Fiche de modélisations n°6

Variables et classes latentes

Kim Antunez

Table des matières

1 Objectif

L'objectif de cette sixième série de modèles est de ...

2 Analyses

TODO

3 Code et résultats

```
#chargement des packages
library(knitr)
library(dplyr) #manipuler les bases de données
library(psych) #EFA
library(lavaan) #CFA et SEM
library(semPlot) #path draw CFA SEM
library(poLCA) #pour les Latent Categorical Variables
```

```
library(ade4) #pour la fonction s5 de plot des classes de CAH
library(RColorBrewer) #palettes de couleur
```

3.1 Correlation coefficients

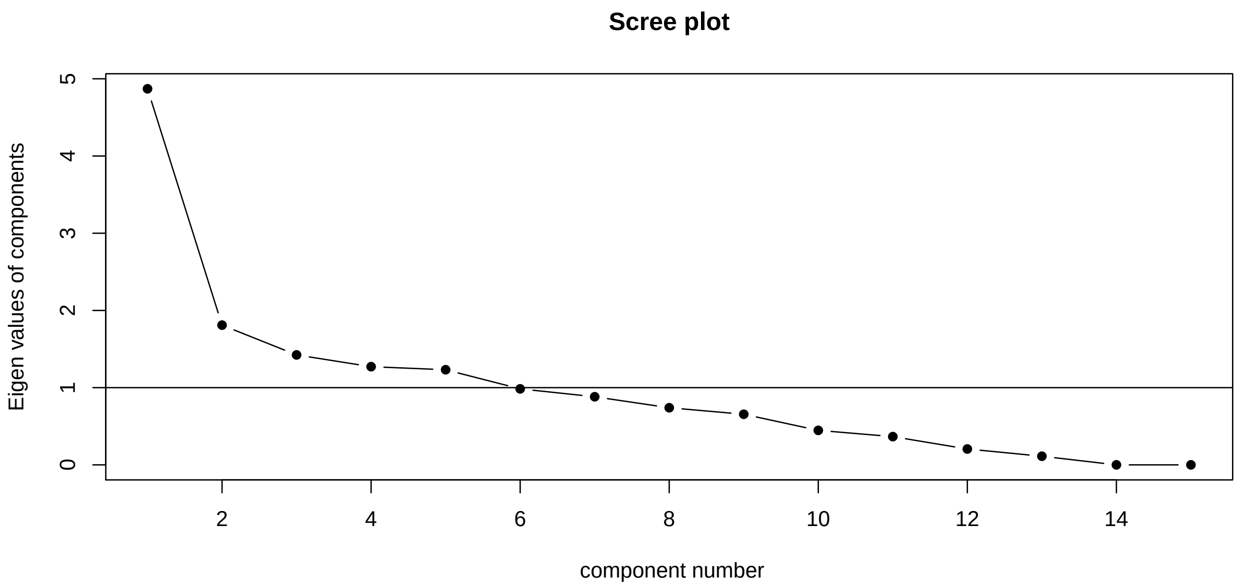
A correlation coefficient suited for dichotomous data and based on this underlying normal strategy is the tetrachoric correlation. It gives us a single number describing the degree of dependence in the table above with the extreme values of 1 if the off-diagonals are 0 and -1 if the diagonals are 0. In addition, we get estimates for the thresholds tau1 and tau2. polycholoric existe aussi pour deux items polytomous.

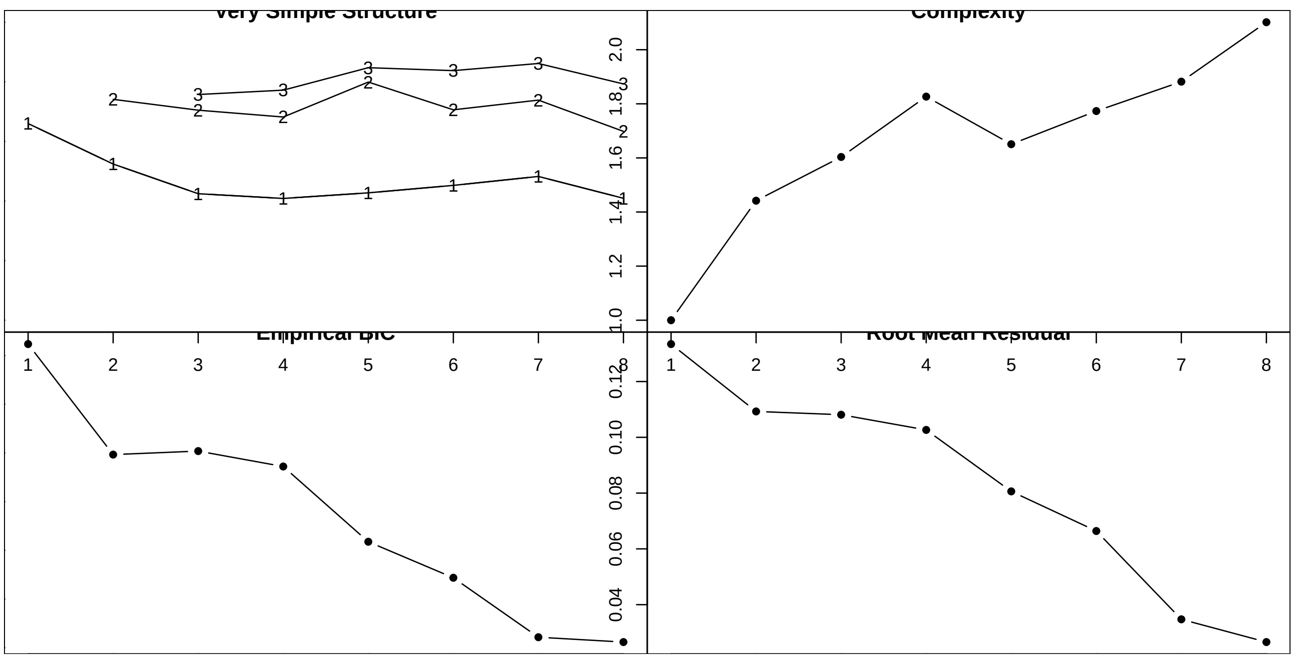
We print out the last six eigenvalues and see that the last eigenvalue is negative. Thus, this matrix does not fulfill the properties of a correlation matrix. The trick is now to apply some smoothing on the correlations.

The final criterion is interpretability.

```
                s_sentpauv  s_risqpauv s_infminidecla   m_nivie1    m_nivie2
s_sentpauv      1.00000000 -0.61169115    -0.47699835  0.4394866  0.17700036
s_risqpauv     -0.61169115  1.00000000    -0.21941563  0.1122074  0.12147720
s_infminidecla -0.47699835 -0.21941563     1.00000000 -0.5473902 -0.27817577
m_nivie1        0.43948663  0.11220743    -0.54739017  1.0000000 -0.27389219
m_nivie2        0.17700036  0.12147720    -0.27817577 -0.2738922  1.00000000
m_nivie3       -0.08999267  0.08204612    -0.03550473 -0.2757136 -0.23815590
m_nivie4       -0.24728331 -0.04516964     0.28020447 -0.2669469 -0.22774028
m_nivie5       -0.36445531 -0.21452006     0.59276364 -0.2736265 -0.23483094
m_locatif      -0.42268192 -0.18495526     0.43699367 -0.2776909 -0.21860883
m_financier    -0.38016835 -0.25518131     0.50749995 -0.2862883 -0.22859129
i_log           0.49288404  0.13004116    -0.42492371  0.6070170  0.17755378
i_rsa           0.44017858  0.04078294    -0.39890908  0.6486328 -0.02929894
i_chom          0.17847358  0.16259637    -0.19360799  0.2291738  0.10878280
i_handi         0.18637178  0.11904696    -0.19205629  0.2186925  0.11589690
i_hlm           0.36717594  0.16949615    -0.37835726  0.4058881  0.17709505
                   m_nivie3    m_nivie4   m_nivie5   m_locatif m_financier
s_sentpauv     -0.089992672 -0.24728331 -0.3644553 -0.42268192 -0.38016835
s_risqpauv      0.082046117 -0.04516964 -0.2145201 -0.18495526 -0.25518131
s_infminidecla -0.035504730  0.28020447  0.5927636  0.43699367  0.50749995
m_nivie1       -0.275713589 -0.26694692 -0.2736265 -0.27769091 -0.28628830
m_nivie2       -0.238155904 -0.22774028 -0.2348309 -0.21860883 -0.22859129
m_nivie3        1.000000000 -0.22648136 -0.2328622 -0.08775699 -0.09442612
m_nivie4       -0.226481357  1.00000000 -0.2210290  0.08268827  0.05524235
m_nivie5       -0.232862235 -0.22102897  1.0000000  0.45649774  0.49781210
m_locatif      -0.087756993  0.08268827  0.4564977  1.00000000  0.61121694
m_financier    -0.094426115  0.05524235  0.4978121  0.61121694  1.00000000
i_log          -0.149289213 -0.33336097 -0.4432850 -0.40224067 -0.40972541
i_rsa          -0.184244388 -0.25252496 -0.3693635 -0.22111111 -0.21187655
i_chom         -0.012275418 -0.13421310 -0.2382888 -0.17770850 -0.18630417
i_handi        -0.006415386 -0.12957374 -0.2503593 -0.10011997 -0.10202599
i_hlm          -0.052358458 -0.20535458 -0.4039782 -0.46037012 -0.37407010
                     i_log       i_rsa      i_chom      i_handi       i_hlm
```

```
s_sentpauv      0.4928840  0.44017858  0.17847358  0.186371776  0.36717594
s_risqpauv      0.1300412  0.04078294  0.16259637  0.119046959  0.16949615
s_infminidecla -0.4249237 -0.39890908 -0.19360799 -0.192056293 -0.37835726
m_nivie1        0.6070170  0.64863281  0.22917379  0.218692458  0.40588808
m_nivie2        0.1775538 -0.02929894  0.10878280  0.115896903  0.17709505
m_nivie3       -0.1492892 -0.18424439 -0.01227542 -0.006415386 -0.05235846
m_nivie4       -0.3333610 -0.25252496 -0.13421310 -0.129573738 -0.20535458
m_nivie5       -0.4432850 -0.36936348 -0.23828876 -0.250359315 -0.40397819
m_locatif      -0.4022407 -0.22111111 -0.17770850 -0.100119971 -0.46037012
m_financier    -0.4097254 -0.21187655 -0.18630417 -0.102025992 -0.37407010
i_log           1.0000000  0.66315036  0.32253411  0.354662651  0.60573548
i_rsa           0.6631504  1.00000000  0.18678644  0.062453472  0.39150437
i_chom          0.3225341  0.18678644  1.00000000  0.093226305  0.19279940
i_handi         0.3546627  0.06245347  0.09322631  1.000000000  0.28122027
i_hlm           0.6057355  0.39150437  0.19279940  0.281220267  1.00000000

 [1] 32.47 12.07  9.49  8.48  8.21  6.56  5.88  4.93  4.37  2.98  2.44  1.37
[13]  0.75  0.00  0.00
```





3.2 Exploratory Factor Analysis (EFA)

However, in order to get an even clearer picture, in EFA we typically apply a rotation on the loadings matrix. Such a rotation does not change the fit of the model; it is only done for interpretation purposes by transforming the loadings. We distinguish between two basic types of rotations: orthogonal (qui implique que les facteurs sont indépendants) and nonorthogonal rotation (comme oblimin).

In practice, EFA with oblique rotation is often used prior to a CFA in order to explore whether the underlying latent structure theory is reflected by the data.

