2019-01-29- LCA data modeling Seth-Josh

1. Loading, setting up

```r
library(tidyverse)
library(poLCA)
library(readxl)
```

Getting data from Google Sheets

```r
library(googlesheets)
library(readr)

g <- gs_title("Observations_segment_Units_1-7_2013-14-with-duplicates-identified")
d <- gs_read(g, col_types =
                  readr::cols(
  `ClassObservation::Observer` = col_character(),
  `ClassObservation::ObsNickname` = col_double(),
  `Teacher::TeacherID` = col_double(),
  `Teacher::First Name` = col_character(),
  `Teacher::Last Name` = col_character(),
  `Teacher::Condition` = col_character(),
  `ClassObservation::Unit` = col_double(),
  `ClassObservation::Date` = col_datetime(format = ""),
  Notes = col_character(),
  ObsNN = col_double(),
  SegNum = col_double(),
  `Segment::StartStamp` = col_datetime(format = ""),
  `Segment::EndStamp` = col_datetime(format = ""),
  fWhole = col_double(),
  fGroups = col_double(),
  fSeat = col_double(),
  sInvented = col_double(),
  sConceptual = col_double(),
  sProcedural = col_double(),
  sEngagement = col_character(),
  tInitSelect = col_double(),
  tCompare = col_double(),
  tDiscussQ = col_double(),
  tPressExplain = col_double(),
  tConnectOthers = col_double(),
  tConnectBigIdeas = col_double(),
  tConventional = col_double(),
  tProcedural = col_double(),
  iPrecision = col_double(),
  iCenter = col_double(),
  iDIsplay = col_double(),
```

```r
    iOther = col_double(),
    iOrder = col_double(),
    iScale = col_double(),
    iGrouping = col_double(),
    iShape = col_double(),
    iShow = col_double(),
    iHide = col_double(),
    iMode = col_double(),
    iMedian = col_double(),
    iMean = col_double(),
    iRange = col_double(),
    iCenterClump = col_double(),
    iDeviation = col_double(),
    iReplicability = col_double(),
    iGeneralizability = col_double(),
    iLinkVisDist = col_double(),
    iLinkImagDist = col_double(),
    ITheoreticalProb = col_double(),
    IEmpiricalProb = col_double(),
    IOdds = col_logical(),
    ISampleSize = col_double(),
    ISamplingDistrib = col_double(),
    ICenterStats = col_double(),
    IVariabilityStats = col_double(),
    `Segment::iIntelligibility` = col_double(),
    `Segment::iModelFit` = col_double(),
    `Segment::iDistribution` = col_double(),
    `Segment::iRandomComponents` = col_double(),
    `Segment::iNonRandomComponents` = col_double(),
    `Segment::iMedianDistr` = col_double(),
    `Segment::iIQRDistr` = col_logical(),
    `Segment::iNewMedian` = col_double(),
    `Segment::iNewIQR` = col_logical(),
    `Segment::iRegions` = col_double(),
    `Segment::iQuantRegions` = col_double(),
    number_of_segments = col_double(),
    `Duplicate Condition` = col_character()
))

d <- dplyr::rename(d, condition = `Teacher::Condition`)

d <- d %>%
    mutate(condition = ifelse(str_detect(condition, "1"), 1,
                              ifelse(str_detect(condition, "2") | condition == 0, 0, NA)))

library(readxl)
u <- read_xlsx("Observations_summary_Units_1-7_2012-13-mod.xlsx")
g1 <- gs_title("Observations_segment_Units_1-7_2012-13-with-duplicates-identified")
d1 <- gs_read(g1)
d1 <- unite(d1, Teacher, `Teacher::First Name`, `Teacher::Last Name`, sep = " ")
d1 <- d1 %>% left_join(u, by = "Teacher")
d1 <- rename(d1, condition = Group)
```

```r
add_one <- function(x) {
    x + 1
}

ds <- d %>%
    dplyr::select(sInvented, sProcedural, sConceptual, tInitSelect, tCompare, tDiscussQ, tConnectBigIde
    map_df(replace_na, 0) %>%
    modify_if(is.numeric, add_one)

ds1 <- d1 %>%
    dplyr::select(sInvented, sProcedural, sConceptual, tInitSelect, tCompare, tDiscussQ, tConnectBigIde
    map_df(replace_na, 0) %>%
    modify_if(is.numeric, add_one)

dd <- bind_rows(ds, ds1)

dds <- filter(dd, `Duplicate Condition` != "D" & `Duplicate Condition` != "d")
```

