LSAT by Riccardo Petrella

First of all load the library and the data set

```
library(ltm)
```

```
## Warning: package 'ltm' was built under R version 4.3.2

## Loading required package: MASS

## Loading required package: msm

## Warning: package 'msm' was built under R version 4.3.2

## Loading required package: polycor

## Warning: package 'polycor' was built under R version 4.3.2
```

```
data(LSAT)
```

```
dsc <- descript(LSAT)
```

Percentage of positive and negative responses

```
dsc$perc
```

```
##            0     1     logit
## Item 1 0.076 0.924 2.4979787
## Item 2 0.291 0.709 0.8905323
## Item 3 0.447 0.553 0.2127994
## Item 4 0.237 0.763 1.1691979
## Item 5 0.130 0.870 1.9009588
```

The percentage of positive response of the FIRST item is very high (also the logit is the highest), followed by the FIFTH item while the THIRD item is the one where we have half positive and half negative proportion of responses.

Information about the n° of people that has -all zeros, -at least a one, - twice one…

```
dsc$items
```

```
##      0  1  2   3   4   5
## Freq 3 20 85 237 357 298
```

For example only three out the total amount of people didn't reply correctly to any item.

Then we can ask for the pairwise association that give us the p-value of the chi square for the test of independence for each 2x2 table. It is used to verify that the items are associated

```r
dsc$pw.ass
```

```
##    Item i Item j p.value
## 1       1      5   0.565
## 2       1      4   0.208
## 3       3      5   0.113
## 4       2      4   0.059
## 5       1      2   0.028
## 6       2      5   0.009
## 7       1      3   0.003
## 8       4      5   0.002
## 9       3      4   7e-04
## 10      2      3   4e-04
```

Not all items are associated: for example items 1-5, 1-4, 3-5, 2-4 aren't associated. This is not the best situation but we can fit the model anyway and try to explain the fact that the model isn't so good by exploiting these information.

First thing that can be done consist in fitting a Rasch model. A model where we assume that the discriminant parameters are all equal to 1 and we estimate only the difficulty parameters. There is an option in the definition of the model which is "constraint" and through this option it is possible to fix the discriminant parameter to a particular and fixed value. if the option constraint=NULL that the parameters are estimated under the assumption that the estimated parameters are ALL EQUAL.

Classical Rasch Model

```r
m1<-rasch(data=LSAT, constraint = cbind(ncol(LSAT) + 1, 1), IRT.param=TRUE)
#constraint = cbind(ncol(LSAT) + 1, 1 ---> this mean that I want the parameter fixed to 1 but I
#can choose the value that I want. If I don't specify anything then it estimate the value but
#all the discriminant parameters are equal)
```

Alternative parametrisation of the Rasch model - GLLVM

```r
m1.rip<-rasch(LSAT, IRT.param=FALSE,constraint = cbind(ncol(LSAT) + 1, 1))
summary(m1) #The discrimination parameter is fixed to 1 and so we don't have the estimation of
```

```
##
## Call:
## rasch(data = LSAT, constraint = cbind(ncol(LSAT) + 1, 1), IRT.param = TRUE)
##
## Model Summary:
##    log.Lik      AIC      BIC
##  -2473.054 4956.108 4980.646
##
## Coefficients:
##                  value std.err   z.vals
## Dffclt.Item 1 -2.8720  0.1287 -22.3066
## Dffclt.Item 2 -1.0630  0.0821 -12.9458
## Dffclt.Item 3 -0.2576  0.0766  -3.3635
## Dffclt.Item 4 -1.3881  0.0865 -16.0478
## Dffclt.Item 5 -2.2188  0.1048 -21.1660
## Dscrmn         1.0000      NA       NA
##
```

```
## Integration:
## method: Gauss-Hermite
## quadrature points: 21
##
## Optimization:
## Convergence: 0
## max(|grad|): 6.3e-05
## quasi-Newton: BFGS
```

```
# the Z and of the std err.
# The only estimated parameters are the difficulty ones and since we are in the IRT
# parametrisation the lowest(highest) parameter correspond to the easiest(difficult) item
summary(m1.rip)
```

```
##
## Call:
## rasch(data = LSAT, constraint = cbind(ncol(LSAT) + 1, 1), IRT.param = FALSE)
##
## Model Summary:
##    log.Lik      AIC      BIC
##  -2473.054 4956.108 4980.646
##
## Coefficients:
##        value std.err  z.vals
## Item1 2.8720  0.1287 22.3066
## Item2 1.0630  0.0821 12.9458
## Item3 0.2576  0.0766  3.3635
## Item4 1.3881  0.0865 16.0478
## Item5 2.2188  0.1048 21.1660
## z     1.0000      NA      NA
##
## Integration:
## method: Gauss-Hermite
## quadrature points: 21
##
## Optimization:
## Convergence: 0
## max(|grad|): 6.3e-05
## quasi-Newton: BFGS
```

```
# All the values are exactly the same, the only thing that change are the sign of the difficulty
# parameter.
```

