Detection of DIF with difNLR package

Adéla Drabinová, Patrícia Martinková

IMPS, Columbia University, New York City, July 11, 2018

Institute of Computer Science of the Czech Academy of Sciences
Faculty of Mathematics and Physics, Charles University
Faculty of Education, Charles University

Table of contents

Motivation

DIF detection methods

Classifications of DIF detection methods with respect to

matching criterion

- based on latent score
  (IRT models)
- based on total score
  (Mantel-Haenszel test, logistic regression)

DIF detection methods

Classifications of DIF detection methods with respect to

matching criterion

- based on latent score
  (IRT models)
- based on total score
  (Mantel-Haenszel test, logistic regression)

possibility of guessing/inattention

- accounting for
  (3PL and 4PL IRT models)
- not accounting for
  (Mantel-Haenszel test, logistic regression)

DIF detection methods

Classifications of DIF detection methods with respect to

matching criterion

- based on latent score
  (IRT models)
- based on total score
  (Mantel-Haenszel test, logistic regression)

possibility of guessing/inattention

- accounting for
  (3PL and 4PL IRT models)
- not accounting for
  (Mantel-Haenszel test, logistic regression)

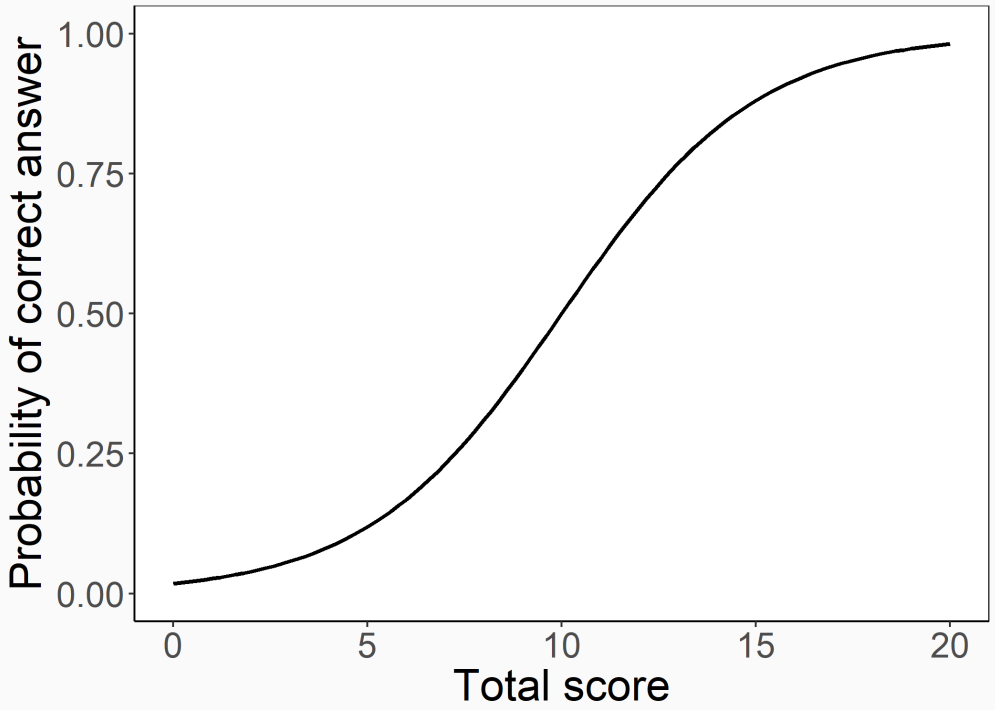
Need for simple tool with low rate of convergence issues and accounting for possibility of guessing/inattention

Generalized logistic models for DIF detection

Generalized logistic regression for DIF detection

$$P(Y_{pi} = 1 | X_p, G_p) = \frac{e^{\alpha_i \quad (X_p - \beta_i \quad)}}{1 + e^{\alpha_i \quad (X_p - \beta_i \quad)}}$$