3. Choosing the number of classes/profiles

Using latent class analysis through the **poLCA** R package.

```r
f <- cbind(sInvented, sProcedural, sConceptual, tInitSelect, tCompare, tDiscussQ, tConnectBigIdeas, tCom

od <- map(1:9, poLCA, formula = f, data = dds, maxiter = 5000, verbose = FALSE, graphs = FALSE) %>%
    map_df(broom::glance)
```
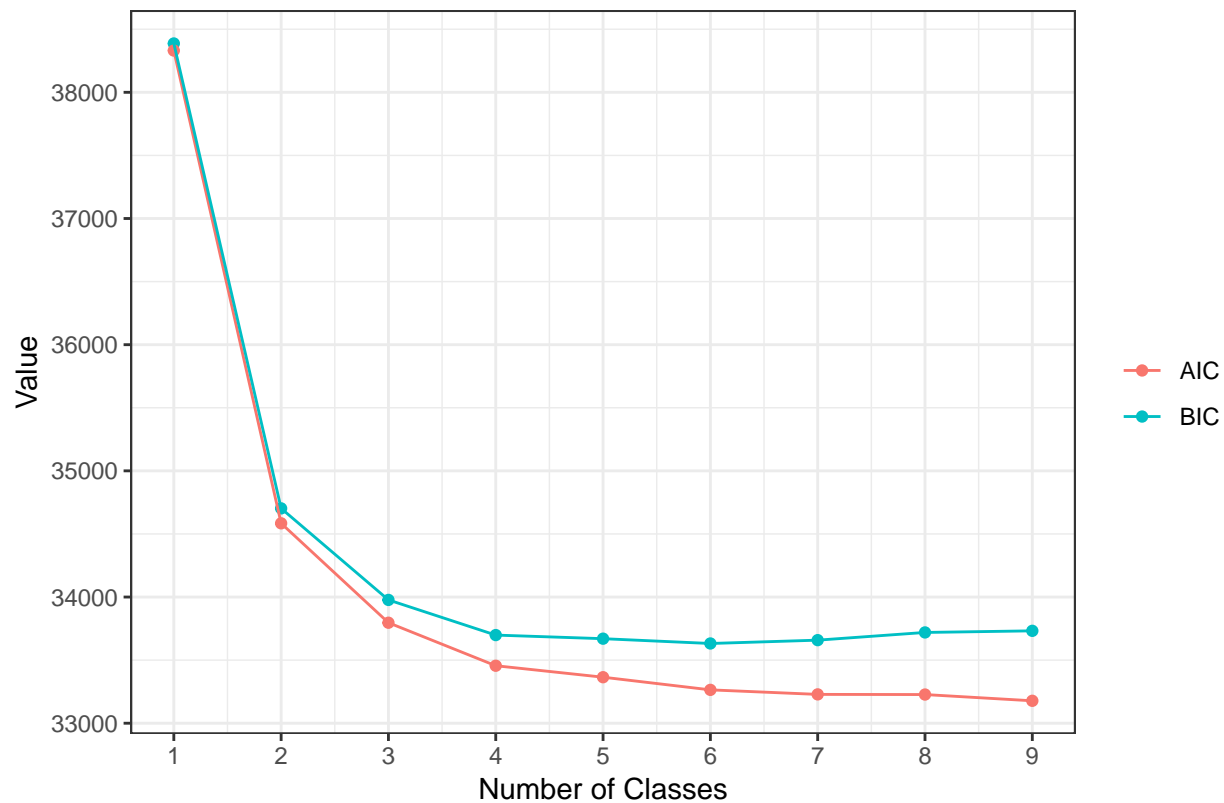
```r
od %>%
    mutate(n_classes = 1:9) %>%
    gather(key, val, BIC, AIC) %>%
    ggplot(aes(x = n_classes, y = val, color = key, group = key)) +
    geom_point() +
    geom_line() +
    scale_x_continuous(breaks = 1:9, labels = 1:9) +
    theme_bw() +
    labs(caption = "Lower values of the AIC & BIC suggest preferred model(s); generally, BIC is more co
    xlab("Number of Classes") +
    ylab("Value") +
    scale_color_discrete("")
```
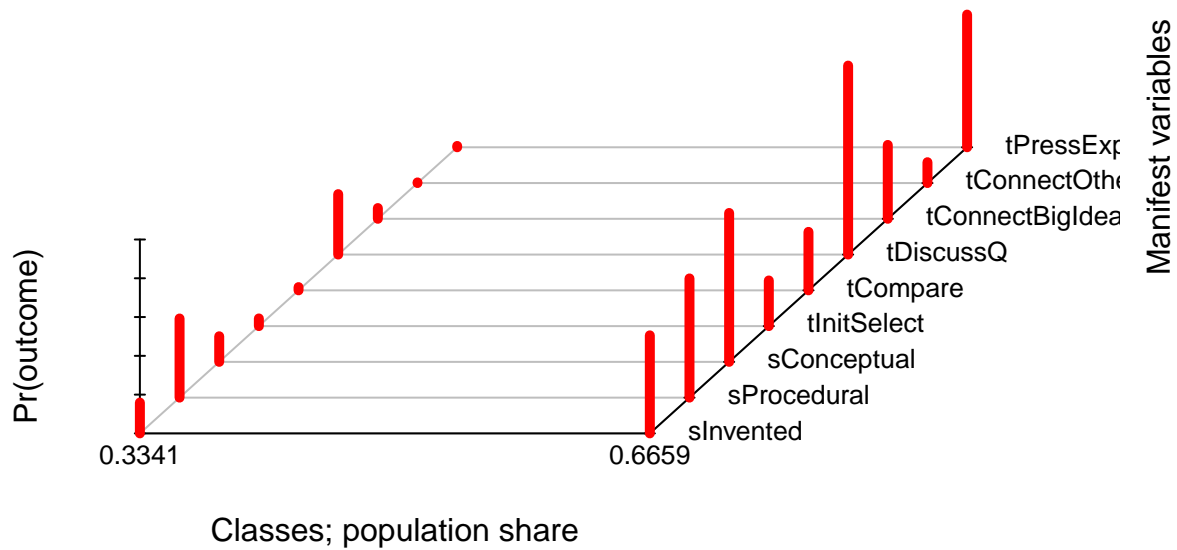
Lower values of the AIC & BIC suggest preferred model(s); generally, BIC is more conservative than AIC

Based on these fit statistics, a three or four class solution seems to exhibit the best fit, though a three-class solution may also be suitable; for comparison, a two-class solution is also explored.

4. Examining 2, 3, 4, and 5 class solutions

```
f <- cbind(sInvented, sProcedural, sConceptual, tInitSelect, tCompare, tDiscussQ, tConnectBigIdeas, tCo

m2 <- poLCA(f, dds, nclass = 2, maxiter = 10000, graphs = TRUE)
```

Pr(outcome)