We have to compute the probability of the median individual given the latent variable equal to 0

```
coef(m1.rip,prob=TRUE,order=TRUE)
```

```
##           beta.i beta P(x=1|z=0)
## Item 3 0.2576109    1  0.5640489
## Item 2 1.0630294    1  0.7432690
## Item 4 1.3880588    1  0.8002822
## Item 5 2.2187785    1  0.9019232
## Item 1 2.8719712    1  0.9464434
```

```
coef(m1,prob=TRUE,order=TRUE) #
```

```
##            Dffclt Dscrmn P(x=1|z=0)
## Item 1 -2.8719712      1  0.9464434
## Item 5 -2.2187785      1  0.9019232
## Item 4 -1.3880588      1  0.8002822
## Item 2 -1.0630294      1  0.7432690
## Item 3 -0.2576109      1  0.5640489
```

The solution is coherent with the one that we found with the comment of the difficulty parameters above.

GOODNESS OF FIT It is likely to have sparse data in this case. In the rash case we have a particular command that we can use to verify the goodness of the model due to the fact that the asymptotic results of the chi-square and LRT don't always hold. The solution consist in using the empirical bootstrap distribution of the Pearson chi-square. This bootstrap option is implemented in the ltm packages only for the rash model.

```
pval.boot<-GoF.rasch(m1,B=199)
pval.boot$Tobs #Empirical value of the Pearson chi-square distribution
```

```
## [1] 30.59541
```

```
pval.boot #P-value
```

```
##
## Bootstrap Goodness-of-Fit using Pearson chi-squared
##
## Call:
## rasch(data = LSAT, constraint = cbind(ncol(LSAT) + 1, 1), IRT.param = TRUE)
##
## Tobs: 30.6
## # data-sets: 200
## p-value: 0.235
```

So according to the bootstrap distribution of this statistics (Pearson chi square), we do not reject the null hypothesis of the one factor model.

It is also possible to compute the asymptotic p-value

```
pval.asint1<-pchisq(pval.boot$Tobs, df=30-1-ncol(LSAT)) #We should put the empirical/observed
#degrees of freedom and not the theoretical ones (30 instead of 32=2^p). So the DF=30-1-n°of
#parameters estimated
pval.asint1
```

```
## [1] 0.8342059
```

Also in this case it is significant, we accept H0.

Even if from the descriptive analysis we found that some items are not associated, this model is still a good model.

We can compute the margins in order to verify the LOCAL goodness of fit of the model. Because even if the model overall is a good model it is possible that there are some problems in terms of residuals (we won't reject the model but we will improve it)

Two-way margins

```r
margins(m1)
```

```
##
## Call:
## rasch(data = LSAT, constraint = cbind(ncol(LSAT) + 1, 1), IRT.param = TRUE)
##
## Fit on the Two-Way Margins
##
## Response: (0,0)
##   Item i Item j Obs   Exp (O-E)^2/E
## 1      2      4  81 98.69      3.17
## 2      1      5  12 18.45      2.25
## 3      3      5  67 80.04      2.12
##
## Response: (1,0)
##   Item i Item j Obs    Exp (O-E)^2/E
## 1      3      5  63  51.62      2.51
## 2      2      4 156 139.78      1.88
## 3      3      4 108  99.42      0.74
##
## Response: (0,1)
##   Item i Item j Obs    Exp (O-E)^2/E
## 1      2      4 210 193.47      1.41
## 2      2      3 135 125.07      0.79
## 3      1      4  53  47.24      0.70
##
## Response: (1,1)
##   Item i Item j Obs    Exp (O-E)^2/E
## 1      2      4 553 568.06      0.40
## 2      3      5 490 501.43      0.26
## 3      2      3 418 427.98      0.23
```

They are all good. All the pairwise association between the item is explaind by the model.