Différences entre EFA et ACP - Les ACP reposent sur des estimations bien plus simples que les EFA (ML, LS). - L'EFA se concentrent sur l'explication des termes en dehors de la diagonales des éléments de grand sigma (explique les covariances) alors que l'ACP se concentre sur la diagonale (explique principalement la variance, même si pas totalement aveugle aux covariances). - Les scores des facteurs sont calculés post hoc alors que ceux de l'ACP est une conséquence directe du SVD. - En EFA, on fixe p avant de faire tourner le modèle, dans l'ACP on choisit le nombre d'axes a posteriori. - En EFA, les rotations peuvent aider à mieux interpréter sans changer la solution, contrairement à la PCA pour laquelle la solution est changée.

PCA et EFA sont deux techniques de réduction de dimensions mais elles ont des différences.

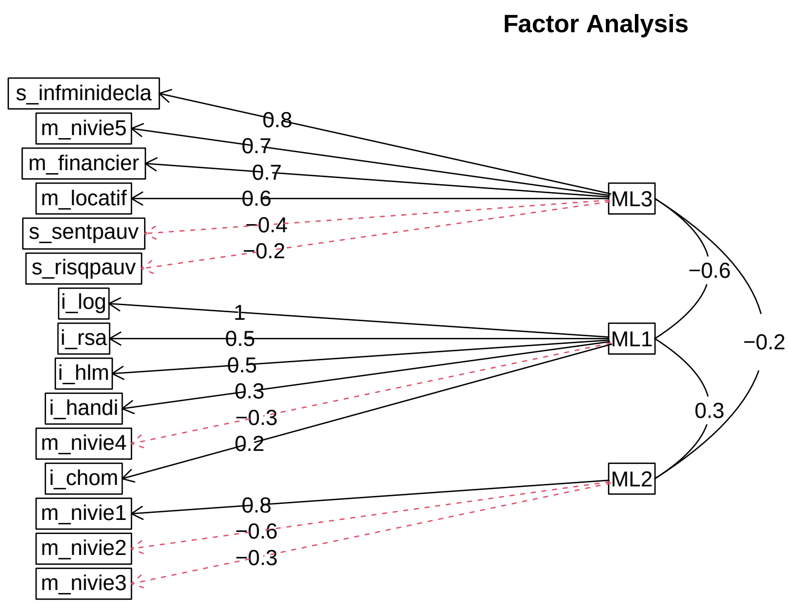
On utilise la PCA quand les variables sont très corrélées cela permet de réduire le nombre de variables observées en un plus petit nombre de composantes principales qui résument un maximum de variances des variables observées On utilise l'EFA pour identifier le nombre de variables latentes (non mesurée directement) et la structure des facteurs qui découlent d'un ensemble de variables. Elle permet d'estimer les facteurs qui influencent les réponses des variables observées.

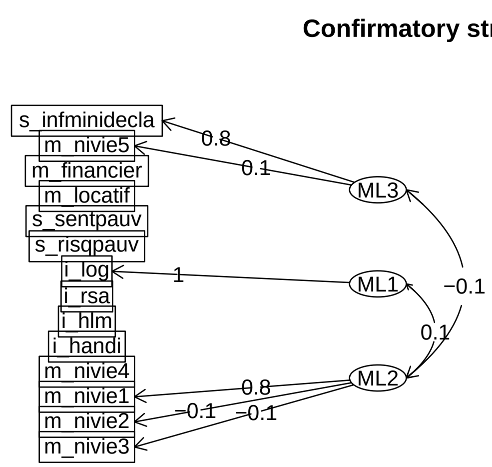
Unlike factor analysis, principal components analysis or PCA makes the assumption that there is no unique variance, the total variance is equal to common variance. Recall that variance can be partitioned into common and unique variance. If there is no unique variance then common variance takes up total variance (see figure below). Additionally, if the total variance is 1, then the common variance is equal to the communality.

— PCA suppose l'absence d'outliers. L'EFA suppose une distribution normale multivariée quand la méthode de ML est utilisée
— Les axes de la PCA tiennent compte de la variance maximale des variables observées alors que les facteurs de l'EFA tiennent compte de la variance commune
— La PCA utilise une matrice de corrélation alors que l'EFA utilise une matrice de corrélation ajustée
— Dans une PCA il y a des 1 sur la diagonale de la matrice de corrélation alors que dans l'EFA la diagonale est ajustée avec les facteurs uniques.
— La PCA minimise la somme des carrés perpendiculaire à la distance aux axes des composantes. L'EFA estime des facteurs qui influencent la réponse à des variables observées.
— Les scores des composantes de la PCA est une combinaison linéaire des variables observées pondérées par les vecteurs propres. Les variables observées de l'EFA sont une combinaison linéaire des facteurs uniques

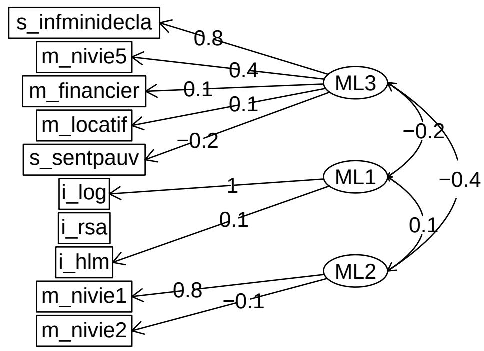
n alternative to factor analysis, which is unfortunately frequently confused withfactoranalysis, isprincipal components analysis. Although the goals ofPCAandFAare similar,PCA is a descriptive model of the data, while FA is a structural model. Psychologiststypically use PCA in a manner similar to factor analysis and thus theprincipalfunctionproduces output that is perhaps more understandable than that produced byprincompin thestatspackage. Table 4 shows a PCA of the Thurstone 9 variable problem rotatedusing thePromaxfunction. Note how the loadings from the factor model are similar butsmaller than the principal component loadings. This is because the PCA model attemptsto account for the entire variance of the correlation matrix, while FA accounts for just thecommon variance. This distinction becomes most important for small correlation matrices.Also note how the goodness of fit statistics, based upon the residual off diagonal elements,is much worse than thefasolution.

```
Factor analysis with Call: fa(r = bdd_fa_tetra$rho, nfactors = 3, rotate = "oblimin", scores = '
    missing = TRUE, impute = "median", fm = "ml", cor = "tet")

Test of the hypothesis that 3 factors are sufficient.
The degrees of freedom for the model is 63  and the objective function was  45.01

The root mean square of the residuals (RMSA) is  0.11
The df corrected root mean square of the residuals is  0.14

 With factor correlations of
      ML3   ML1   ML2
ML3  1.00 -0.55 -0.21
ML1 -0.55  1.00  0.35
ML2 -0.21  0.35  1.00
```

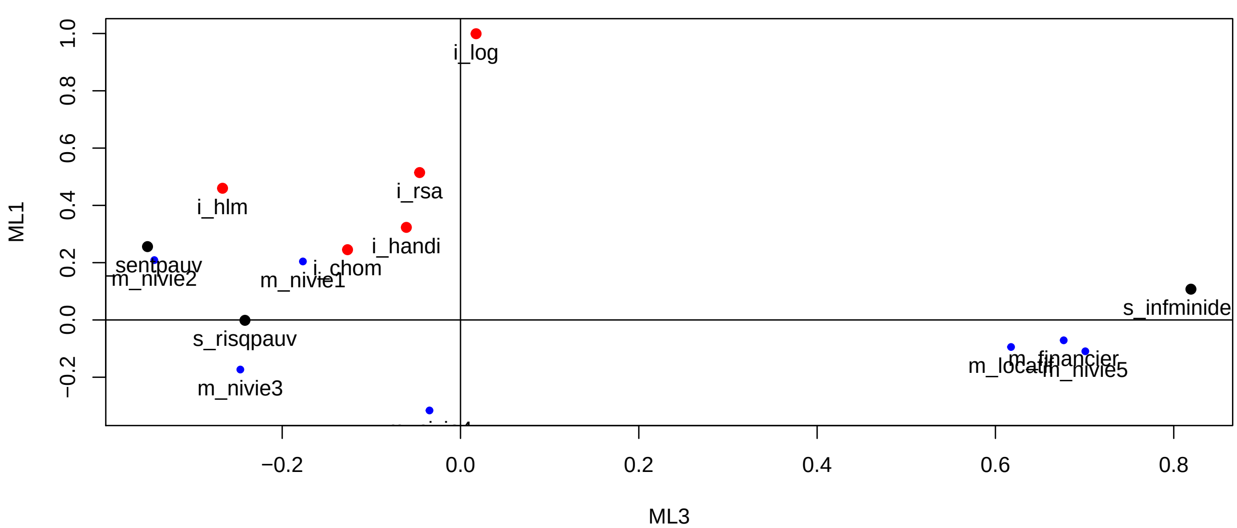

Factor Analysis
s_infminidecla
m_nivie5
m_financier
m_locatif
s_sentpauv
s_risqpauv
i_log
i_rsa
i_hlm
i_handi
m_nivie4
i_chom
m_nivie1
m_nivie2
m_nivie3
ML3
ML1
ML2
0.8
0.7
0.7
0.6
−0.4
−0.2
1
0.5
0.5
0.3
−0.3
0.2
0.8
−0.6
−0.3
−0.6
−0.2
0.3


Confirmatory structure
s_infminidecla
m_nivie5
m_financier
m_locatif
s_sentpauv
s_risqpauv
i_log
i_rsa
i_hlm
i_handi
m_nivie4
m_nivie1
m_nivie2
m_nivie3
ML3
ML1
ML2
0.8
0.1
1
0.8
−0.1
−0.1
−0.1
0.1