Classes; population share

0.3341   0.6659

Manifest variables

tPressEx
tConnectOth
tConnectBigIdea
tDiscussQ
tCompare
tInitSelect
sConceptual
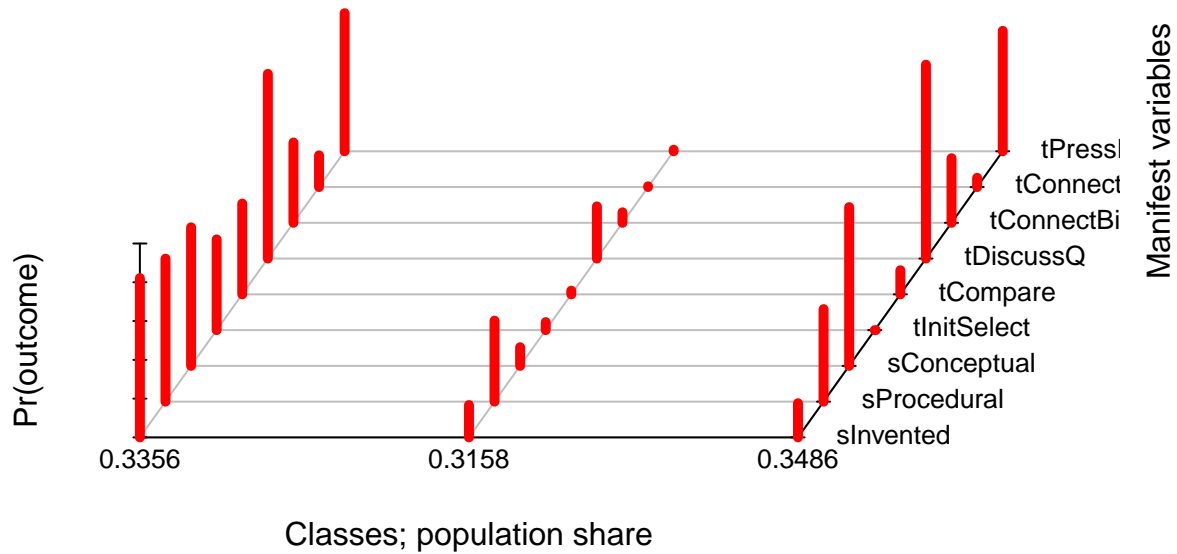sProcedural
sInvented

```
## Conditional item response (column) probabilities,
##  by outcome variable, for each class (row)
##
## $sInvented
##          Pr(1)  Pr(2)
## class 1:  0.8410 0.1590
## class 2:  0.4943 0.5057
##
## $sProcedural
##          Pr(1)  Pr(2)
## class 1:  0.5919 0.4081
## class 2:  0.3853 0.6147
##
## $sConceptual
##          Pr(1)  Pr(2)
## class 1:  0.8679 0.1321
## class 2:  0.2320 0.7680
##
## $tInitSelect
##          Pr(1)  Pr(2)
## class 1:  0.9605 0.0395
## class 2:  0.7645 0.2355
##
## $tCompare
##          Pr(1)  Pr(2)
## class 1:  0.9833 0.0167
```

```
## class 2:   0.6983 0.3017
##
## $tDiscussQ
##             Pr(1)  Pr(2)
## class 1:   0.6882 0.3118
## class 2:   0.0252 0.9748
##
## $tConnectBigIdeas
##             Pr(1)  Pr(2)
## class 1:   0.9444 0.0556
## class 2:   0.6189 0.3811
##
## $tConnectOthers
##             Pr(1)  Pr(2)
## class 1:   0.9967 0.0033
## class 2:   0.8924 0.1076
##
## $tPressExplain
##             Pr(1)  Pr(2)
## class 1:   0.9930 0.0070
## class 2:   0.3157 0.6843
##
## Estimated class population shares
##   0.3341 0.6659
##
## Predicted class memberships (by modal posterior prob.)
##   0.3213 0.6787
##
## ============================================================
## Fit for 2 latent classes:
## ============================================================
## number of observations: 3753
## number of estimated parameters: 19
## residual degrees of freedom: 492
## maximum log-likelihood: -17273.09
##
## AIC(2): 34584.18
## BIC(2): 34702.56
## G^2(2): 2000.883 (Likelihood ratio/deviance statistic)
## X^2(2): 5120.786 (Chi-square goodness of fit)
##
```

```r
m3 <- poLCA(f, dds, nclass = 3, maxiter = 10000, graphs = TRUE)
```

```
## Conditional item response (column) probabilities,
##  by outcome variable, for each class (row)
##
## $sInvented
##          Pr(1)  Pr(2)
## class 1:  0.1784 0.8216
## class 2:  0.8332 0.1668
## class 3:  0.8238 0.1762
##
## $sProcedural
##          Pr(1)  Pr(2)
## class 1:  0.2626 0.7374
## class 2:  0.5823 0.4177
## class 3:  0.5229 0.4771
##
## $sConceptual
##          Pr(1)  Pr(2)
## class 1:  0.2855 0.7145
## class 2:  0.9038 0.0962
## class 3:  0.1813 0.8187
##
## $tInitSelect
##          Pr(1)  Pr(2)
## class 1:  0.5323 0.4677
## class 2:  0.9587 0.0413
## class 3:  1.0000 0.0000
```

```
## 
## $tCompare
##            Pr(1)  Pr(2)
## class 1:  0.5317 0.4683
## class 2:  0.9809 0.0191
## class 3:  0.8758 0.1242
## 
## $tDiscussQ
##            Pr(1)  Pr(2)
## class 1:  0.0474 0.9526
## class 2:  0.7309 0.2691
## class 3:  0.0000 1.0000
## 
## $tConnectBigIdeas
##            Pr(1)  Pr(2)
## class 1:  0.5858 0.4142
## class 2:  0.9457 0.0543
## class 3:  0.6668 0.3332
## 
## $tConnectOthers
##            Pr(1)  Pr(2)
## class 1:  0.8364 0.1636
## class 2:  0.9962 0.0038
## class 3:  0.9523 0.0477
## 
## $tPressExplain
##            Pr(1)  Pr(2)
## class 1:  0.2878 0.7122
## class 2:  0.9921 0.0079
## class 3:  0.3788 0.6212
## 
## Estimated class population shares
##   0.3356 0.3158 0.3486
## 
## Predicted class memberships (by modal posterior prob.)
##   0.3288 0.3176 0.3536
## 
## =============================================================
## Fit for 3 latent classes:
## =============================================================
## number of observations: 3753
## number of estimated parameters: 29
## residual degrees of freedom: 482
## maximum log-likelihood: -16869.22
## 
## AIC(3): 33796.45
## BIC(3): 33977.13
## G^2(3): 1193.151 (Likelihood ratio/deviance statistic)
## X^2(3): 2276.576 (Chi-square goodness of fit)
## 
```
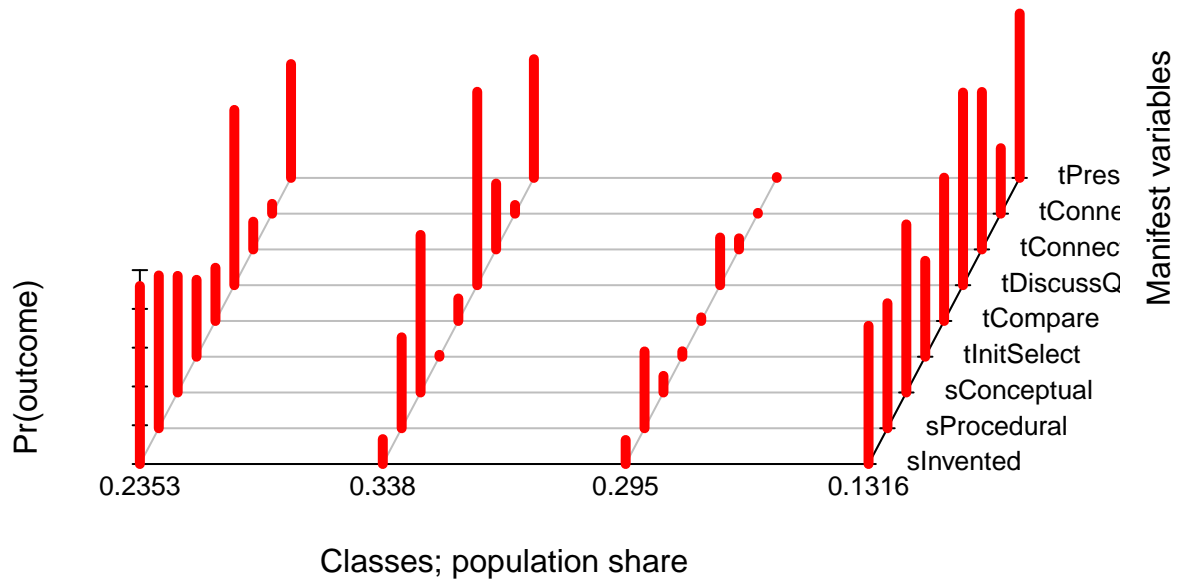
```
m4 <- poLCA(f, dds, nclass = 4, maxiter = 10000, graphs = TRUE)
```



```
## Conditional item response (column) probabilities,
##  by outcome variable, for each class (row)
##
## $sInvented
##            Pr(1)  Pr(2)
## class 1:  0.0818 0.9182
## class 2:  0.8716 0.1284
## class 3:  0.8761 0.1239
## class 4:  0.2874 0.7126
##
## $sProcedural
##            Pr(1)  Pr(2)
## class 1:  0.2120 0.7880
## class 2:  0.5310 0.4690
## class 3:  0.6039 0.3961
## class 4:  0.3553 0.6447
##
## $sConceptual
##            Pr(1)  Pr(2)
## class 1:  0.3986 0.6014
## class 2:  0.1874 0.8126
## class 3:  0.9144 0.0856
## class 4:  0.1331 0.8669
```