Three-way margins

```r
margins(m1,type="three-way",nprint=2)
```

```
##
## Call:
## rasch(data = LSAT, constraint = cbind(ncol(LSAT) + 1, 1), IRT.param = TRUE)
##
## Fit on the Three-Way Margins
##
## Response: (0,0,0)
##   Item i Item j Item k Obs   Exp (O-E)^2/E
## 1      2      3      4  48 66.07      4.94 ***
## 2      1      3      5   6 13.58      4.23 ***
##
## Response: (1,0,0)
##   Item i Item j Item k Obs   Exp (O-E)^2/E
## 1      1      2      4  70 82.01      1.76
## 2      2      4      5  28 22.75      1.21
```

```
##
## Response: (0,1,0)
##   Item i Item j Item k Obs   Exp (O-E)^2/E
## 1      1      2      5   3  7.73      2.90
## 2      3      4      5  37 45.58      1.61
##
## Response: (1,1,0)
##   Item i Item j Item k Obs    Exp (O-E)^2/E
## 1      3      4      5  48  36.91      3.33
## 2      1      2      4 144 126.35      2.47
##
## Response: (0,0,1)
##   Item i Item j Item k Obs   Exp (O-E)^2/E
## 1      1      3      5  41 34.58      1.19
## 2      2      4      5  64 72.26      0.94
##
## Response: (1,0,1)
##   Item i Item j Item k Obs    Exp (O-E)^2/E
## 1      1      2      4 190 174.87      1.31
## 2      1      2      3 126 114.66      1.12
##
## Response: (0,1,1)
##   Item i Item j Item k Obs   Exp (O-E)^2/E
## 1      1      2      5  42 34.35      1.70
## 2      1      4      5  46 38.23      1.58
##
## Response: (1,1,1)
##   Item i Item j Item k Obs    Exp (O-E)^2/E
## 1      3      4      5 397 416.73      0.93
## 2      2      3      4 343 361.18      0.91
##
## '***' denotes a chi-squared residual greater than 3.5
```

```r
margins(m1,type="three-way",nprint=3)
```

```
##
## Call:
## rasch(data = LSAT, constraint = cbind(ncol(LSAT) + 1, 1), IRT.param = TRUE)
##
## Fit on the Three-Way Margins
##
## Response: (0,0,0)
##   Item i Item j Item k Obs   Exp (O-E)^2/E
## 1      2      3      4  48 66.07      4.94 ***
## 2      1      3      5   6 13.58      4.23 ***
## 3      2      4      5  17 26.43      3.36
##
## Response: (1,0,0)
##   Item i Item j Item k Obs   Exp (O-E)^2/E
## 1      1      2      4  70 82.01      1.76
## 2      2      4      5  28 22.75      1.21
## 3      2      3      4  81 72.98      0.88
##
## Response: (0,1,0)
```

```
##   Item i Item j Item k Obs   Exp (O-E)^2/E
## 1      1      2      5   3  7.73      2.90
## 2      3      4      5  37 45.58      1.61
## 3      1      3      4   5  8.61      1.51
##
## Response: (1,1,0)
##   Item i Item j Item k Obs    Exp (O-E)^2/E
## 1      3      4      5  48  36.91      3.33
## 2      1      2      4 144 126.35      2.47
## 3      1      3      5  57  46.76      2.24
##
## Response: (0,0,1)
##   Item i Item j Item k Obs    Exp (O-E)^2/E
## 1      1      3      5  41  34.58      1.19
## 2      2      4      5  64  72.26      0.94
## 3      2      3      4 108 101.01      0.48
##
## Response: (1,0,1)
##   Item i Item j Item k Obs    Exp (O-E)^2/E
## 1      1      2      4 190 174.87      1.31
## 2      1      2      3 126 114.66      1.12
## 3      2      4      5 128 117.03      1.03
##
## Response: (0,1,1)
##   Item i Item j Item k Obs    Exp (O-E)^2/E
## 1      1      2      5  42  34.35      1.70
## 2      1      4      5  46  38.23      1.58
## 3      3      4      5 281 262.32      1.33
##
## Response: (1,1,1)
##   Item i Item j Item k Obs    Exp (O-E)^2/E
## 1      3      4      5 397 416.73      0.93
## 2      2      3      4 343 361.18      0.91
## 3      2      3      5 377 394.23      0.75
##
## '***' denotes a chi-squared residual greater than 3.5
```

There are some problems for items (2,3,4) and (1,3,5). The model doesn't explain completely the association between some items.