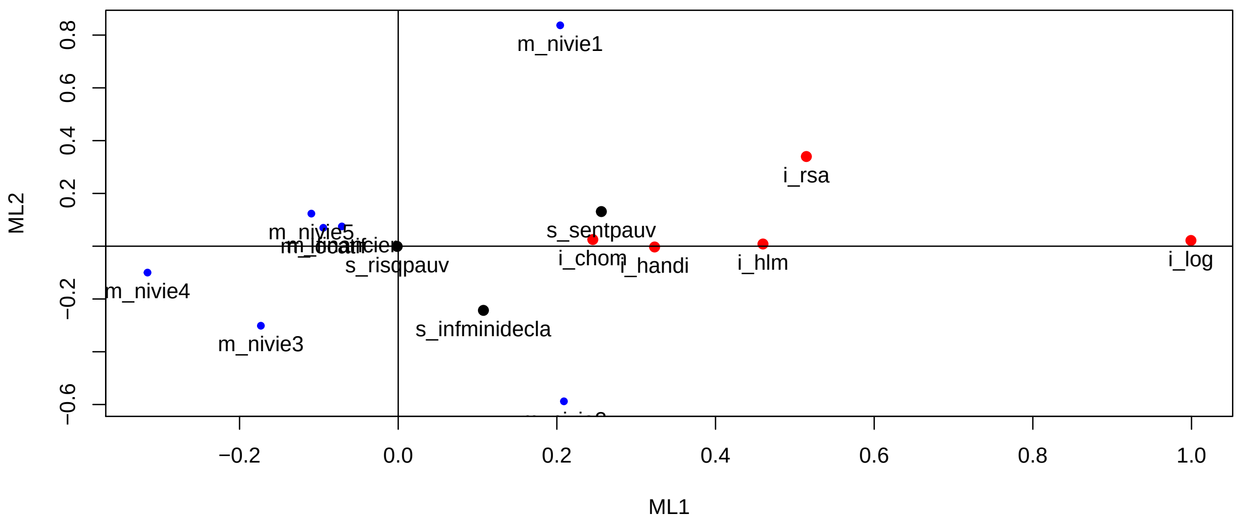
Confirmatory structure



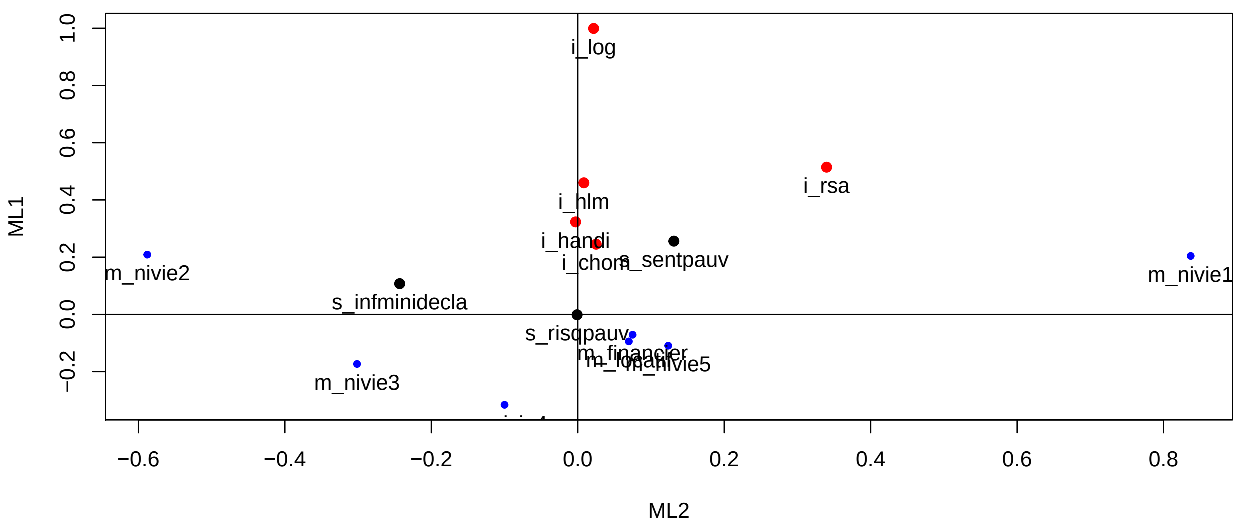
Plot



Plot



Plot



Loadings:

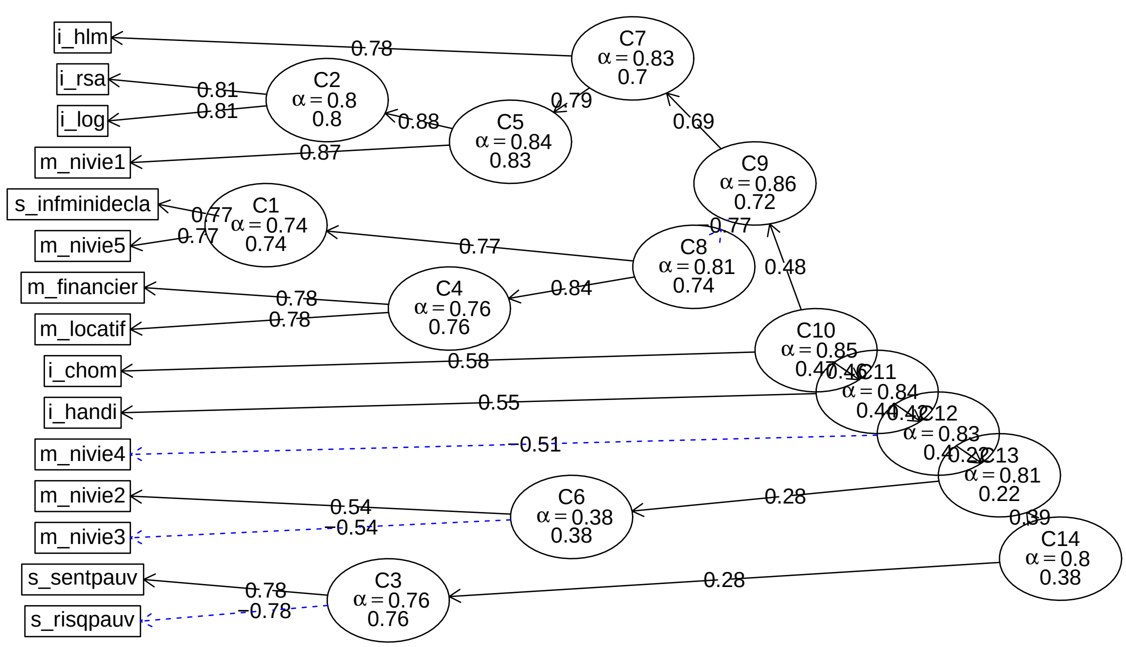
	ML3	ML1	ML2
s_sentpauv	-0.351	0.256	
s_risqpauv	-0.242		
s_infminidecla	0.819		-0.243
m_nivie1		0.204	0.837
m_nivie2	-0.343	0.209	-0.588
m_nivie3	-0.247		-0.301
m_nivie4		-0.316	
m_nivie5	0.701		
m_locatif	0.618		
m_financier	0.677		

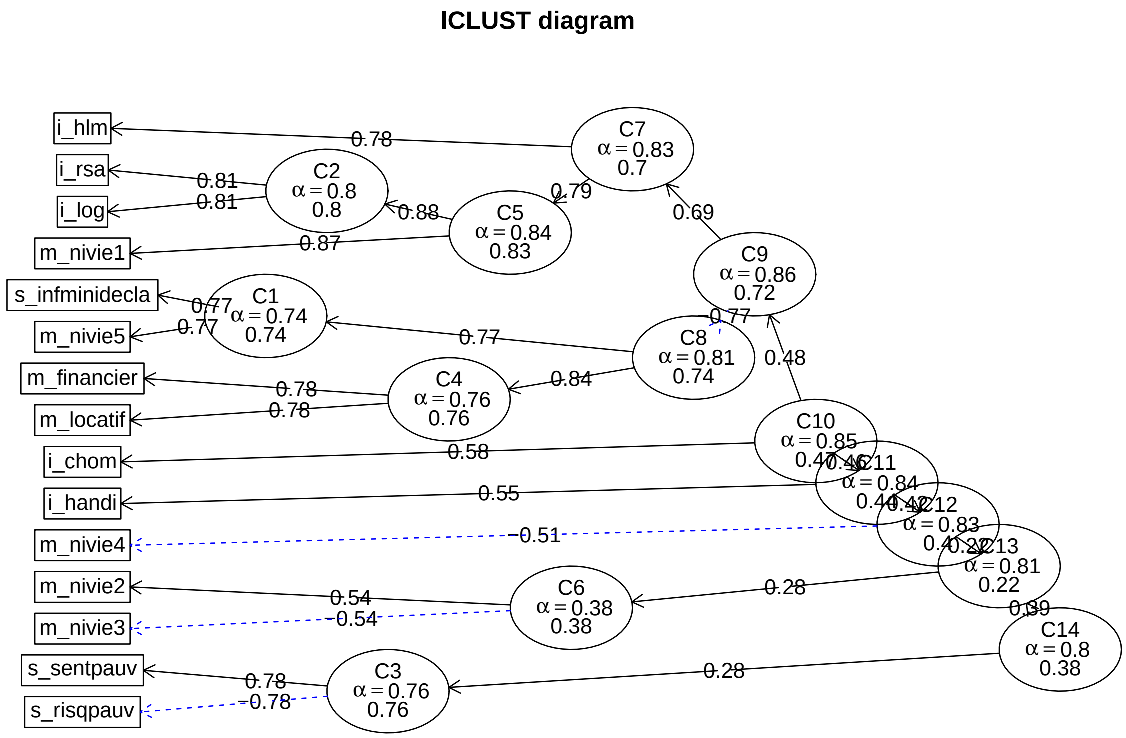
```
i_log                  0.999
i_rsa                  0.515  0.340
i_chom                 0.245
i_handi                0.323
i_hlm          -0.267  0.460

                 ML3   ML1   ML2
SS loadings    2.488 1.958 1.366
Proportion Var 0.166 0.131 0.091
Cumulative Var 0.166 0.296 0.387
```

```
    s_sentpauv      s_risqpauv s_infminidecla        m_nivie1        m_nivie2
          0.35            0.06           0.71            1.00            0.42
      m_nivie3        m_nivie4       m_nivie5       m_locatif     m_financier
          0.14            0.12           0.56            0.44            0.50
         i_log           i_rsa         i_chom         i_handi           i_hlm
          1.00            0.54           0.12            0.13            0.42
```

Premier type de clustering (de variables et non d'individus) avec iclust

ICLUST using tetrachoric correlations





3.3 Latent Categorical Variables

Source : https://m-clark.github.io/sem/mixture-models.html

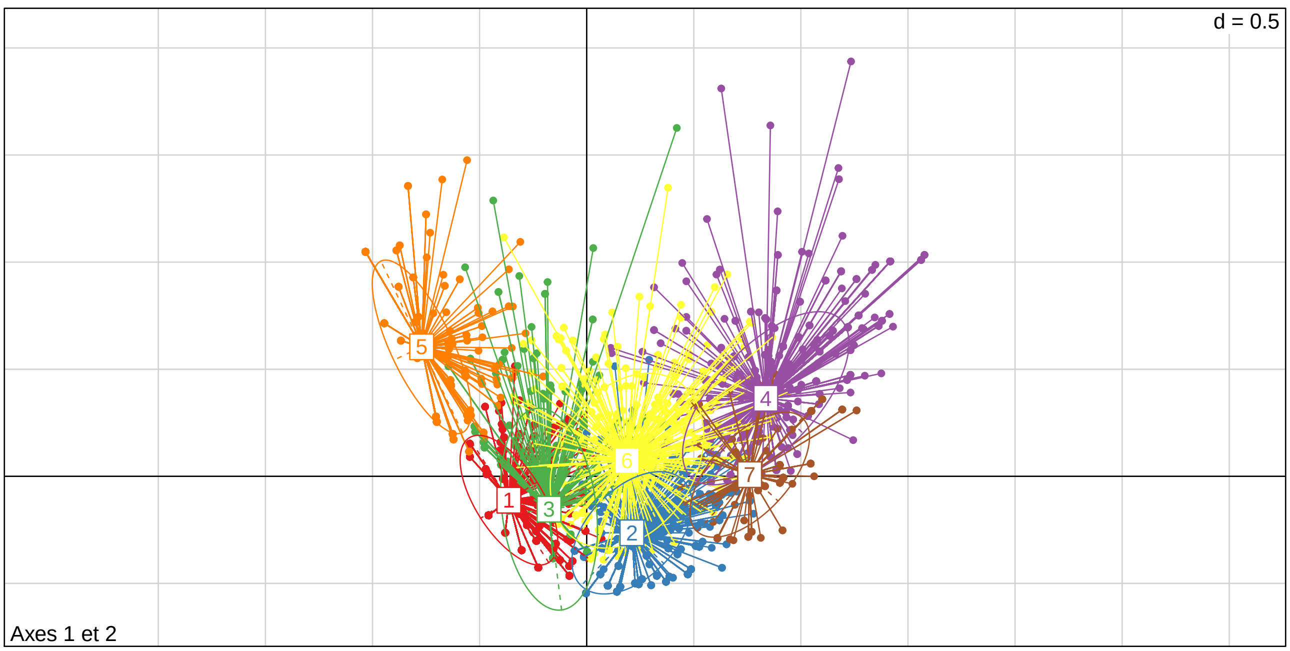
Documentation https://raw.githubusercontent.com/dlinzer/poLCA/master/inst/doc/poLCA-manual-1-4.pdf

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

$s_sentpauvrisque
           Pr(1)  Pr(2)  Pr(3)
class 1:  0.7975 0.1845 0.0180
class 2:  0.3633 0.3845 0.2521
class 3:  0.8155 0.1845 0.0000
class 4:  0.2344 0.2650 0.5006
class 5:  0.9471 0.0487 0.0042
class 6:  0.5667 0.3142 0.1191
class 7:  0.2428 0.3217 0.4355