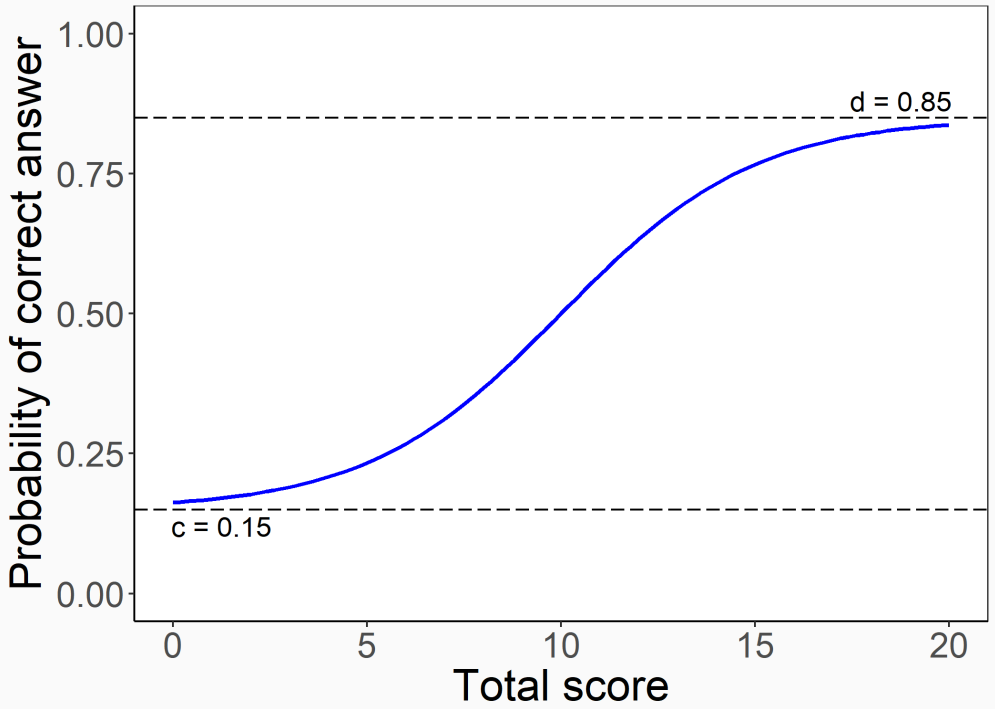
= probability of correct answer by person p on item i
X_p total score, G_p group membership



Generalized logistic regression for DIF detection

$$P(Y_{pi} = 1|X_p, G_p) = c_i \quad + (d_i \quad - c_i \quad) \frac{e^{\alpha_i \quad (X_p - \beta_i \quad)}}{1 + e^{\alpha_i \quad (X_p - \beta_i \quad)}}$$