We can improve it by: -relaxing the assumption that the discriminant parameter is = 1 for each item. -we estimate the discriminant parameter but we assume that they are all equal for each individual. So we assume that all the items discriminate in the same way among smart and not smart individuals.

IRT parametrisation

```
m2<-rasch(LSAT)
summary(m2)
```

```
##
## Call:
## rasch(data = LSAT)
##
## Model Summary:
```

```
##    log.Lik      AIC      BIC
##  -2466.938 4945.875 4975.322
##
## Coefficients:
##                  value std.err   z.vals
## Dffclt.Item 1 -3.6153  0.3266 -11.0680
## Dffclt.Item 2 -1.3224  0.1422  -9.3009
## Dffclt.Item 3 -0.3176  0.0977  -3.2518
## Dffclt.Item 4 -1.7301  0.1691 -10.2290
## Dffclt.Item 5 -2.7802  0.2510 -11.0743
## Dscrmn         0.7551  0.0694  10.8757
##
## Integration:
## method: Gauss-Hermite
## quadrature points: 21
##
## Optimization:
## Convergence: 0
## max(|grad|): 2.9e-05
## quasi-Newton: BFGS
```

GLLVM parametrisation

```r
m2.rip<-rasch(LSAT,IRT.param=FALSE)
summary(m2.rip)
```

```
##
## Call:
## rasch(data = LSAT, IRT.param = FALSE)
##
## Model Summary:
##    log.Lik      AIC      BIC
##  -2466.938 4945.875 4975.322
##
## Coefficients:
##        value std.err  z.vals
## Item1 2.7300  0.1304 20.9291
## Item2 0.9986  0.0792 12.6123
## Item3 0.2399  0.0718  3.3418
## Item4 1.3065  0.0846 15.4357
## Item5 2.0994  0.1054 19.9099
## z     0.7551  0.0694 10.8757
##
## Integration:
## method: Gauss-Hermite
## quadrature points: 21
##
## Optimization:
## Convergence: 0
## max(|grad|): 2.9e-05
## quasi-Newton: BFGS
```

In this case the results of the two different parametrisation are a little bit different (but the order is the same) because since we don't fix the value of the discrimination parameter there is a different estimation procedure that depends also on the discrimination parameter.

To go from one parametrisation to the other the formula are:

```
summary(m2)$coefficients[1,1]*summary(m2)$coefficients[6,1] #To get the GLLVM parameter
```

```
## [1] -2.730013
```

```
summary(m2.rip)$coefficients[1,1] #To get the IRT parameter
```

```
## [1] 2.730013
```

In this package for the Rash model there are a lot of possibilities like for example it is possible to implement the ANOVA TEST. In this case the ANOVA TEST check the hypothesis that all the discriminant parameters are equal and fixed to 1 in the population.

```
anova(m1,m2) #m1 th H0 hypothesis and m2 is the H1 one.
```

```
##
##  Likelihood Ratio Table
##        AIC     BIC  log.Lik   LRT df p.value
## m1 4956.11 4980.65 -2473.05
## m2 4945.88 4975.32 -2466.94 12.23  1  <0.001
```

We reject the hypothesis under model m1 –> all the discriminant parameters are not all equal to 1, m2 is a better model than m1 and this can be seen also from AIC and BIC.

Now we can estimate the **latent trait model**.

IRT parametrisation.

```
m3<-ltm(LSAT ~ z1)
summary(m3)
```

```
##
## Call:
## ltm(formula = LSAT ~ z1)
##
## Model Summary:
##    log.Lik      AIC      BIC
##  -2466.653 4953.307 5002.384
##
## Coefficients:
##                 value std.err  z.vals
## Dffclt.Item 1 -3.3597  0.8669 -3.8754
## Dffclt.Item 2 -1.3696  0.3073 -4.4565
## Dffclt.Item 3 -0.2799  0.0997 -2.8083
## Dffclt.Item 4 -1.8659  0.4341 -4.2982
## Dffclt.Item 5 -3.1236  0.8700 -3.5904
## Dscrmn.Item 1  0.8254  0.2581  3.1983
## Dscrmn.Item 2  0.7229  0.1867  3.8721
```

```
## Dscrmn.Item 3  0.8905  0.2326  3.8281
## Dscrmn.Item 4  0.6886  0.1852  3.7186
## Dscrmn.Item 5  0.6575  0.2100  3.1306
##
## Integration:
## method: Gauss-Hermite
## quadrature points: 21
##
## Optimization:
## Convergence: 0
## max(|grad|): 0.024
## quasi-Newton: BFGS
```

The discrimination parameters are all estimated now and they are all significant.