$s_infminidecla
           Pr(1)  Pr(2)
class 1:  0.2609 0.7391
class 2:  0.9326 0.0674
class 3:  0.8074 0.1926
class 4:  0.9585 0.0415
class 5:  0.0549 0.9451
```

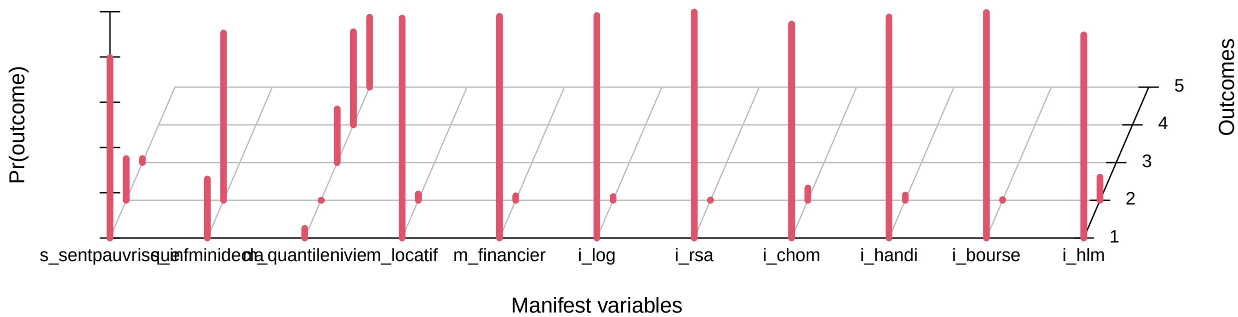
```
class 6:  0.2037 0.7963
class 7:  0.8820 0.1180

$m_quantilenivie
           Pr(1)  Pr(2)  Pr(3)  Pr(4)  Pr(5)
class 1:  0.0426 0.0000 0.2371 0.4115 0.3088
class 2:  0.3647 0.2495 0.3079 0.0780 0.0000
class 3:  0.2029 0.1859 0.4139 0.1973 0.0000
class 4:  0.0875 0.8997 0.0100 0.0028 0.0000
class 5:  0.0055 0.0000 0.0736 0.1714 0.7496
class 6:  0.3286 0.1387 0.3331 0.1153 0.0843
class 7:  0.3483 0.5438 0.0861 0.0166 0.0052

$m_locatif
           Pr(1)  Pr(2)
class 1:  0.9718 0.0282
class 2:  0.9972 0.0028
class 3:  0.8742 0.1258
class 4:  0.9891 0.0109
class 5:  0.6464 0.3536
class 6:  0.9701 0.0299
class 7:  1.0000 0.0000

$m_financier
           Pr(1)  Pr(2)
class 1:  0.9799 0.0201
class 2:  0.9942 0.0058
class 3:  0.8932 0.1068
class 4:  0.9880 0.0120
class 5:  0.5842 0.4158
class 6:  0.9519 0.0481
class 7:  1.0000 0.0000

$i_log
           Pr(1)  Pr(2)
class 1:  0.9833 0.0167
class 2:  0.8503 0.1497
class 3:  0.9825 0.0175
class 4:  0.1656 0.8344
class 5:  0.9859 0.0141
class 6:  0.3947 0.6053
class 7:  0.0000 1.0000

$i_rsa
           Pr(1)  Pr(2)
class 1:  0.9986 0.0014
class 2:  0.9928 0.0072
class 3:  0.9876 0.0124
class 4:  0.6244 0.3756
```

```
class 5:  0.9990 0.0010
class 6:  0.9158 0.0842
class 7:  1.0000 0.0000

$i_chom
           Pr(1)  Pr(2)
class 1:  0.9456 0.0544
class 2:  0.8446 0.1554
class 3:  0.9244 0.0756
class 4:  0.7552 0.2448
class 5:  0.9563 0.0437
class 6:  0.7543 0.2457
class 7:  0.8060 0.1940

$i_handi
           Pr(1)  Pr(2)
class 1:  0.9766 0.0234
class 2:  0.9236 0.0764
class 3:  0.9377 0.0623
class 4:  0.9118 0.0882
class 5:  0.9763 0.0237
class 6:  0.8723 0.1277
class 7:  0.6214 0.3786

$i_bourse
           Pr(1)  Pr(2)
class 1:  0.9966 0.0034
class 2:  0.9816 0.0184
class 3:  0.9489 0.0511
class 4:  0.8052 0.1948
class 5:  0.9850 0.0150
class 6:  0.8358 0.1642
class 7:  0.9396 0.0604

$i_hlm
           Pr(1)  Pr(2)
class 1:  0.8976 0.1024
class 2:  0.6983 0.3017
class 3:  0.9402 0.0598
class 4:  0.4508 0.5492
class 5:  0.9863 0.0137
class 6:  0.5807 0.4193
class 7:  0.1471 0.8529

Estimated class population shares
 0.2626 0.2348 0.0995 0.1301 0.1486 0.0774 0.047

Predicted class memberships (by modal posterior prob.)
 0.3305 0.253 0.0751 0.1368 0.0997 0.0659 0.039
```

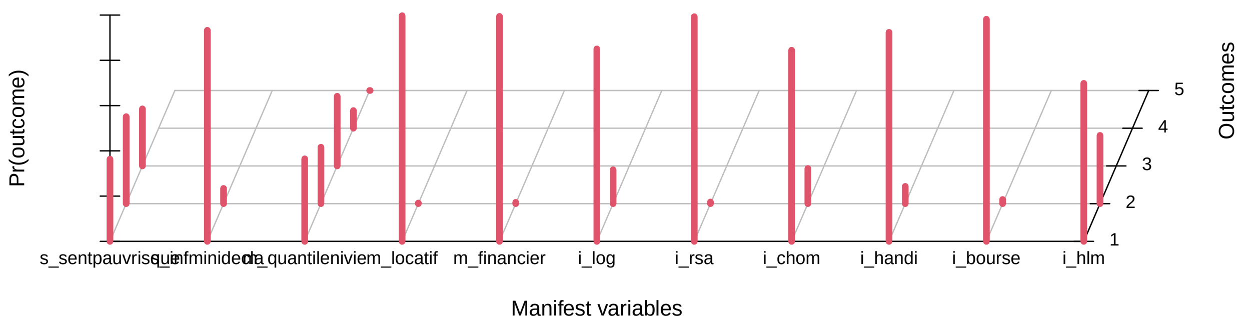
```
============================================================
Fit for 7 latent classes:
============================================================
number of observations: 10555
number of estimated parameters: 111
residual degrees of freedom: 7568
maximum log-likelihood: -55120.34

AIC(7): 110462.7
BIC(7): 111269
G^2(7): 2375.8 (Likelihood ratio/deviance statistic)
X^2(7): 189592.9 (Chi-square goodness of fit)

ALERT: iterations finished, MAXIMUM LIKELIHOOD NOT FOUND
```



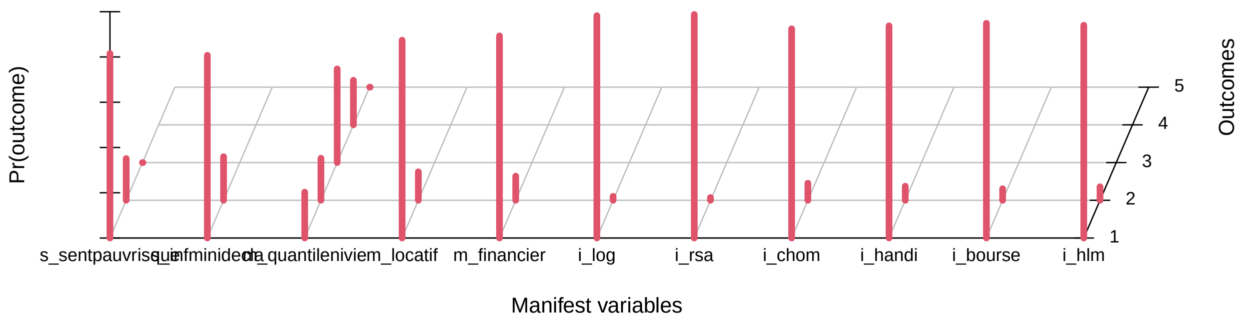
Classe 1 : part de la population = 26.3 %



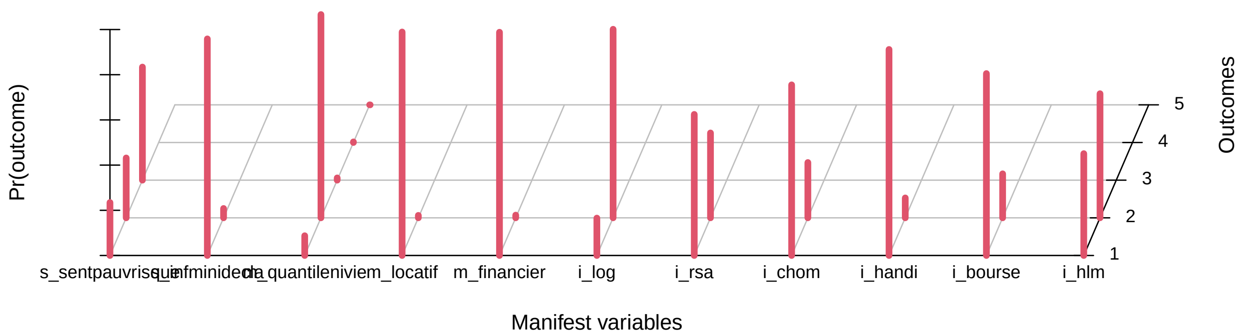
Classe 2 : part de la population = 23.5 %

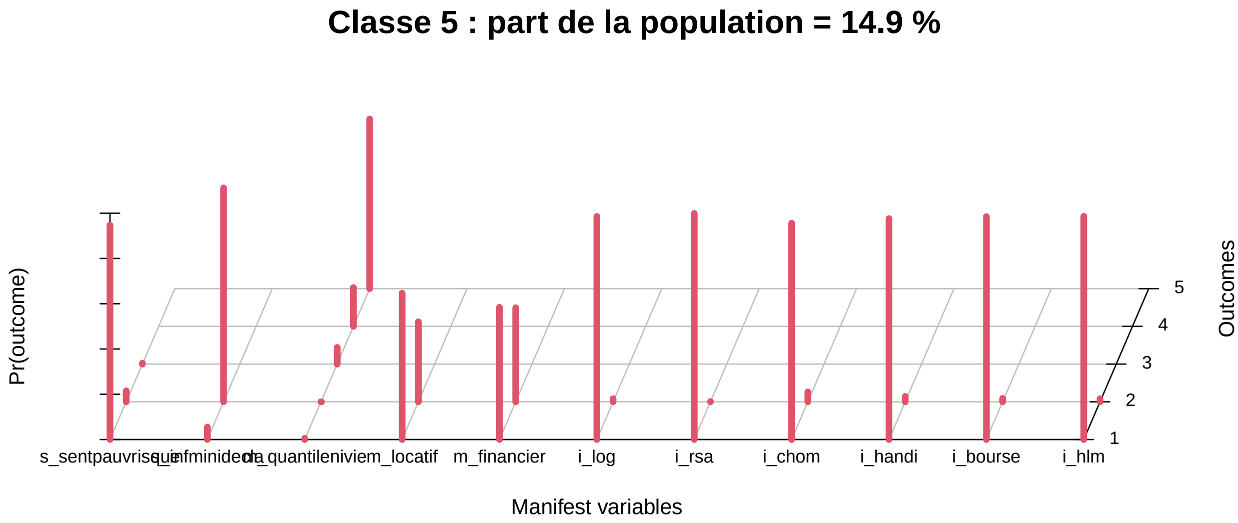


Classe 3 : part de la population = 9.9 %



Classe 4 : part de la population = 13 %

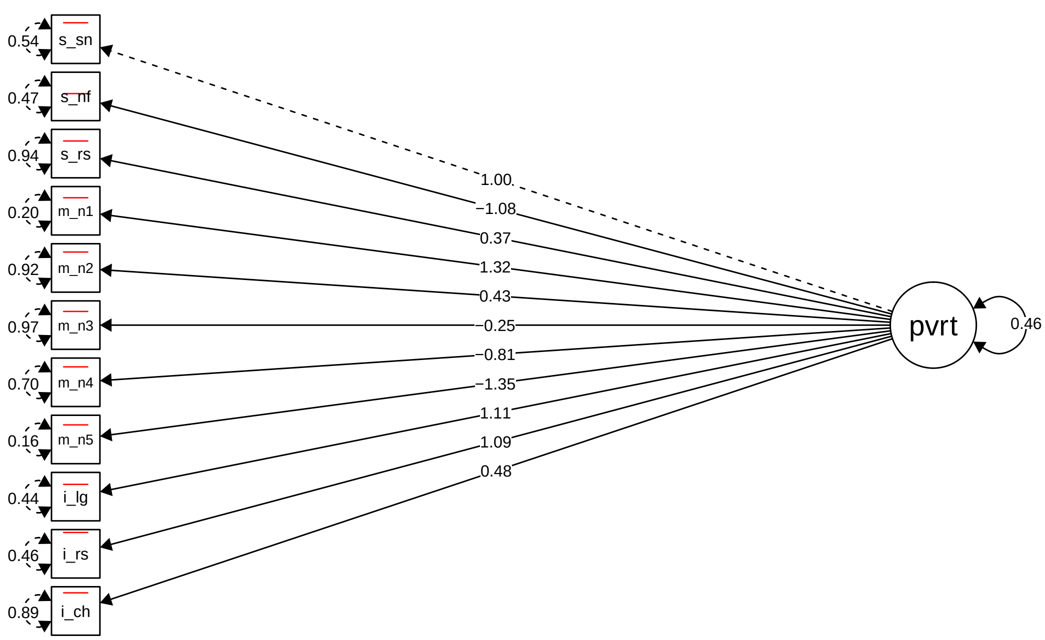




3.4 Confirmatory factor analysis (CFA) des dimensions de la pauvreté

EFA and CFA are mathematically very similar, since we have the same fundamental equation in both cases. In EFA we assumed uncorrelated factors by setting

3.4.1 Modèle sans dimensions de la pauvreté



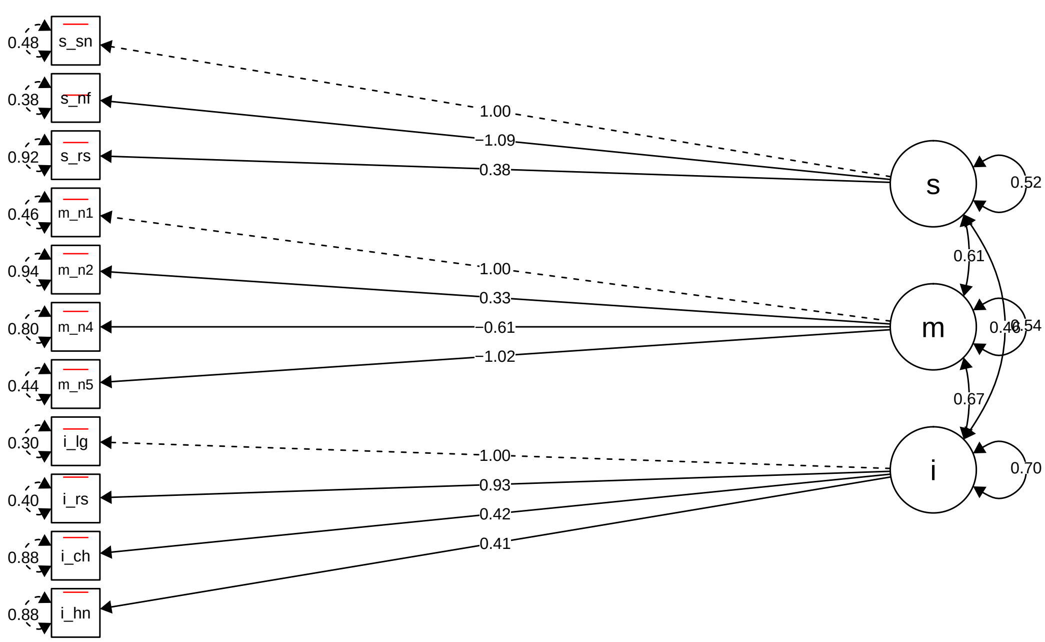