```
## 
## $tInitSelect
##           Pr(1)  Pr(2)
## class 1:  0.6033 0.3967
## class 2:  0.9903 0.0097
## class 3:  0.9720 0.0280
## class 4:  0.5053 0.4947
## 
## $tCompare
##           Pr(1)  Pr(2)
## class 1:  0.7255 0.2745
## class 2:  0.8842 0.1158
## class 3:  0.9814 0.0186
## class 4:  0.2609 0.7391
## 
## $tDiscussQ
##           Pr(1)  Pr(2)
## class 1:  0.0959 0.9041
## class 2:  0.0025 0.9975
## class 3:  0.7540 0.2460
## class 4:  0.0066 0.9934
## 
## $tConnectBigIdeas
##           Pr(1)  Pr(2)
## class 1:  0.8569 0.1431
## class 2:  0.6602 0.3398
## class 3:  0.9428 0.0572
## class 4:  0.1879 0.8121
## 
## $tConnectOthers
##           Pr(1)  Pr(2)
## class 1:  0.9493 0.0507
## class 2:  0.9552 0.0448
## class 3:  0.9959 0.0041
## class 4:  0.6622 0.3378
## 
## $tPressExplain
##           Pr(1)  Pr(2)
## class 1:  0.4128 0.5872
## class 2:  0.3876 0.6124
## class 3:  0.9956 0.0044
## class 4:  0.1523 0.8477
## 
## Estimated class population shares
##   0.2353 0.338 0.295 0.1316
## 
## Predicted class memberships (by modal posterior prob.)
##   0.2443 0.3429 0.2963 0.1164
## 
## ==========================================================
## Fit for 4 latent classes:
## ==========================================================
## number of observations: 3753
## number of estimated parameters: 39
```

```
## residual degrees of freedom: 472
## maximum log-likelihood: -16688.75
##
## AIC(4): 33455.5
## BIC(4): 33698.49
## G^2(4): 832.2073 (Likelihood ratio/deviance statistic)
## X^2(4): 1870.209 (Chi-square goodness of fit)
##
```

```
m5 <- poLCA(f, dds, nclass = 5, maxiter = 10000, graphs = TRUE)
```
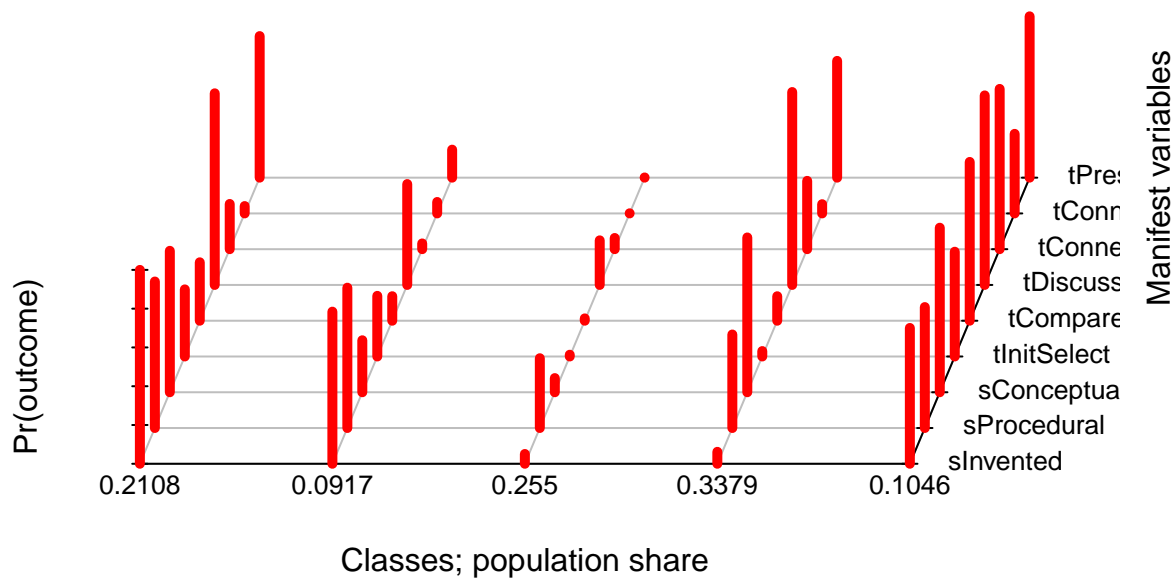


```
## Conditional item response (column) probabilities,
##  by outcome variable, for each class (row)
##
## $sInvented
##            Pr(1)  Pr(2)
## class 1:  0.0000 1.0000
## class 2:  0.2154 0.7846
## class 3:  0.9494 0.0506
## class 4:  0.9383 0.0617
## class 5:  0.2995 0.7005
##
## $sProcedural
##            Pr(1)  Pr(2)
## class 1:  0.2444 0.7556
```

```
## class 2:   0.2761 0.7239
## class 3:   0.6400 0.3600
## class 4:   0.5177 0.4823
## class 5:   0.3762 0.6238
##
## $sConceptual
##            Pr(1)  Pr(2)
## class 1:   0.2704 0.7296
## class 2:   0.7318 0.2682
## class 3:   0.9269 0.0731
## class 4:   0.2016 0.7984
## class 5:   0.1513 0.8487
##
## $tInitSelect
##            Pr(1)  Pr(2)
## class 1:   0.6538 0.3462
## class 2:   0.6871 0.3129
## class 3:   0.9908 0.0092
## class 4:   0.9720 0.0280
## class 5:   0.4595 0.5405
##
## $tCompare
##            Pr(1)  Pr(2)
## class 1:   0.6986 0.3014
## class 2:   0.8736 0.1264
## class 3:   0.9880 0.0120
## class 4:   0.8737 0.1263
## class 5:   0.1810 0.8190
##
## $tDiscussQ
##            Pr(1)  Pr(2)
## class 1:   0.0115 0.9885
## class 2:   0.4795 0.5205
## class 3:   0.7698 0.2302
## class 4:   0.0049 0.9951
## class 5:   0.0224 0.9776
##
## $tConnectBigIdeas
##            Pr(1)  Pr(2)
## class 1:   0.7666 0.2334
## class 2:   0.9702 0.0298
## class 3:   0.9420 0.0580
## class 4:   0.6474 0.3526
## class 5:   0.1732 0.8268
##
## $tConnectOthers
##            Pr(1)  Pr(2)
## class 1:   0.9619 0.0381
## class 2:   0.9408 0.0592
## class 3:   0.9992 0.0008
## class 4:   0.9523 0.0477
## class 5:   0.5894 0.4106
##
## $tPressExplain
```

```
##            Pr(1)  Pr(2)
## class 1:  0.2697 0.7303
## class 2:  0.8555 0.1445
## class 3:  0.9989 0.0011
## class 4:  0.3976 0.6024
## class 5:  0.1678 0.8322
##
## Estimated class population shares
##   0.2108 0.0917 0.255 0.3379 0.1046
##
## Predicted class memberships (by modal posterior prob.)
##   0.2427 0.0759 0.2635 0.3269 0.0909
##
## ============================================================
## Fit for 5 latent classes:
## ============================================================
## number of observations: 3753
## number of estimated parameters: 49
## residual degrees of freedom: 462
## maximum log-likelihood: -16633.33
##
## AIC(5): 33364.65
## BIC(5): 33669.94
## G^2(5): 721.3581 (Likelihood ratio/deviance statistic)
## X^2(5): 1170.447 (Chi-square goodness of fit)
##
```