GLLVM parametrisation

```r
m3.rip<-ltm(LSAT ~ z1, IRT.param=FALSE)
summary(m3.rip)
```

```
##
## Call:
## ltm(formula = LSAT ~ z1, IRT.param = FALSE)
##
## Model Summary:
##    log.Lik      AIC      BIC
##  -2466.653 4953.307 5002.384
##
## Coefficients:
##                      value std.err  z.vals
## (Intercept).Item 1 2.7730  0.2057 13.4824
## (Intercept).Item 2 0.9902  0.0900 10.9987
## (Intercept).Item 3 0.2492  0.0763  3.2681
## (Intercept).Item 4 1.2848  0.0990 12.9711
## (Intercept).Item 5 2.0536  0.1354 15.1620
## z1.Item 1          0.8254  0.2581  3.1983
## z1.Item 2          0.7229  0.1867  3.8721
## z1.Item 3          0.8905  0.2326  3.8281
## z1.Item 4          0.6886  0.1852  3.7186
## z1.Item 5          0.6575  0.2100  3.1306
##
## Integration:
## method: Gauss-Hermite
## quadrature points: 21
##
## Optimization:
## Convergence: 0
## max(|grad|): 0.024
## quasi-Newton: BFGS
```

The values of the discrimination parameters are the same as in the IRT parametrisation. We can comment these parameters as loadings then we have that the first three items have higher loading in ability then the last ones. In education topic it is more natural to interpret them as discriminant and difficulty parameters.

ANOVA TEST Under H0 holds m2 while the alternative hypothesis is that these estimated parameters are significantly different one to each other

```r
anova(m2,m3)
```

```
##
##  Likelihood Ratio Table
##        AIC     BIC  log.Lik  LRT df p.value
## m2 4945.88 4975.32 -2466.94
## m3 4953.31 5002.38 -2466.65 0.57  4   0.967
```

Accept H0, so the discriminant parameters are all NOT SIGNIFICANTLY DIFFERENT. m2 is the reference, the best model.

So we check if the margins are better:

```r
margins(m2)
```

```
##
## Call:
## rasch(data = LSAT)
##
## Fit on the Two-Way Margins
##
## Response: (0,0)
##   Item i Item j Obs   Exp (O-E)^2/E
## 1      1      3  47 42.47      0.48
## 2      1      5  12 14.55      0.45
## 3      2      4  81 87.21      0.44
##
## Response: (1,0)
##   Item i Item j Obs    Exp (O-E)^2/E
## 1      3      5  63  58.48      0.35
## 2      2      4 156 149.79      0.26
## 3      4      5  85  88.43      0.13
##
## Response: (0,1)
##   Item i Item j Obs    Exp (O-E)^2/E
## 1      1      3  29  33.53      0.61
## 2      2      4 210 203.79      0.19
## 3      1      5  64  61.45      0.11
##
## Response: (1,1)
##   Item i Item j Obs    Exp (O-E)^2/E
## 1      2      4 553 559.21      0.07
## 2      3      5 490 494.53      0.04
## 3      1      3 524 519.47      0.04
```

```r
margins(m2, type="three-way",nprint=2)
```