```
lavaan 0.6-8 ended normally after 32 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          22
```

```
  Number of observations                         10632

Model Test User Model:
                                              Standard      Robust
  Test Statistic                               963.997     542.800
  Degrees of freedom                                44          44
  P-value (Chi-square)                           0.000       0.000
  Scaling correction factor                                  1.855
  Shift parameter                                           23.260
       simple second-order correction

Model Test Baseline Model:

  Test statistic                             26883.282   14636.449
  Degrees of freedom                                55          55
  P-value                                        0.000       0.000
  Scaling correction factor                                  1.840

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.966       0.966
  Tucker-Lewis Index (TLI)                       0.957       0.957

  Robust Comparative Fit Index (CFI)                            NA
  Robust Tucker-Lewis Index (TLI)                               NA

Root Mean Square Error of Approximation:

  RMSEA                                          0.044       0.033
  90 Percent confidence interval - lower         0.042       0.030
  90 Percent confidence interval - upper         0.047       0.035
  P-value RMSEA <= 0.05                          1.000       1.000

  Robust RMSEA                                                  NA
  90 Percent confidence interval - lower                        NA
  90 Percent confidence interval - upper                        NA

Standardized Root Mean Square Residual:

  SRMR                                           0.361       0.361

Parameter Estimates:

  Standard errors                           Robust.sem
  Information                                 Expected
  Information saturated (h1) model        Unstructured

Latent Variables:
```

```
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  pauvrete =~
    s_sentpauv        1.000                               0.678    0.678
    s_infminidecla   -1.077    0.023  -47.555    0.000   -0.730   -0.730
    s_risqpauv        0.366    0.023   15.690    0.000    0.248    0.248
    m_nivie1          1.320    0.026   50.397    0.000    0.895    0.895
    m_nivie2          0.427    0.024   17.572    0.000    0.290    0.290
    m_nivie3         -0.250    0.023  -10.901    0.000   -0.169   -0.169
    m_nivie4         -0.805    0.026  -30.429    0.000   -0.546   -0.546
    m_nivie5         -1.351    0.031  -43.800    0.000   -0.916   -0.916
    i_log             1.105    0.023   48.199    0.000    0.749    0.749
    i_rsa             1.089    0.027   39.780    0.000    0.738    0.738
    i_chom            0.484    0.024   19.782    0.000    0.328    0.328

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .s_sentpauv        0.000                               0.000    0.000
   .s_infminidecla    0.000                               0.000    0.000
   .s_risqpauv        0.000                               0.000    0.000
   .m_nivie1          0.000                               0.000    0.000
   .m_nivie2          0.000                               0.000    0.000
   .m_nivie3          0.000                               0.000    0.000
   .m_nivie4          0.000                               0.000    0.000
   .m_nivie5          0.000                               0.000    0.000
   .i_log             0.000                               0.000    0.000
   .i_rsa             0.000                               0.000    0.000
   .i_chom            0.000                               0.000    0.000
    pauvrete          0.000                               0.000    0.000

Thresholds:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    s_sentpauv|t1     0.998    0.015   68.218    0.000    0.998    0.998
    s_infmindcl|t1    0.147    0.012   12.060    0.000    0.147    0.147
    s_risqpauv|t1     0.710    0.013   53.206    0.000    0.710    0.710
    m_nivie1|t1       0.735    0.013   54.735    0.000    0.735    0.735
    m_nivie2|t1       0.948    0.014   65.979    0.000    0.948    0.948
    m_nivie3|t1       0.781    0.014   57.419    0.000    0.781    0.781
    m_nivie4|t1       0.911    0.014   64.251    0.000    0.911    0.911
    m_nivie5|t1       0.846    0.014   60.991    0.000    0.846    0.846
    i_log|t1          0.686    0.013   51.743    0.000    0.686    0.686
    i_rsa|t1          1.563    0.019   80.419    0.000    1.563    1.563
    i_chom|t1         1.150    0.016   73.813    0.000    1.150    1.150

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .s_sentpauv        0.541                               0.541    0.541
   .s_infminidecla    0.467                               0.467    0.467
   .s_risqpauv        0.938                               0.938    0.938
   .m_nivie1          0.200                               0.200    0.200
```

```
   .m_nivie2          0.916                                     0.916    0.916
   .m_nivie3          0.971                                     0.971    0.971
   .m_nivie4          0.702                                     0.702    0.702
   .m_nivie5          0.162                                     0.162    0.162
   .i_log             0.439                                     0.439    0.439
   .i_rsa             0.456                                     0.456    0.456
   .i_chom            0.892                                     0.892    0.892
    pauvrete          0.459    0.016   28.933    0.000         1.000    1.000

Scales y*:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    s_sentpauv        1.000                                     1.000    1.000
    s_infminidecla    1.000                                     1.000    1.000
    s_risqpauv        1.000                                     1.000    1.000
    m_nivie1          1.000                                     1.000    1.000
    m_nivie2          1.000                                     1.000    1.000
    m_nivie3          1.000                                     1.000    1.000
    m_nivie4          1.000                                     1.000    1.000
    m_nivie5          1.000                                     1.000    1.000
    i_log             1.000                                     1.000    1.000
    i_rsa             1.000                                     1.000    1.000
    i_chom            1.000                                     1.000    1.000
```

3.4.2 Modèle avec 3 dimensions de la pauvreté (i,m,s)




```
lavaan 0.6-8 ended normally after 48 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          25
```

```
  Number of observations                         10632

Model Test User Model:
                                              Standard      Robust
  Test Statistic                               627.336     493.780
  Degrees of freedom                                41          41
  P-value (Chi-square)                           0.000       0.000
  Scaling correction factor                                  1.307
  Shift parameter                                           13.778
       simple second-order correction

Model Test Baseline Model:

  Test statistic                             27207.377   18634.349
  Degrees of freedom                                55          55
  P-value                                        0.000       0.000
  Scaling correction factor                                  1.461

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.978       0.976
  Tucker-Lewis Index (TLI)                       0.971       0.967

  Robust Comparative Fit Index (CFI)                            NA
  Robust Tucker-Lewis Index (TLI)                               NA

Root Mean Square Error of Approximation:

  RMSEA                                          0.037       0.032
  90 Percent confidence interval - lower         0.034       0.030
  90 Percent confidence interval - upper         0.039       0.035
  P-value RMSEA <= 0.05                          1.000       1.000

  Robust RMSEA                                                  NA
  90 Percent confidence interval - lower                        NA
  90 Percent confidence interval - upper                        NA

Standardized Root Mean Square Residual:

  SRMR                                           0.282       0.282

Parameter Estimates:

  Standard errors                           Robust.sem
  Information                                 Expected
  Information saturated (h1) model        Unstructured

Latent Variables:
```

```
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  s =~
    s_sentpauv        1.000                               0.718    0.718
    s_infminidecla   -1.094    0.023  -46.995    0.000   -0.786   -0.786
    s_risqpauv        0.384    0.023   16.543    0.000    0.276    0.276
  m =~
    m_nivie1          1.000                               0.735    0.735
    m_nivie2          0.332    0.019   17.868    0.000    0.244    0.244
    m_nivie4         -0.608    0.017  -35.196    0.000   -0.447   -0.447
    m_nivie5         -1.019    0.017  -61.290    0.000   -0.749   -0.749
  i =~
    i_log             1.000                               0.837    0.837
    i_rsa             0.927    0.022   42.606    0.000    0.776    0.776
    i_chom            0.423    0.020   20.645    0.000    0.353    0.353
    i_handi           0.412    0.023   17.725    0.000    0.345    0.345

