 0.2141 0.1280 0.1870 0.0000

class 5: 0.1208 0.3052 0.2095 0.2972 0.0673

class 6: 0.3579 0.3588 0.0000 0.2115 0.0718

$cl.w7

Pr(1) Pr(2) Pr(3) Pr(4)

class 1: 0.0000 0.5867 0.1242 0.2891

class 2: 0.7488 0.0516 0.1637 0.0359

class 3: 0.0200 0.1557 0.2571 0.5673

class 4: 0.3321 0.3176 0.3503 0.0000

class 5: 0.3935 0.2134 0.2505 0.1426

class 6: 0.9282 0.0000 0.0718 0.0000

$cl.w8

Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)

class 1: 0.3886 0.0000 0.4065 0.1488 0.0562

class 2: 0.0515 0.3275 0.0000 0.1021 0.5189

class 3: 0.0000 0.0637 0.2326 0.6024 0.1013

class 4: 0.2453 0.0706 0.1631 0.1000 0.4210

class 5: 0.2770 0.1426 0.0510 0.1231 0.4063

class 6: 0.0000 0.5734 0.0718 0.0718 0.2830

$cl.w9

Pr(1) Pr(2) Pr(3) Pr(4)

class 1: 0.0290 0.3907 0.0000 0.5803

class 2: 0.1583 0.0498 0.7734 0.0185

class 3: 0.0967 0.7781 0.0415 0.0837

class 4: 0.5310 0.2305 0.0508 0.1876

class 5: 0.4110 0.2626 0.1446 0.1819

class 6: 0.0719 0.1431 0.7850 0.0000

$cl.w10

Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)

class 1: 0.6819 0.0717 0.0241 0.2222 0.0000

class 2: 0.0389 0.0461 0.6421 0.2144 0.0585

class 3: 0.0000 0.6628 0.0804 0.2568 0.0000

class 4: 0.3829 0.0980 0.1292 0.3899 0.0000

class 5: 0.2782 0.0692 0.3380 0.2594 0.0552

class 6: 0.0000 0.2152 0.1433 0.2114 0.4301

$cl.w11

Pr(1) Pr(2) Pr(3) Pr(4)

class 1: 0.0943 0.8230 0.0827 0.0000

class 2: 0.8433 0.1395 0.0000 0.0172

class 3: 0.1120 0.4035 0.4845 0.0000

class 4: 0.1963 0.6487 0.1550 0.0000

class 5: 0.8801 0.0515 0.0000 0.0684

class 6: 0.0000 0.0000 0.2117 0.7883

$cl.w12

Pr(1) Pr(2) Pr(3) Pr(4)

class 1: 0.0184 0.0525 0.4494 0.4798

class 2: 0.8305 0.0370 0.0000 0.1326

class 3: 0.0611 0.5543 0.0000 0.3846

class 4: 0.0270 0.2635 0.2879 0.4216

class 5: 0.5143 0.0733 0.0735 0.3389

class 6: 0.5408 0.1548 0.0773 0.2272

$cl.w13

Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)

class 1: 0.5291 0.0406 0.0833 0.1989 0.1481

class 2: 0.1242 0.3073 0.2642 0.2369 0.0674

class 3: 0.6621 0.0213 0.1217 0.0977 0.0972

class 4: 0.4567 0.1660 0.1563 0.0683 0.1527

class 5: 0.1443 0.1230 0.3275 0.2294 0.1758

class 6: 0.7851 0.1432 0.0000 0.0717 0.0000

Estimated class population shares

0.2972 0.203 0.1904 0.16 0.1013 0.048

Predicted class memberships (by modal posterior prob.)

0.3 0.2069 0.1931 0.1552 0.0966 0.0483

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

Fit for 6 latent classes:

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

number of observations: 290

number of fully observed cases: 239

number of estimated parameters: 257

residual degrees of freedom: 33

maximum log-likelihood: -3970.409

AIC(6): 8454.819

BIC(6): 9397.978

G^2(6): 4121.598 (Likelihood ratio/deviance statistic)

X^2(6): 11423087 (Chi-square goodness of fit)