5. Examining predictors of the 4-class solution

```r
f <- cbind(sInvented, sProcedural, sConceptual, tInitSelect, tCompare, tDiscussQ, tConnectBigIdeas, tCon

m2 <- poLCA(f, dds, nclass = 2, maxiter = 10000, graphs = TRUE)
```
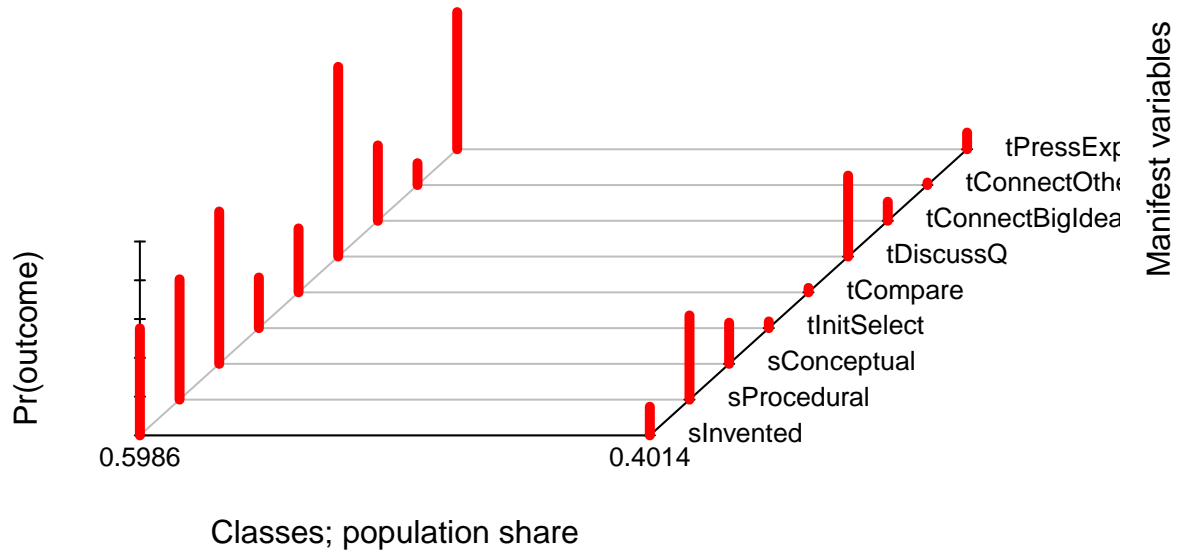
Classes; population share

```
## Conditional item response (column) probabilities,
##  by outcome variable, for each class (row)
##
## $sInvented
##           Pr(1) Pr(2)
## class 1:  0.448 0.552
## class 2:  0.852 0.148
##
## $sProcedural
##            Pr(1)  Pr(2)
## class 1:  0.3795 0.6205
## class 2:  0.5658 0.4342
##
## $sConceptual
##            Pr(1)  Pr(2)
## class 1:  0.2143 0.7857
## class 2:  0.7875 0.2125
##
## $tInitSelect
##            Pr(1)  Pr(2)
## class 1:  0.7390 0.2610
## class 2:  0.9657 0.0343
##
## $tCompare
##            Pr(1)  Pr(2)
## class 1:  0.6704 0.3296
```
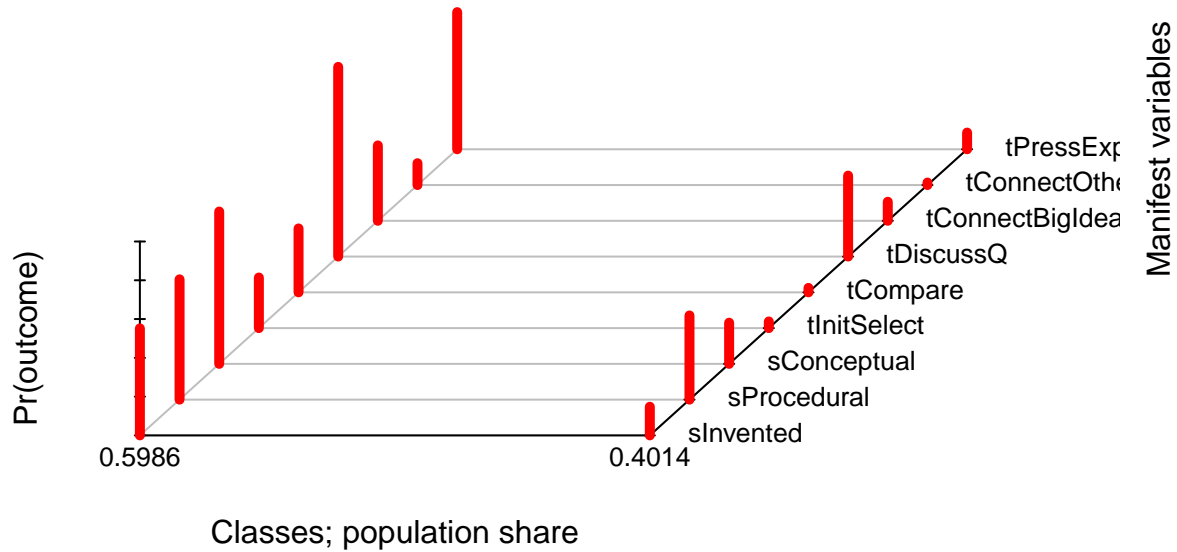
14

```
## class 2:  0.9771 0.0229
##
## $tDiscussQ
##            Pr(1)  Pr(2)
## class 1:  0.0216 0.9784
## class 2:  0.5825 0.4175
##
## $tConnectBigIdeas
##            Pr(1)  Pr(2)
## class 1:  0.6112 0.3888
## class 2:  0.9014 0.0986
##
## $tConnectOthers
##            Pr(1)  Pr(2)
## class 1:  0.8872 0.1128
## class 2:  0.9870 0.0130
##
## $tPressExplain
##            Pr(1)  Pr(2)
## class 1:  0.2926 0.7074
## class 2:  0.9138 0.0862
##
## Estimated class population shares
##   0.5986 0.4014
##
## Predicted class memberships (by modal posterior prob.)
##   0.6011 0.3989
##
## ============================================================
## Fit for 2 latent classes:
## ============================================================
## 2 / 1
##             Coefficient  Std. error  t value  Pr(>|t|)
## (Intercept)    33.07481     0.00973 3399.863         0
## condition     -16.94078     0.01946 -870.694         0
## ============================================================
## number of observations: 3753
## number of estimated parameters: 20
## residual degrees of freedom: 491
## maximum log-likelihood: -17134.56
##
## AIC(2): 34309.12
## BIC(2): 34433.73
## X^2(2): 4611.619 (Chi-square goodness of fit)
##
```

```r
m2 <- poLCA(f, dds, nclass = 2, maxiter = 10000, graphs = TRUE)
```