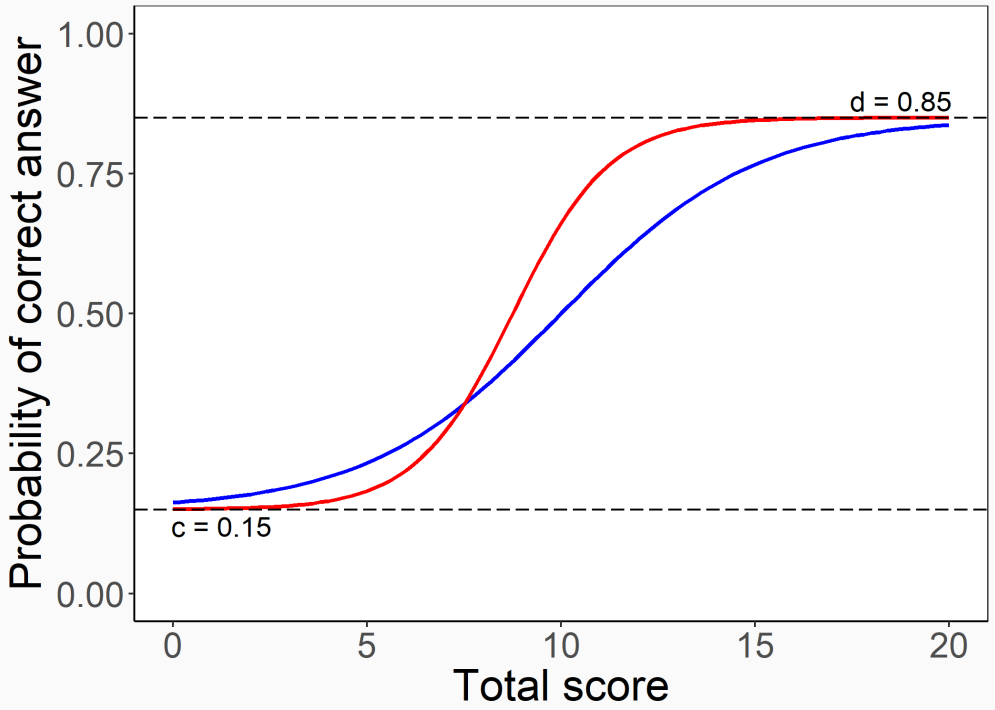
= probability of correct answer by person p on item i
X_p total score, G_p group membership



Generalized logistic regression for DIF detection

$$P(Y_{pi} = 1|X_p, G_p) = c_i + (d_i - c_i) \frac{e^{\alpha_{iG_p}(X_p - \beta_{iG_p})}}{1 + e^{\alpha_{iG_p}(X_p - \beta_{iG_p})}}$$

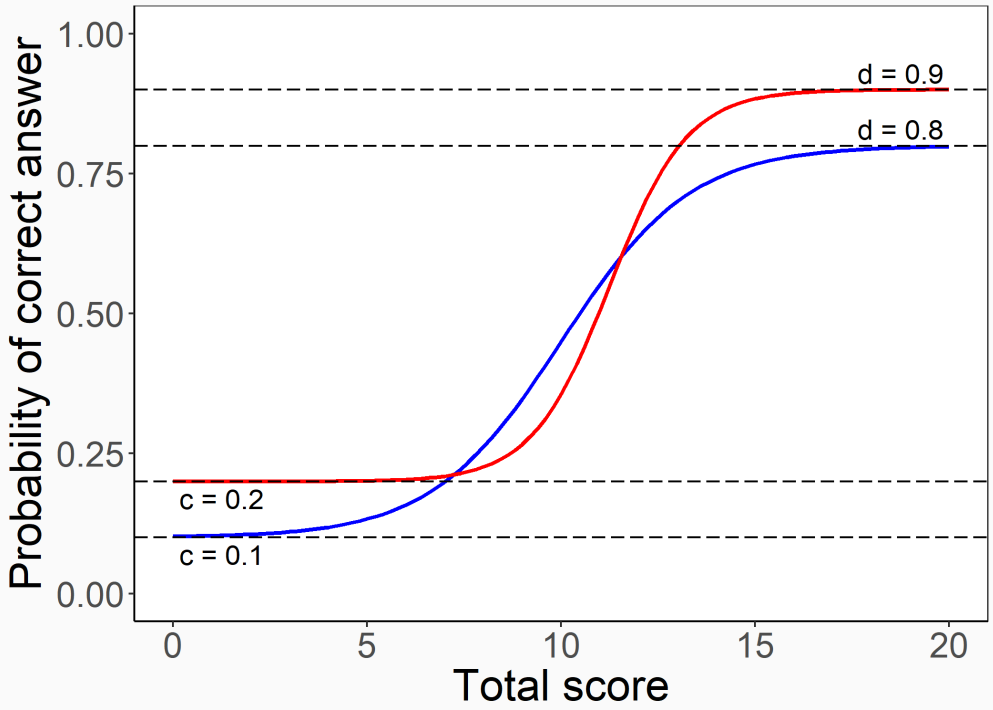
= probability of correct answer by person p on item i
X_p total score, G_p group membership