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  s ~~
    m                 0.608    0.012   50.380    0.000    1.152    1.152
    i                 0.460    0.012   39.775    0.000    0.766    0.766
  m ~~
    i                 0.671    0.010   66.111    0.000    1.090    1.090

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .s_sentpauv        0.000                               0.000    0.000
   .s_infminidecla    0.000                               0.000    0.000
   .s_risqpauv        0.000                               0.000    0.000
   .m_nivie1          0.000                               0.000    0.000
   .m_nivie2          0.000                               0.000    0.000
   .m_nivie4          0.000                               0.000    0.000
   .m_nivie5          0.000                               0.000    0.000
   .i_log             0.000                               0.000    0.000
   .i_rsa             0.000                               0.000    0.000
   .i_chom            0.000                               0.000    0.000
   .i_handi           0.000                               0.000    0.000
    s                 0.000                               0.000    0.000
    m                 0.000                               0.000    0.000
    i                 0.000                               0.000    0.000

Thresholds:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    s_sentpauv|t1     0.998    0.015   68.218    0.000    0.998    0.998
    s_infmindcl|t1    0.147    0.012   12.060    0.000    0.147    0.147
    s_risqpauv|t1     0.710    0.013   53.206    0.000    0.710    0.710
    m_nivie1|t1       0.735    0.013   54.735    0.000    0.735    0.735
    m_nivie2|t1       0.948    0.014   65.979    0.000    0.948    0.948
    m_nivie4|t1       0.911    0.014   64.251    0.000    0.911    0.911
```

```
    m_nivie5|t1        0.846    0.014   60.991    0.000    0.846    0.846
    i_log|t1           0.686    0.013   51.743    0.000    0.686    0.686
    i_rsa|t1           1.563    0.019   80.419    0.000    1.563    1.563
    i_chom|t1          1.150    0.016   73.813    0.000    1.150    1.150
    i_handi|t1         1.453    0.018   79.896    0.000    1.453    1.453

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .s_sentpauv        0.484                                0.484    0.484
   .s_infminidecla    0.383                                0.383    0.383
   .s_risqpauv        0.924                                0.924    0.924
   .m_nivie1          0.459                                0.459    0.459
   .m_nivie2          0.940                                0.940    0.940
   .m_nivie4          0.800                                0.800    0.800
   .m_nivie5          0.438                                0.438    0.438
   .i_log             0.300                                0.300    0.300
   .i_rsa             0.398                                0.398    0.398
   .i_chom            0.875                                0.875    0.875
   .i_handi           0.881                                0.881    0.881
    s                 0.516    0.019   27.332    0.000    1.000    1.000
    m                 0.541    0.234    2.313    0.021    1.000    1.000
    i                 0.700    0.021   33.791    0.000    1.000    1.000

Scales y*:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    s_sentpauv        1.000                                1.000    1.000
    s_infminidecla    1.000                                1.000    1.000
    s_risqpauv        1.000                                1.000    1.000
    m_nivie1          1.000                                1.000    1.000
    m_nivie2          1.000                                1.000    1.000
    m_nivie4          1.000                                1.000    1.000
    m_nivie5          1.000                                1.000    1.000
    i_log             1.000                                1.000    1.000
    i_rsa             1.000                                1.000    1.000
    i_chom            1.000                                1.000    1.000
    i_handi           1.000                                1.000    1.000
```

- p-valeur du test du chi-2 de 0, très mauvais car un résultat non significatif veut dire que le modèle "fits" mais il ne faut pas faire très attention à cette statistique car elle est très souvent significative quand l'échantillon est grand.
- Le CFI doit être supérieur à 0,95. Ici 0,98 OK
- Le RMSEA doit être dans l'intervalle [0.05,0.10]. Ici [0,03, 0.04] pas tout à fait le cas donc.
- Le SRMR doit être inférieur à 0.08, pas du tout le cas ici 0,24.

3.4.3 Comment améliorer le modèle ?

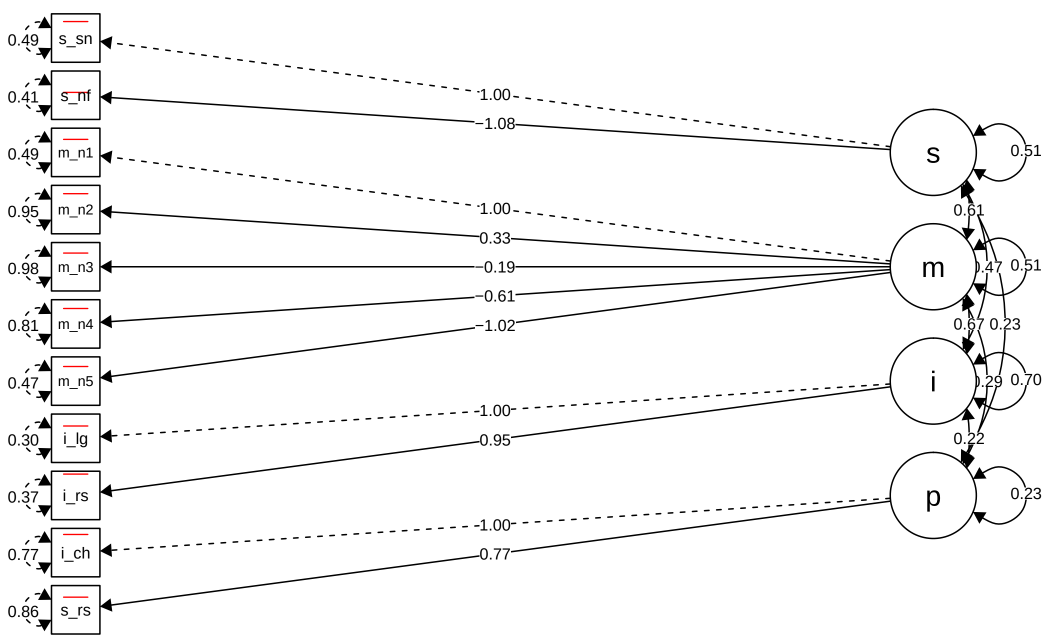
```
  lhs op            rhs    est    se       z pvalue ci.lower ci.upper
1   s =~     s_sentpauv  1.000 0.000      NA     NA    1.000    1.000
2   s =~ s_infminidecla -1.094 0.023 -46.995      0   -1.139   -1.048
3   s =~     s_risqpauv  0.384 0.023  16.543      0    0.338    0.429
```

```
4   m =~       m_nivie1  1.000 0.000      NA     NA    1.000    1.000
5   m =~       m_nivie2  0.332 0.019  17.868      0    0.296    0.368
6   m =~       m_nivie4 -0.608 0.017 -35.196      0   -0.642   -0.575
```

```
  id        lhs            rhs  nobs row col obs.freq obs.prop est.prop      X2
1  1 s_sentpauv s_infminidecla 10632   1   1     4464    0.420    0.419   0.034
2  1 s_sentpauv s_infminidecla 10632   2   1     1474    0.139    0.140   0.103
3  1 s_sentpauv s_infminidecla 10632   1   2     4477    0.421    0.422   0.034
4  1 s_sentpauv s_infminidecla 10632   2   2      217    0.020    0.019   0.746
5  2 s_sentpauv     s_risqpauv 10632   1   1     6400    0.602    0.656  47.207
6  2 s_sentpauv     s_risqpauv 10632   2   1     1691    0.159    0.105 294.669
```

```
          lhs  op        rhs        mi    epc sepc.lv sepc.all sepc.nox
76          m  =~      i_log 54307.649 69.986  51.465   51.465   51.465
1           s  =~ s_sentpauv 12575.682 23.967  17.218   17.218   17.218
40 s_sentpauv ~*~ s_sentpauv 12575.682 23.967  23.967    1.000    1.000
73          m  =~ s_sentpauv  6894.099 11.120   8.177    8.177    8.177
69          s  =~      i_log  6001.029  8.809   6.328    6.328    6.328
80          i  =~ s_sentpauv  2330.673  3.285   2.748    2.748    2.748
```

3.4.4 Quelques indicateurs écartés



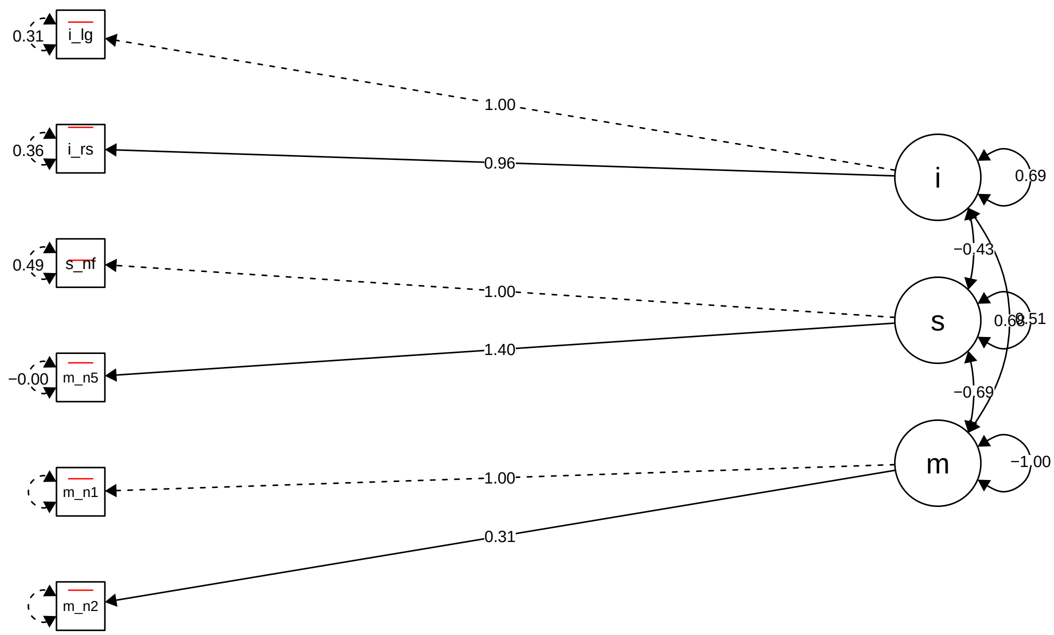

```
lavaan 0.6-8 ended normally after 49 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          28

  Number of observations                           10632

Model Test User Model:
```

```
                                                  Standard      Robust
  Test Statistic                                   842.670     471.213
  Degrees of freedom                                    38          38
  P-value (Chi-square)                               0.000       0.000
  Scaling correction factor                                      1.868
  Shift parameter                                               19.985
       simple second-order correction

Parameter Estimates:

  Standard errors                               Robust.sem
  Information                                     Expected
  Information saturated (h1) model             Unstructured

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  s =~
    s_sentpauv        1.000
    s_infminidecla   -1.081    0.023  -46.539    0.000
  m =~
    m_nivie1          1.000
    m_nivie2          0.326    0.019   17.484    0.000
    m_nivie3         -0.188    0.017  -11.099    0.000
    m_nivie4         -0.609    0.017  -35.067    0.000
    m_nivie5         -1.019    0.017  -60.847    0.000
  i =~
    i_log             1.000
    i_rsa             0.949    0.023   41.217    0.000
  p =~
    i_chom            1.000
    s_risqpauv        0.769    0.056   13.642    0.000

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  s ~~
    m                 0.611    0.012   49.610    0.000
    i                 0.465    0.012   37.843    0.000
    p                 0.232    0.015   15.696    0.000
  m ~~
    i                 0.673    0.010   65.976    0.000
    p                 0.287    0.015   19.141    0.000
  i ~~
    p                 0.219    0.016   13.691    0.000