Pr(outcome)

Classes; population share

Manifest variables

tPressExp
tConnectOth
tConnectBigIdea
tDiscussQ
tCompare
tInitSelect
sConceptual
sProcedural
sInvented

0.5986   0.4014

```
## Conditional item response (column) probabilities,
##  by outcome variable, for each class (row)
##
## $sInvented
##           Pr(1) Pr(2)
## class 1:  0.448 0.552
## class 2:  0.852 0.148
##
## $sProcedural
##            Pr(1)  Pr(2)
## class 1:  0.3795 0.6205
## class 2:  0.5658 0.4342
##
## $sConceptual
##            Pr(1)  Pr(2)
## class 1:  0.2143 0.7857
## class 2:  0.7875 0.2125
##
## $tInitSelect
##            Pr(1)  Pr(2)
## class 1:  0.7390 0.2610
## class 2:  0.9657 0.0343
##
## $tCompare
##            Pr(1)  Pr(2)
## class 1:  0.6703 0.3297
```
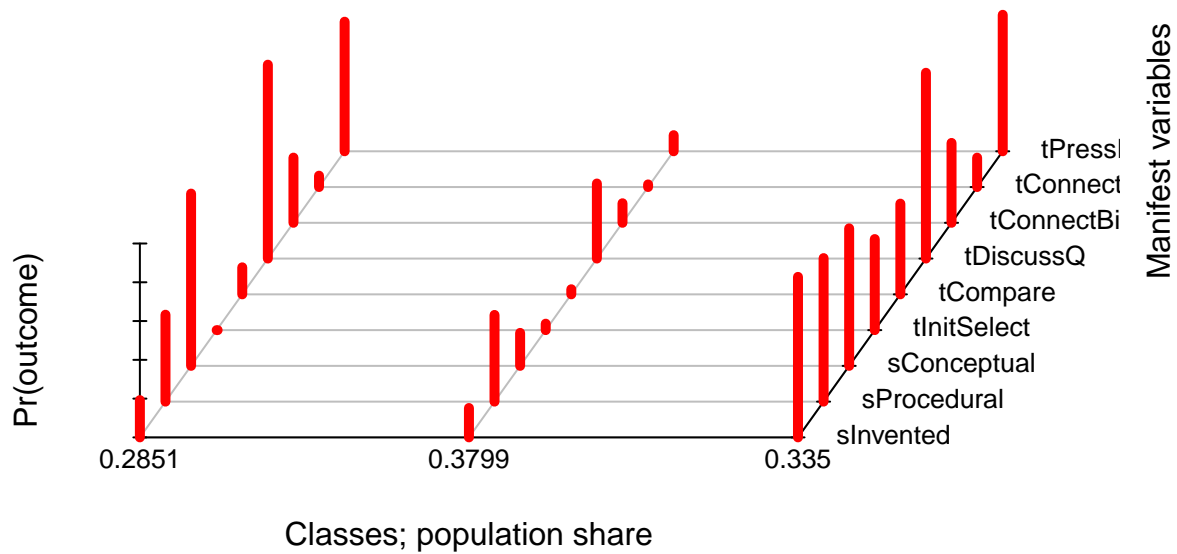
```
## class 2:  0.9771 0.0229
##
## $tDiscussQ
##             Pr(1)  Pr(2)
## class 1:  0.0216 0.9784
## class 2:  0.5825 0.4175
##
## $tConnectBigIdeas
##             Pr(1)  Pr(2)
## class 1:  0.6112 0.3888
## class 2:  0.9014 0.0986
##
## $tConnectOthers
##             Pr(1)  Pr(2)
## class 1:  0.8872 0.1128
## class 2:  0.9870 0.0130
##
## $tPressExplain
##             Pr(1)  Pr(2)
## class 1:  0.2926 0.7074
## class 2:  0.9138 0.0862
##
## Estimated class population shares
##   0.5986 0.4014
##
## Predicted class memberships (by modal posterior prob.)
##   0.6011 0.3989
##
## ============================================================
## Fit for 2 latent classes:
## ============================================================
## 2 / 1
##            Coefficient  Std. error  t value  Pr(>|t|)
## (Intercept)    34.38782     0.00973 3534.825         0
## condition     -17.59728     0.01946 -904.436         0
## ============================================================
## number of observations: 3753
## number of estimated parameters: 20
## residual degrees of freedom: 491
## maximum log-likelihood: -17134.56
##
## AIC(2): 34309.12
## BIC(2): 34433.73
## X^2(2): 4611.618 (Chi-square goodness of fit)
##
```

```r
m3 <- poLCA(f, dds, nclass = 3, maxiter = 10000, graphs = TRUE)
```

Manifest variables: tPressl, tConnect, tConnectBi, tDiscussQ, tCompare, tInitSelect, sConceptual, sProcedural, sInvented

Pr(outcome)