Generalized logistic regression for DIF detection

$$P(Y_{pi} = 1|X_p, G_p) = c_{iG_p} + (d_{iG_p} - c_{iG_p}) \frac{e^{\alpha_{iG_p}(X_p - \beta_{iG_p})}}{1 + e^{\alpha_{iG_p}(X_p - \beta_{iG_p})}}$$

= probability of correct answer by person p on item i
X_p total score, G_p group membership



Generalized logistic model for DDF detection

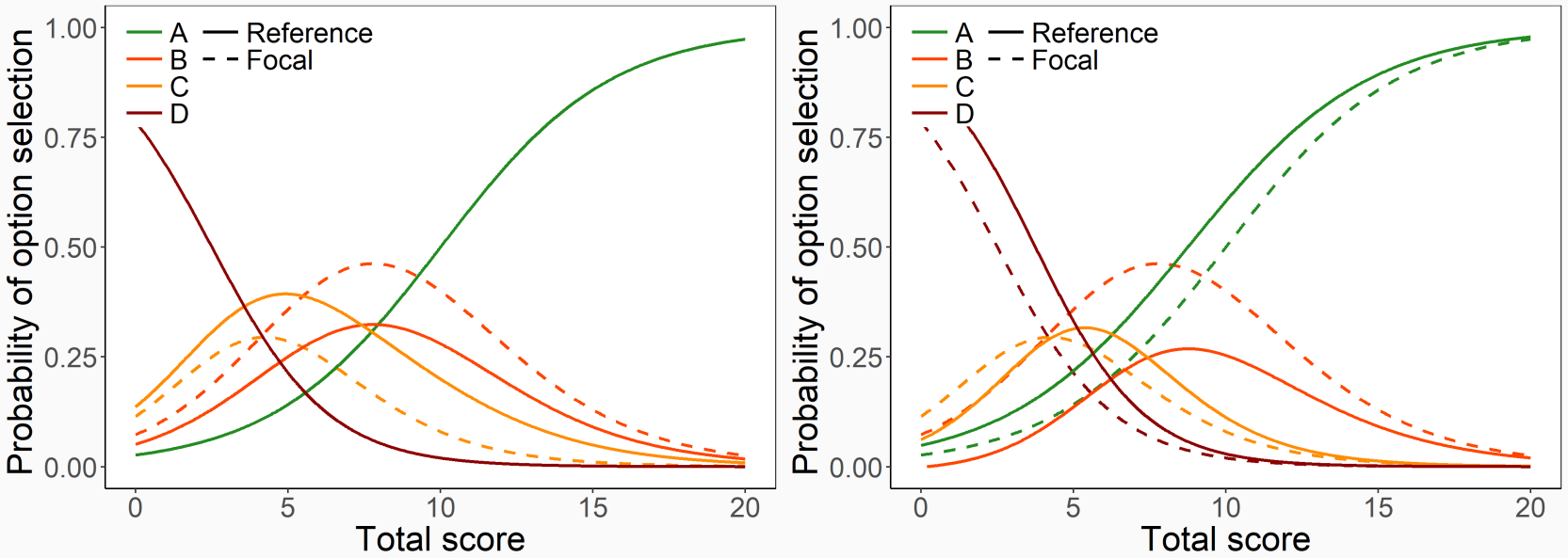
Multinomial regression

$$\mathrm{P}(Y_{pi} = k|X_p, G_p) = \frac{e^{\alpha_{iG_pk}(X_p - \beta_{iG_pk})}}{1 + \sum_{l=1}^{K-1} e^{\alpha_{iG_pl}(X_p - \beta_{iG_pl})}} \qquad \text{(distractor)}$$

$$\mathrm{P}(Y_{pi} = K|X_p, G_p) = \frac{1}{1 + \sum_{l=1}^{K-1} e^{\alpha_{iG_pl}(X_p - \beta_{iG_pl})}} \qquad \text{(correct answer)}$$