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .s_sentpauv        0.000
   .s_infminidecla    0.000
   .m_nivie1          0.000
```

```
   .m_nivie2          0.000
   .m_nivie3          0.000
   .m_nivie4          0.000
   .m_nivie5          0.000
   .i_log             0.000
   .i_rsa             0.000
   .i_chom            0.000
   .s_risqpauv        0.000
    s                 0.000
    m                 0.000
    i                 0.000
    p                 0.000

Thresholds:
                   Estimate  Std.Err  z-value  P(>|z|)
    s_sentpauv|t1     0.998    0.015   68.218    0.000
    s_infmindcl|t1    0.147    0.012   12.060    0.000
    m_nivie1|t1       0.735    0.013   54.735    0.000
    m_nivie2|t1       0.948    0.014   65.979    0.000
    m_nivie3|t1       0.781    0.014   57.419    0.000
    m_nivie4|t1       0.911    0.014   64.251    0.000
    m_nivie5|t1       0.846    0.014   60.991    0.000
    i_log|t1          0.686    0.013   51.743    0.000
    i_rsa|t1          1.563    0.019   80.419    0.000
    i_chom|t1         1.150    0.016   73.813    0.000
    s_risqpauv|t1     0.710    0.013   53.206    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .s_sentpauv        0.491
   .s_infminidecla    0.405
   .m_nivie1          0.492
   .m_nivie2          0.946
   .m_nivie3          0.982
   .m_nivie4          0.812
   .m_nivie5          0.472
   .i_log             0.301
   .i_rsa             0.370
   .i_chom            0.768
   .s_risqpauv        0.863
    s                 0.509    0.019   26.616    0.000
    m                 0.508    0.216    2.352    0.019
    i                 0.699    0.023   30.747    0.000
    p                 0.232    0.030    7.818    0.000

Scales y*:
                   Estimate  Std.Err  z-value  P(>|z|)
    s_sentpauv        1.000
    s_infminidecla    1.000
```

```
    m_nivie1         1.000
    m_nivie2         1.000
    m_nivie3         1.000
    m_nivie4         1.000
    m_nivie5         1.000
    i_log            1.000
    i_rsa            1.000
    i_chom           1.000
    s_risqpauv       1.000
```

3.4.5 Carrément à partir de l'EFA

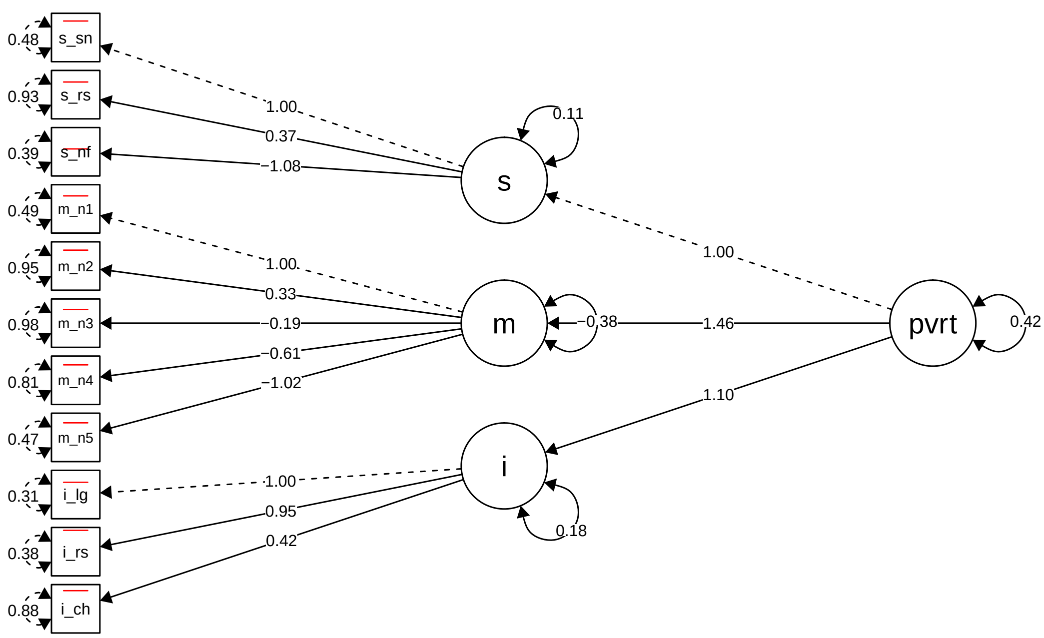



```
lavaan 0.6-8 did NOT end normally after 74 iterations
** WARNING ** Estimates below are most likely unreliable

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          15

  Number of observations                           10632

Model Test User Model:

  Test statistic                                      NA
  Degrees of freedom                                  NA

Parameter Estimates:

  Standard errors                             Robust.sem
  Information                                   Expected
```

```
  Information saturated (h1) model        Unstructured

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  i =~
    i_log             1.000
    i_rsa             0.960       NA
  s =~
    s_infminidecla    1.000
    m_nivie5          1.404       NA
  m =~
    m_nivie1          1.000
    m_nivie2          0.308       NA

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  i ~~
    s                -0.431       NA
    m                 0.682       NA
  s ~~
    m                -0.692       NA

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .i_log             0.000
   .i_rsa             0.000
   .s_infminidecla    0.000
   .m_nivie5          0.000
   .m_nivie1          0.000
   .m_nivie2          0.000
    i                 0.000
    s                 0.000
    m                 0.000

Thresholds:
                   Estimate  Std.Err  z-value  P(>|z|)
    i_log|t1          0.686       NA
    i_rsa|t1          1.563       NA
    s_infmindcl|t1    0.147       NA
    m_nivie5|t1       0.846       NA
    m_nivie1|t1       0.735       NA
    m_nivie2|t1       0.948       NA

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .i_log             0.308
   .i_rsa             0.363
   .s_infminidecla    0.492
   .m_nivie5         -0.001
```

```
    .m_nivie1          2.000
    .m_nivie2          1.095
     i                 0.692       NA
     s                 0.508       NA
     m                -1.000       NA

Scales y*:
                   Estimate  Std.Err  z-value  P(>|z|)
    i_log              1.000
    i_rsa              1.000
    s_infminidecla     1.000
    m_nivie5           1.000
    m_nivie1           1.000
    m_nivie2           1.000
```

3.4.6 CFA hiérarchique

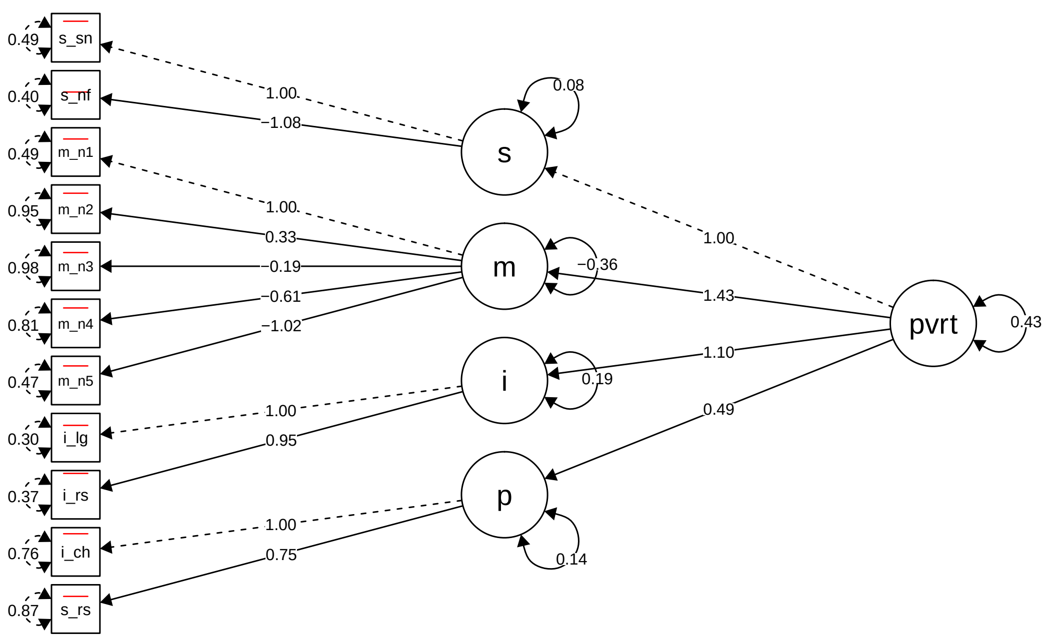


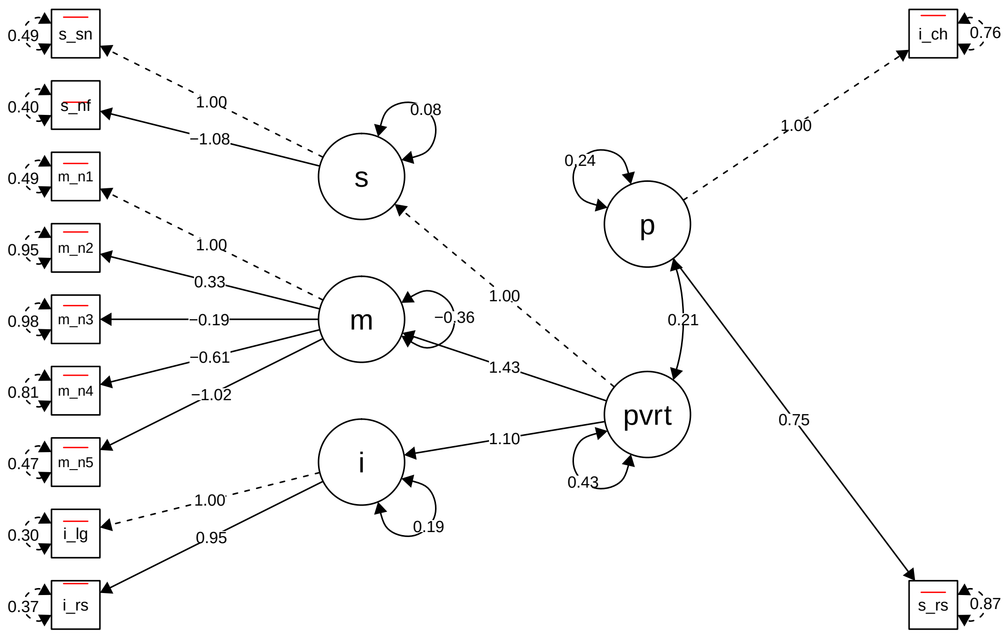

```
lavaan 0.6-8 ended normally after 48 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          25

  Number of observations                           10632

Model Test User Model:
                                              Standard      Robust
  Test Statistic                               844.834     485.183
  Degrees of freedom                                41          41
  P-value (Chi-square)                           0.000       0.000
```

```
  Scaling correction factor                                  1.820
  Shift parameter                                           21.025
       simple second-order correction

Model Test Baseline Model:

  Test statistic                             26883.282   14636.449
  Degrees of freedom                                55          55
  P-value                                        0.000       0.000
  Scaling correction factor                                  1.840

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.970       0.970
  Tucker-Lewis Index (TLI)                       0.960       0.959

  Robust Comparative Fit Index (CFI)                            NA
  Robust Tucker-Lewis Index (TLI)                               NA

Root Mean Square Error of Approximation:

  RMSEA                                          0.043       0.032
  90 Percent confidence interval - lower         0.040       0.029
  90 Percent confidence interval - upper         0.045       0.035
  P-value RMSEA <= 0.05                          1.000       1.000

  Robust RMSEA                                                  NA
  90 Percent confidence interval - lower                        NA
  90 Percent confidence interval - upper                        NA

Standardized Root Mean Square Residual:

  SRMR                                           0.356       0.356

Parameter Estimates:

  Standard errors                           Robust.sem
  Information                                 Expected
  Information saturated (h1) model          Unstructured

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  s =~
    s_sentpauv        1.000                               0.723    0.723
    s_risqpauv        0.370    0.023   16.032    0.000    0.268    0.268
    s_infminidecla   -1.083    0.023  -46.589    0.000   -0.783   -0.783
  m =~
    m_nivie1          1.000                               0.713    0.713
    m_nivie2          0.326    0.019   17.518    0.000    0.232    0.232
```

```
    m_nivie3         -0.187    0.017  -11.072    0.000   -0.133   -0.133
    m_nivie4         -0.609    0.017  -35.076    0.000   -0.434   -0.434
    m_nivie5         -1.019    0.017  -60.940    0.000   -0.727   -0.727
  i =~
    i_log             1.000                               0.832    0.832
    i_rsa             0.949    0.023   41.747    0.000    0.789    0.789
    i_chom            0.424    0.021   20.400    0.000    0.353    0.353
  pauvrete =~
    s                 1.000                               0.892    0.892
    m                 1.464    0.034   43.339    0.000    1.325    1.325
    i                 1.104    0.024   45.324    0.000    0.856    0.856

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .s_sentpauv        0.000                               0.000    0.000
   .s_risqpauv        0.000                               0.000    0.000
   .s_infminidecla    0.000                               0.000    0.000
   .m_nivie1          0.000                               0.000    0.000
   .m_nivie2          0.000                               0.000    0.000
   .m_nivie3          0.000                               0.000    0.000
   .m_nivie4          0.000                               0.000    0.000
   .m_nivie5          0.000                               0.000    0.000
   .i_log             0.000                               0.000    0.000
   .i_rsa             0.000                               0.000    0.000
   .i_chom            0.000                               0.000    0.000
   .s                 0.000                               0.000    0.000
   .m                 0.000                               0.000    0.000
   .i                 0.000                               0.000    0.000
    pauvrete          0.000                               0.000    0.000

Thresholds:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    s_sentpauv|t1     0.998    0.015   68.218    0.000    0.998    0.998
    s_risqpauv|t1     0.710    0.013   53.206    0.000    0.710    0.710
    s_infmindcl|t1    0.147    0.012   12.060    0.000    0.147    0.147
    m_nivie1|t1       0.735    0.013   54.735    0.000    0.735    0.735
    m_nivie2|t1       0.948    0.014   65.979    0.000    0.948    0.948
    m_nivie3|t1       0.781    0.014   57.419    0.000    0.781    0.781
    m_nivie4|t1       0.911    0.014   64.251    0.000    0.911    0.911
    m_nivie5|t1       0.846    0.014   60.991    0.000    0.846    0.846
    i_log|t1          0.686    0.013   51.743    0.000    0.686    0.686
    i_rsa|t1          1.563    0.019   80.419    0.000    1.563    1.563
    i_chom|t1         1.150    0.016   73.813    0.000    1.150    1.150

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .s_sentpauv        0.477                               0.477    0.477
   .s_risqpauv        0.928                               0.928    0.928
   .s_infminidecla    0.387                               0.387    0.387
```

```
    .m_nivie1         0.492                                       0.492    0.492
    .m_nivie2         0.946                                       0.946    0.946
    .m_nivie3         0.982                                       0.982    0.982
    .m_nivie4         0.812                                       0.812    0.812
    .m_nivie5         0.472                                       0.472    0.472
    .i_log            0.308                                       0.308    0.308
    .i_rsa            0.377                                       0.377    0.377
    .i_chom           0.876                                       0.876    0.876
    .s                0.107    0.014    7.816    0.000    0.204    0.204
    .m               -0.384    0.216   -1.773    0.076   -0.755   -0.755
    .i                0.185    0.020    9.057    0.000    0.267    0.267
     pauvrete         0.416    0.016   26.469    0.000    1.000    1.000

Scales y*:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    s_sentpauv        1.000                               1.000    1.000
    s_risqpauv        1.000                               1.000    1.000
    s_infminidecla    1.000                               1.000    1.000
    m_nivie1          1.000                               1.000    1.000
    m_nivie2          1.000                               1.000    1.000
    m_nivie3          1.000                               1.000    1.000
    m_nivie4          1.000                               1.000    1.000
    m_nivie5          1.000                               1.000    1.000
    i_log             1.000                               1.000    1.000
    i_rsa             1.000                               1.000    1.000
    i_chom            1.000                               1.000    1.000
```

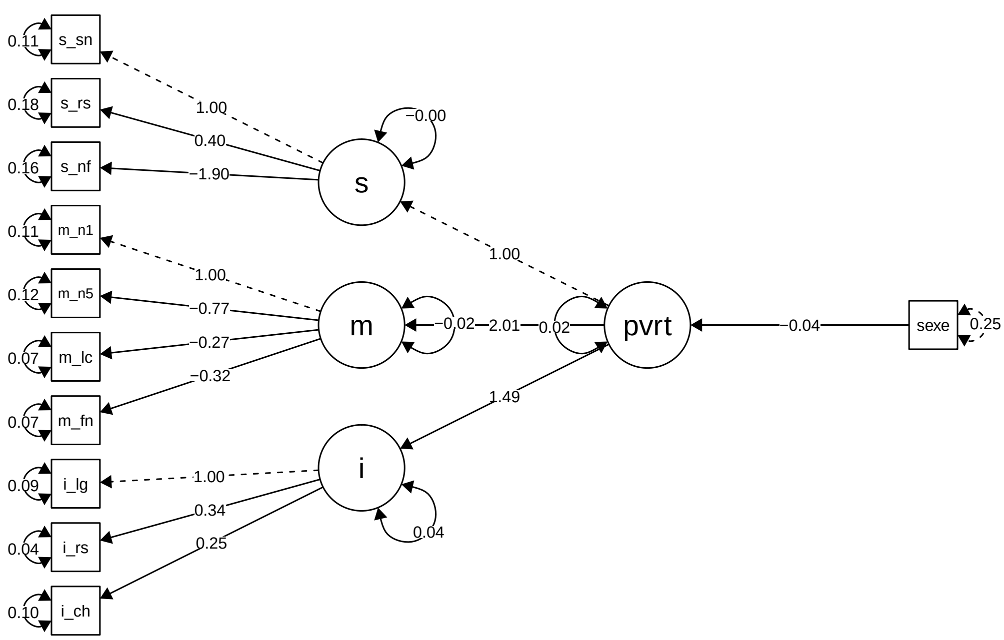




3.5 CFA avec des covariables (MIMIC)

MIMIC stands for multiple indicators multiple independent causes (Jöreskog and Goldberger, 1975) and is a general structural latent variable concept where CFA is extended in terms of linking covariates with latent variables. MIMIC models can be used to control for sociodemographic or other types of covariates in CFA and more general SEM specifications.

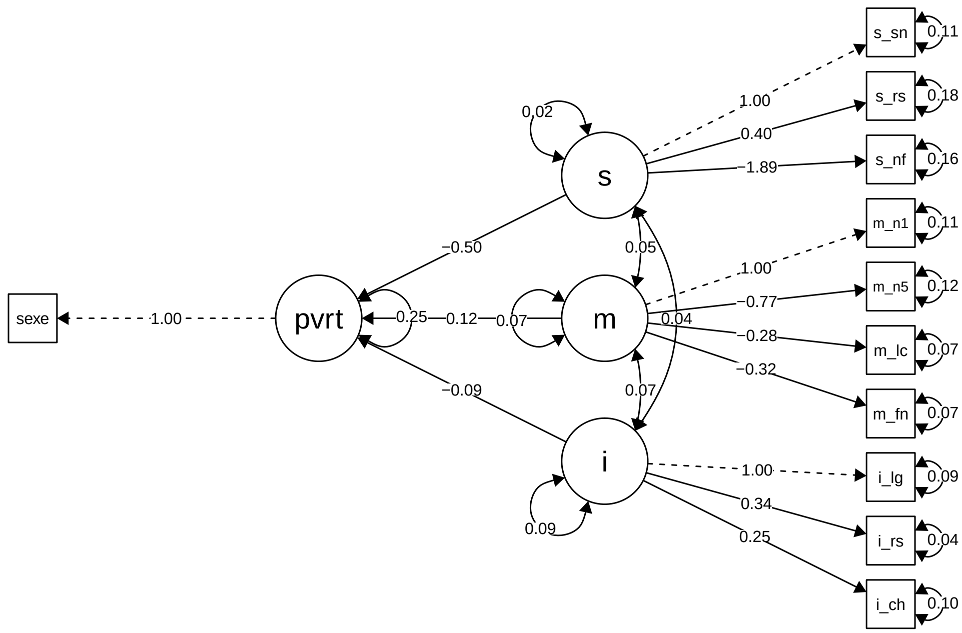
Remarque : ne marche qu'avec les covariates exogènes de moins de 2 facteurs (à transformer en indicatrices j'imagine)



3.6 Structural equation models (SEM)

Structural equation models (SEM) integrate confirmatory factor analysis (CFA) into a larger path analytic framework. Formally, we extend the basic CFA expression (measurement model) by an additional linear specification reflecting dependencies among the latent variables (structural model).

Remarque : ne marche pas pour les facteurs non ordonnés (en gros, considère les facteurs comme des variables numériques)



4 Notes méthodologiques

Pour ces modèles quatre vagues du Baromètre ont été empilées : 2016, 2017, 2018 et 2019 (12 114 observations). Le nombre d'observations utilisées est différent dans chaque modèle, il s'agit uniquement des individus où toutes les variables utilisées dans les modèles sont renseignées (voir notes en bas des tableaux).

Bibliographie

— https://stats.idre.ucla.edu/spss/seminars/efa-spss/ https://support.sas.com/resources/papers/proceedings/proceedings/sugi30/203-30.pdf https://community.jmp.com/t5/JMP-Blog/Principal-components-or-factor-analysis/ba-p/38347 bases de l'EFA

— En bouquins : https://books.google.es/books?hl=fr&lr=&id=qKrumJ4CsboC&oi=fnd&pg=PT180&ots=TDmmzvQP5X&sig=7gFjzxbPC49Tz7IkGT-4gXMzx8U&redir_esc=y#v=onepage&q&f=false et slides https://slideplayer.com/slide/5080/

— https://m-clark.github.io/posts/2020-04-10-psych-explained/

— https://cran.r-project.org/web/packages/psychTools/vignettes/factor.pdf

— https://rstudio-pubs-static.s3.amazonaws.com/363499_73a1c1a94da148b6ad81e6eb8dc1b771.html

— https://en.wikipedia.org/wiki/Factor_analysis

— Analyse en facteurs communs et spécifiques docs en Français. https://www.rocq.inria.fr/axis/m odulad/archives/numero-37/Chaventetal-37/Chaventetal-37.pdf http://grumlidesforets.free.fr/cours%20psycho/M1%20psycho/chapitre2/chapitre2.pdf http://jeanalain.monfort.free.fr/Dicostat2005/A/Analyse_en_facteurs_communs_etc.pdf https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjX-NWssbbxAhUNxoUKHXMzBxEQFnoECAkQAA&url=http%3A%2F%2Fwww.normalesup.org%2F~carpenti%2FNotes%2FAnalyse-factorielle%2FAnalyse-Factorielle-2011.doc&usg=AOvVaw04RFWMowmry0JRVMNZqR7h http://jeanalain.monfort.free.fr/Dicostat2005/A/Analyse_en_facteurs_communs_etc.pdf https://www.psychometrie.jlroulin.fr/cours/aide_quizz.html?H.html https://www.persee.fr/doc/hism_0982-1783_1997_num_12_3_1544 http://psychologie.psyblogs.net/2012/01/cours-theories-de-lintelligence-en.html