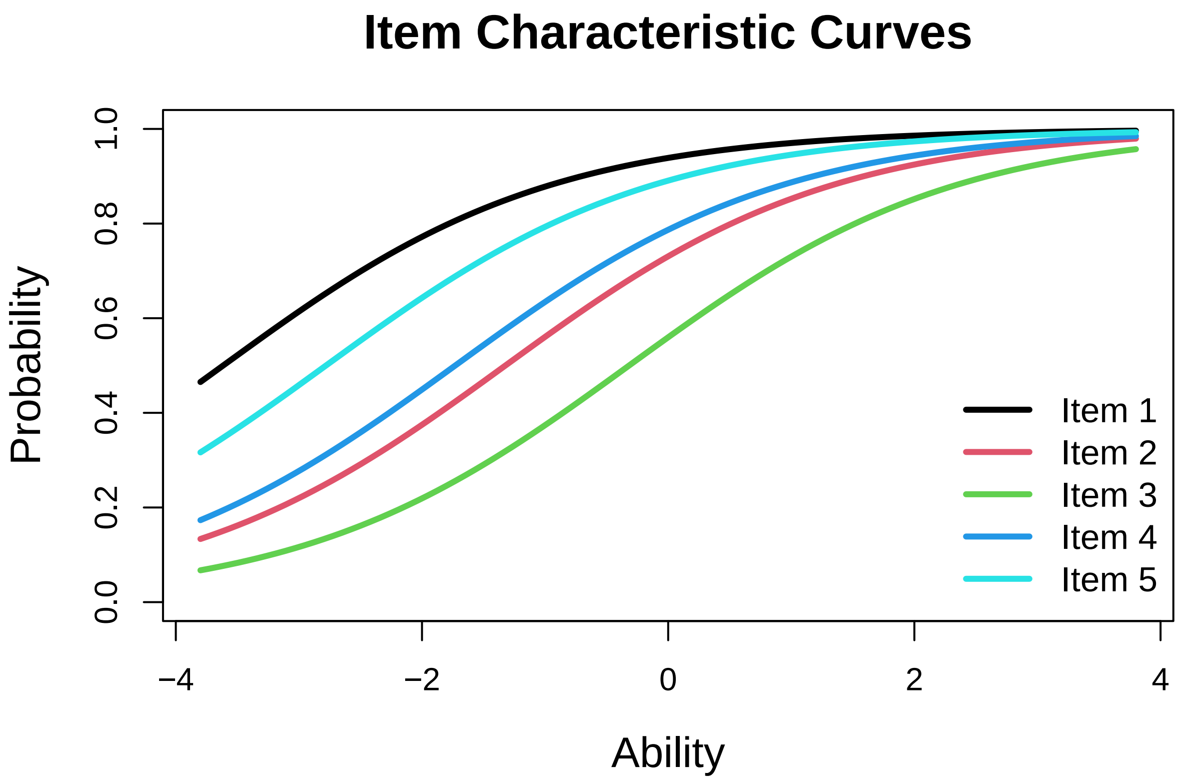
```
##
## Call:
```

```
## rasch(data = LSAT)
##
## Fit on the Three-Way Margins
##
## Response: (0,0,0)
##   Item i Item j Item k Obs   Exp (O-E)^2/E
## 1      1      3      5   6  9.40      1.23
## 2      3      4      5  30 25.85      0.67
##
## Response: (1,0,0)
##   Item i Item j Item k Obs   Exp (O-E)^2/E
## 1      2      4      5  28 22.75      1.21
## 2      2      3      4  81 74.44      0.58
##
## Response: (0,1,0)
##   Item i Item j Item k Obs  Exp (O-E)^2/E
## 1      1      2      5   3 7.58      2.76
## 2      1      3      4   5 9.21      1.92
##
## Response: (1,1,0)
##   Item i Item j Item k Obs   Exp (O-E)^2/E
## 1      2      4      5  51 57.49      0.73
## 2      3      4      5  48 42.75      0.64
##
## Response: (0,0,1)
##   Item i Item j Item k Obs    Exp (O-E)^2/E
## 1      1      3      5  41  33.07      1.90
## 2      2      3      4 108 101.28      0.45
##
## Response: (1,0,1)
##   Item i Item j Item k Obs    Exp (O-E)^2/E
## 1      2      3      4 210 218.91      0.36
## 2      1      2      4 190 185.56      0.11
##
## Response: (0,1,1)
##   Item i Item j Item k Obs   Exp (O-E)^2/E
## 1      1      3      5  23 28.38      1.02
## 2      1      4      5  46 42.51      0.29
##
## Response: (1,1,1)
##   Item i Item j Item k Obs    Exp (O-E)^2/E
## 1      1      2      4 520 526.36      0.08
## 2      1      2      3 398 393.30      0.06
```

They are good. Everything is perfect.

We can plot the item characteristic curves

```r
plot(m2, legend = TRUE, cx = "bottomright", lwd = 3, cex.main = 1.5,
     cex.lab = 1.3, cex =1.1)
```



For m2 the curves don't cross (they are all parallel) because they have all the same discriminant parameter. In terms of probability of correct response for the median individual, item 1 is the easiest one because the probability of getting a positive response is the highest. Item 3 is the most difficult. As we already found out.

The last thing that we need to do is to compute the **FACTOR SCORES**.