= probability of option selection by person p on item i

X_p total score, G_p group membership



difNLR package

difNLR package

difNLR: DIF and DDF detection by non-linear regression models

- R package
- Version 1.2.2 on ▸ CRAN

```
install.packages("difNLR")
```

- The newest development version on ▸ GitHub

```
devtools::install_github("drabinova/difNLR")
```

- Run it with

```
library("difNLR")
```

- Try some features online

https://shiny.cs.cas.cz/ShinyItemAnalysis/

DIF detection - `difNLR()` function

What needs to be specified:

```
difNLR(Data, group, focal.name, model, ...)
```

Options

- 11 predefined models

```
"2PL", "3PLc", "4PL", ...
```

- Arguments `constraints` and `type` to specify further models

```
model = "4PL", constraints = "a", type = "cd"
```

- Models can be item specific

DIF detection - `difNLR()` function

Example (Drabinová & Martinková, 2018)

Generated data with 4 DIF items (5, 8, 11 and 15)

```
dim(df)
[1] 1000   16

head(df)
  I1 I2 I3 I4 I5 I6 I7 I8 I9 I10 I11 I12 I13 I14 I15 group
1  0  1  1  1  1  1  0  0  1   1   1   1   1   0   1     0
2  0  1  1  0  1  1  0  0  1   1   1   0   0   1   0     0
3  0  1  0  0  1  1  0  0  1   0   1   0   1   0   1     0
4  1  1  1  0  1  1  0  0  1   1   1   1   0   0   1     0
5  1  1  0  1  1  1  0  0  1   1   1   1   1   0   1     0
6  0  1  0  0  1  0  0  0  1   0   0   0   0   0   1     0

DataDIF <- df[, 1:15]
groupDIF <- df[, 16]
```

Drabinová, A. & Martinková, P. (2018) difNLR: Generalized logistic regression models for DIF and DDF detection. Submitted.

```
(fit1 <- difNLR(DataDIF, groupDIF,
                focal.name = 1,
                model = "4PL",
                type = "all"))
```

```
(fit1 <- difNLR(DataDIF, groupDIF,
                focal.name = 1,
                model = "4PL",
                type = "all"))
```

```
Detection of all types of differential item functioning
using generalized logistic regression model

Generalized logistic regression
likelihood ratio chi-square statistics based on 4PL model

Parameters were estimated with non-linear least squares

Item purification was not applied
No p-value adjustment for multiple comparisons

    Chisq-value P-value
I1   6.2044      0.1844
I2   0.2802      0.9911
I3   2.7038      0.6086
I4   5.8271      0.2124
I5  48.0052      0.0000 ***
I6   7.2060      0.1254
I7   3.2390      0.5187
I8  16.8991      0.0020 **
I9   2.1595      0.7064
I10  4.6866      0.3210
I11 69.5328      0.0000 ***
I12  8.1931      0.0848 .
I13  2.5850      0.6295
I14  2.9478      0.5666
I15 20.6589      0.0004 ***

Sign. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Detection thresholds: 9.4877 (significance level: 0.05)

Items detected as DIF items:
 I5
 I8
 I11
 I15
```

```r
(fit1 <- difNLR(DataDIF, groupDIF,
                focal.name = 1,
                model = "4PL",
                type = "all"))

# coefficients
round(coef(fit1), 3)
```

```r
(fit1 <- difNLR(DataDIF, groupDIF,
                focal.name = 1,
                model = "4PL",
                type = "all"))

# coefficients
round(coef(fit1), 3)
```

```
        a      b     c     d   aDif   bDif  cDif   dDif
I1  1.484  1.294 0.049 1.000  0.000  0.000 0.000  0.000
I2  1.176  0.153 0.000 1.000  0.000  0.000 0.000  0.000
I3  1.281  1.766 0.001 1.000  0.000  0.000 0.000  0.000
I4  1.450  0.421 0.000 1.000  0.000  0.000 0.000  0.000
I5  1.965 -1.147 0.000 0.868 -0.408  0.769 0.023 -0.006
I6  1.458 -0.527 0.000 0.954  0.000  0.000 0.000  0.000
I7  0.888  1.392 0.000 1.000  0.000  0.000 0.000  0.000
I8  1.162  1.407 0.000 0.866 -0.117  0.974 0.007  0.134
I9  1.482 -1.337 0.000 0.928  0.000  0.000 0.000  0.000
I10 1.375 -0.570 0.007 0.967  0.000  0.000 0.000  0.000
I11 1.071 -1.027 0.000 0.969  1.173 -0.499 0.000  0.011
I12 1.051  1.560 0.080 1.000  0.000  0.000 0.000  0.000
I13 1.009  1.348 0.084 1.000  0.000  0.000 0.000  0.000
I14 1.093  1.659 0.141 1.000  0.000  0.000 0.000  0.000
I15 0.875 -0.565 0.000 0.945  0.205  0.348 0.000 -0.142
```

```r
(fit1 <- difNLR(DataDIF, groupDIF,
                focal.name = 1,
                model = "4PL",
                type = "all"))

# coefficients
round(coef(fit1), 3)


# fit measures
AIC(fit1, item = 5)
BIC(fit1, item = 5)
logLik(fit1, item = 5)
```

```r
(fit1 <- difNLR(DataDIF, groupDIF,
                focal.name = 1,
                model = "4PL",
                type = "all"))

# coefficients
round(coef(fit1), 3)


# fit measures
AIC(fit1, item = 5)
BIC(fit1, item = 5)
logLik(fit1, item = 5)
```

```
[1] 1166.81

[1] 1206.072

'log Lik.' -575.4048 (df=8)
```