Classes; population share

0.2851   0.3799   0.335

```
## Conditional item response (column) probabilities,
##  by outcome variable, for each class (row)
##
## $sInvented
##            Pr(1)  Pr(2)
## class 1:  0.8077 0.1923
## class 2:  0.8484 0.1516
## class 3:  0.1720 0.8280
##
## $sProcedural
##            Pr(1)  Pr(2)
## class 1:  0.5517 0.4483
## class 2:  0.5528 0.4472
## class 3:  0.2597 0.7403
##
## $sConceptual
##            Pr(1)  Pr(2)
## class 1:  0.1115 0.8885
## class 2:  0.8298 0.1702
## class 3:  0.2908 0.7092
##
## $tInitSelect
##            Pr(1)  Pr(2)
## class 1:  1.0000 0.0000
## class 2:  0.9667 0.0333
## class 3:  0.5303 0.4697
```

```
## 
## $tCompare
##            Pr(1)  Pr(2)
## class 1:  0.8606 0.1394
## class 2:  0.9736 0.0264
## class 3:  0.5322 0.4678
## 
## $tDiscussQ
##            Pr(1)  Pr(2)
## class 1:  0.0000 1.0000
## class 2:  0.6128 0.3872
## class 3:  0.0416 0.9584
## 
## $tConnectBigIdeas
##            Pr(1)  Pr(2)
## class 1:  0.6637 0.3363
## class 2:  0.8990 0.1010
## class 3:  0.5879 0.4121
## 
## $tConnectOthers
##            Pr(1)  Pr(2)
## class 1:  0.9413 0.0587
## class 2:  0.9864 0.0136
## class 3:  0.8483 0.1517
## 
## $tPressExplain
##            Pr(1)  Pr(2)
## class 1:  0.3315 0.6685
## class 2:  0.9168 0.0832
## class 3:  0.2960 0.7040
## 
## Estimated class population shares
##   0.2851 0.3799 0.335
## 
## Predicted class memberships (by modal posterior prob.)
##   0.2963 0.373 0.3307
## 
## ============================================================
## Fit for 3 latent classes:
## ============================================================
## 2 / 1
##             Coefficient  Std. error  t value  Pr(>|t|)
## (Intercept)    39.30355     0.01422 2764.643         0
## condition     -19.72629     0.02843 -693.780         0
## ============================================================
## 3 / 1
##             Coefficient  Std. error  t value  Pr(>|t|)
## (Intercept)     7.15673     0.01783  401.303         0
## condition      -3.49761     0.03567  -98.062         0
## ============================================================
## number of observations: 3753
## number of estimated parameters: 31
## residual degrees of freedom: 480
## maximum log-likelihood: -16786.96
```

```
## 
## AIC(3): 33635.92
## BIC(3): 33829.06
## X^2(3): 2415.283 (Chi-square goodness of fit)
## 
```
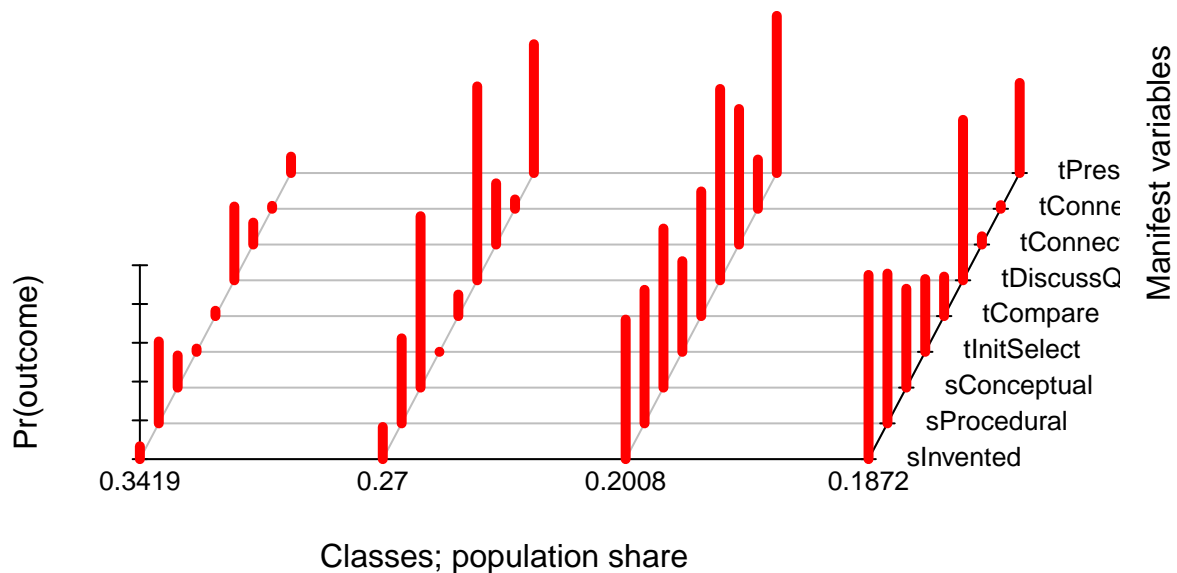
```
m4 <- poLCA(f, dds, nclass = 4, maxiter = 10000, graphs = TRUE)
```



```
## Conditional item response (column) probabilities,
##  by outcome variable, for each class (row)
## 
## $sInvented
##           Pr(1)  Pr(2)
## class 1:  0.9338 0.0662
## class 2:  0.8343 0.1657
## class 3:  0.2802 0.7198
## class 4:  0.0500 0.9500
## 
## $sProcedural
##           Pr(1)  Pr(2)
## class 1:  0.5778 0.4222
## class 2:  0.5615 0.4385
## class 3:  0.3112 0.6888
## class 4:  0.2276 0.7724
## 
```

```
## $sConceptual
##            Pr(1)  Pr(2)
## class 1:  0.8343 0.1657
## class 2:  0.1160 0.8840
## class 3:  0.1798 0.8202
## class 4:  0.4901 0.5099
##
## $tInitSelect
##            Pr(1)  Pr(2)
## class 1:  0.9833 0.0167
## class 2:  0.9996 0.0004
## class 3:  0.5325 0.4675
## class 4:  0.6246 0.3754
##
## $tCompare
##            Pr(1)  Pr(2)
## class 1:  0.9726 0.0274
## class 2:  0.8892 0.1108
## class 3:  0.3572 0.6428
## class 4:  0.7964 0.2036
##
## $tDiscussQ
##            Pr(1)  Pr(2)
## class 1:  0.6196 0.3804
## class 2:  0.0000 1.0000
## class 3:  0.0131 0.9869
## class 4:  0.1724 0.8276
##
## $tConnectBigIdeas
##            Pr(1)  Pr(2)
## class 1:  0.8869 0.1131
## class 2:  0.6843 0.3157
## class 3:  0.3009 0.6991
## class 4:  0.9575 0.0425
##
## $tConnectOthers
##            Pr(1)  Pr(2)
## class 1:  0.9849 0.0151
## class 2:  0.9518 0.0482
## class 3:  0.7459 0.2541
## class 4:  0.9811 0.0189
##
## $tPressExplain
##            Pr(1)  Pr(2)
## class 1:  0.9154 0.0846
## class 2:  0.3357 0.6643
## class 3:  0.1896 0.8104
## class 4:  0.5355 0.4645
##
## Estimated class population shares
##   0.3419 0.27 0.2008 0.1872
##
## Predicted class memberships (by modal posterior prob.)
##   0.3227 0.2822 0.1881 0.207
```

```