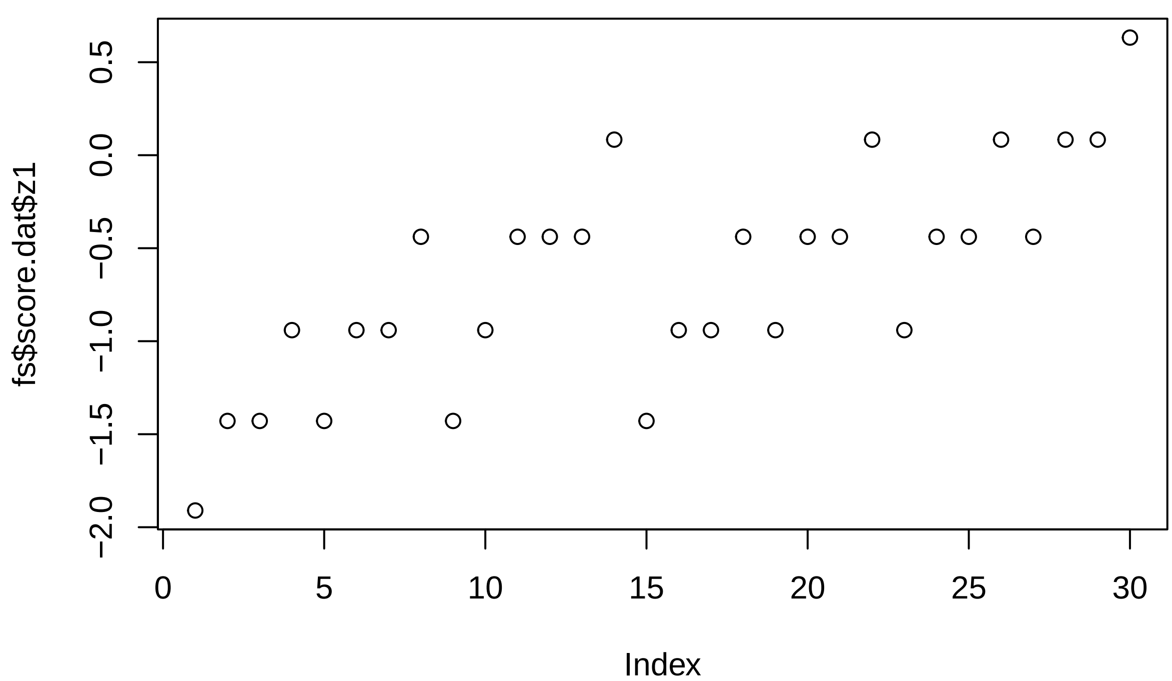
We can use the expected posterior:

```
fs<-factor.scores(m2,method="EAP")
fs
```

```
##
## Call:
## rasch(data = LSAT)
##
## Scoring Method: Expected A Posteriori
##
## Factor-Scores for observed response patterns:
##    Item 1 Item 2 Item 3 Item 4 Item 5 Obs   Exp     z1 se.z1
## 1       0      0      0      0      0   3 2.364 -1.910 0.797
## 2       0      0      0      0      1   6 5.468 -1.429 0.800
## 3       0      0      0      1      0   2 2.474 -1.429 0.800
## 4       0      0      0      1      1  11 8.249 -0.941 0.809
## 5       0      0      1      0      0   1 0.852 -1.429 0.800
## 6       0      0      1      0      1   1 2.839 -0.941 0.809
## 7       0      0      1      1      0   3 1.285 -0.941 0.809
## 8       0      0      1      1      1   4 6.222 -0.439 0.823
```

```
## 9        0       1       0       0       0   1    1.819 -1.429 0.800
## 10       0       1       0       0       1   8    6.063 -0.941 0.809
## 11       0       1       0       1       1  16   13.288 -0.439 0.823
## 12       0       1       1       0       1   3    4.574 -0.439 0.823
## 13       0       1       1       1       0   2    2.070 -0.439 0.823
## 14       0       1       1       1       1  15   14.749  0.084 0.841
## 15       1       0       0       0       0  10   10.273 -1.429 0.800
## 16       1       0       0       0       1  29   34.249 -0.941 0.809
## 17       1       0       0       1       0  14   15.498 -0.941 0.809
## 18       1       0       0       1       1  81   75.060 -0.439 0.823
## 19       1       0       1       0       0   3    5.334 -0.941 0.809
## 20       1       0       1       0       1  28   25.834 -0.439 0.823
## 21       1       0       1       1       0  15   11.690 -0.439 0.823
## 22       1       0       1       1       1  80   83.310  0.084 0.841
## 23       1       1       0       0       0  16   11.391 -0.941 0.809
## 24       1       1       0       0       1  56   55.171 -0.439 0.823
## 25       1       1       0       1       0  21   24.965 -0.439 0.823
## 26       1       1       0       1       1 173  177.918  0.084 0.841
## 27       1       1       1       0       0  11    8.592 -0.439 0.823
## 28       1       1       1       0       1  61   61.235  0.084 0.841
## 29       1       1       1       1       0  28   27.709  0.084 0.841
## 30       1       1       1       1       1 298  295.767  0.632 0.864
```

```
plot(fs$score.dat$z1)
```