## 
## ================================================================
## Fit for 4 latent classes:
## ================================================================
## 2 / 1
##            Coefficient  Std. error   t value  Pr(>|t|)
## (Intercept)  -35.37547     0.01501 -2356.945         0
## condition     17.81916     0.02999   594.148         0
## ================================================================
## 3 / 1
##            Coefficient  Std. error   t value  Pr(>|t|)
## (Intercept)  -34.31754     0.01741 -1970.863         0
## condition     17.14220     0.03485   491.923         0
## ================================================================
## 4 / 1
##            Coefficient  Std. error   t value  Pr(>|t|)
## (Intercept)  -30.46994     0.01814 -1679.387         0
## condition     15.18333     0.03632   418.068         0
## ================================================================
## number of observations: 3753
## number of estimated parameters: 42
## residual degrees of freedom: 469
## maximum log-likelihood: -16575.95
## 
## AIC(4): 33235.89
## BIC(4): 33497.56
## X^2(4): 1716.269 (Chi-square goodness of fit)
## 
## ALERT: estimation algorithm automatically restarted with new initial values
## 
```

```r
m5 <- poLCA(f, dds, nclass = 5, maxiter = 10000, graphs = TRUE)
```
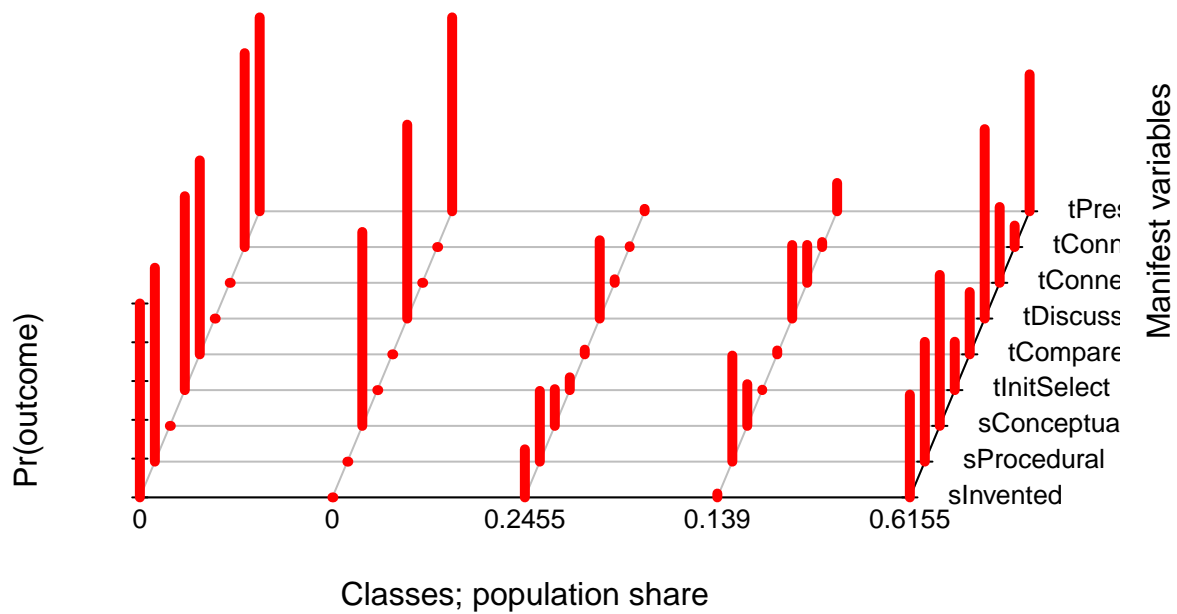
```
## Conditional item response (column) probabilities,
##  by outcome variable, for each class (row)
##
## $sInvented
##           Pr(1)  Pr(2)
## class 1:  0.0000 1.0000
## class 2:  1.0000 0.0000
## class 3:  0.7522 0.2478
## class 4:  0.9802 0.0198
## class 5:  0.4700 0.5300
##
## $sProcedural
##           Pr(1)  Pr(2)
## class 1:  0.0000 1.0000
## class 2:  1.0000 0.0000
## class 3:  0.6352 0.3648
## class 4:  0.4532 0.5468
## class 5:  0.3824 0.6176
##
## $sConceptual
##           Pr(1)  Pr(2)
## class 1:  1.0000 0.0000
## class 2:  0.0000 1.0000
## class 3:  0.8106 0.1894
## class 4:  0.7851 0.2149
## class 5:  0.2215 0.7785
```

```
## 
## $tInitSelect
##            Pr(1)  Pr(2)
## class 1:  0.0000 1.0000
## class 2:  1.0000 0.0000
## class 3:  0.9341 0.0659
## class 4:  0.9980 0.0020
## class 5:  0.7506 0.2494
## 
## $tCompare
##            Pr(1)  Pr(2)
## class 1:  0.0000 1.0000
## class 2:  1.0000 0.0000
## class 3:  0.9759 0.0241
## class 4:  0.9801 0.0199
## class 5:  0.6786 0.3214
## 
## $tDiscussQ
##            Pr(1)  Pr(2)
## class 1:  1.0000 0.0000
## class 2:  0.0000 1.0000
## class 3:  0.5961 0.4039
## class 4:  0.6207 0.3793
## class 5:  0.0230 0.9770
## 
## $tConnectBigIdeas
##            Pr(1)  Pr(2)
## class 1:  1.0000 0.0000
## class 2:  1.0000 0.0000
## class 3:  0.9818 0.0182
## class 4:  0.8047 0.1953
## class 5:  0.6089 0.3911
## 
## $tConnectOthers
##            Pr(1)  Pr(2)
## class 1:  0.0000 1.0000
## class 2:  1.0000 0.0000
## class 3:  0.9959 0.0041
## class 4:  0.9742 0.0258
## class 5:  0.8893 0.1107
## 
## $tPressExplain
##            Pr(1)  Pr(2)
## class 1:  0.0000 1.0000
## class 2:  0.0000 1.0000
## class 3:  0.9875 0.0125
## class 4:  0.8540 0.1460
## class 5:  0.2938 0.7062
## 
## Estimated class population shares
##   0 0 0.2455 0.139 0.6155
## 
## Predicted class memberships (by modal posterior prob.)
##   0.2587 0.1343 0.607
```

```
## 
## =============================================================
## Fit for 5 latent classes:
## =============================================================
## 2 / 1
##             Coefficient  Std. error      t value  Pr(>|t|)
## (Intercept)   -38.02780           0 -41159132315         0
## condition      18.55307           0  10040394549         0
## =============================================================
## 3 / 1
##             Coefficient  Std. error   t value  Pr(>|t|)
## (Intercept)  -164.16553     0.15967 -1028.157         0
## condition      92.69443     0.31934   290.269         0
## =============================================================
## 4 / 1
##             Coefficient  Std. error  t value  Pr(>|t|)
## (Intercept)    51.24144     0.31058  164.988         0
## condition     -16.99041     0.62115  -27.353         0
## =============================================================
## 5 / 1
##             Coefficient  Std. error  t value  Pr(>|t|)
## (Intercept)   -40.11479     0.15108 -265.527         0
## condition      31.12865     0.30215  103.023         0
## =============================================================
## number of observations: 3753
## number of estimated parameters: 53
## residual degrees of freedom: 458
## maximum log-likelihood: -16941.05
## 
## AIC(5): 33988.09
## BIC(5): 34318.3
## X^2(5): 4662.233 (Chi-square goodness of fit)
## 
## ALERT: estimation algorithm automatically restarted with new initial values
## 
```