There are some expected frequencies lower than 5 for some response pattern so there are some sparse data – the model in this case is not affected because they are few.

The lowest score is associated to the response pattern 0000 while the highest is associated to the response pattern 1111.

The sufficiency principle holds so we can compute the components and the total scores.

```
resp.pattern<-fs$score.dat[,1:5]
total.score<-apply(resp.pattern,1,sum)
total.score
```

```
##  [1] 0 1 1 2 1 2 2 3 1 2 3 3 3 4 1 2 2 3 2 3 3 4 2 3 3 4 3 4 4 5
```

```
round(fs$score.dat[order(total.score),],3) # Score according to the conditional expectation
```

```
##    Item 1 Item 2 Item 3 Item 4 Item 5 Obs     Exp     z1 se.z1
## 1       0      0      0      0      0   3   2.364 -1.910 0.797
## 2       0      0      0      0      1   6   5.468 -1.429 0.800
## 3       0      0      0      1      0   2   2.474 -1.429 0.800
## 5       0      0      1      0      0   1   0.852 -1.429 0.800
## 9       0      1      0      0      0   1   1.819 -1.429 0.800
## 15      1      0      0      0      0  10  10.273 -1.429 0.800
## 4       0      0      0      1      1  11   8.249 -0.941 0.809
## 6       0      0      1      0      1   1   2.839 -0.941 0.809
## 7       0      0      1      1      0   3   1.285 -0.941 0.809
## 10      0      1      0      0      1   8   6.063 -0.941 0.809
## 16      1      0      0      0      1  29  34.249 -0.941 0.809
## 17      1      0      0      1      0  14  15.498 -0.941 0.809
## 19      1      0      1      0      0   3   5.334 -0.941 0.809
## 23      1      1      0      0      0  16  11.391 -0.941 0.809
## 8       0      0      1      1      1   4   6.222 -0.439 0.823
## 11      0      1      0      1      1  16  13.288 -0.439 0.823
## 12      0      1      1      0      1   3   4.574 -0.439 0.823
## 13      0      1      1      1      0   2   2.070 -0.439 0.823
## 18      1      0      0      1      1  81  75.060 -0.439 0.823
## 20      1      0      1      0      1  28  25.834 -0.439 0.823
## 21      1      0      1      1      0  15  11.690 -0.439 0.823
## 24      1      1      0      0      1  56  55.171 -0.439 0.823
## 25      1      1      0      1      0  21  24.965 -0.439 0.823
## 27      1      1      1      0      0  11   8.592 -0.439 0.823
## 14      0      1      1      1      1  15  14.749  0.084 0.841
## 22      1      0      1      1      1  80  83.310  0.084 0.841
## 26      1      1      0      1      1 173 177.918  0.084 0.841
## 28      1      1      1      0      1  61  61.235  0.084 0.841
## 29      1      1      1      1      0  28  27.709  0.084 0.841
## 30      1      1      1      1      1 298 295.767  0.632 0.864
```

```
#ordered according to the total score.
```

Factor score of the two response pattern that didn't occur in the data set:

```
factor.scores(m2,resp.pattern=rbind(c(0,1,1,0,0),c(0,1,0,1,0)))
```

```
##
```

```
## Call:
## rasch(data = LSAT)
##
## Scoring Method: Empirical Bayes
##
## Factor-Scores for specified response patterns:
##   Item 1 Item 2 Item 3 Item 4 Item 5 Obs   Exp     z1 se.z1
## 1      0      1      1      0      0   0 0.944 -0.959 0.801
## 2      0      1      0      1      0   0 2.744 -0.959 0.801
```

Scores are very small because of course there are zero observations for these patterns.