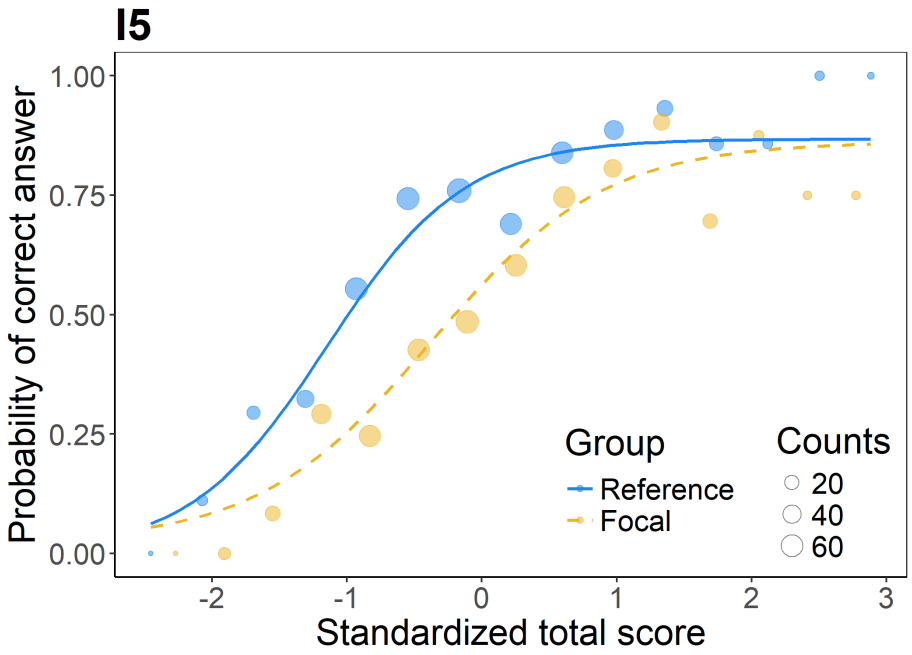
```r
(fit1 <- difNLR(DataDIF, groupDIF,
                focal.name = 1,
                model = "4PL",
                type = "all"))

# coefficients
round(coef(fit1), 3)


# fit measures
AIC(fit1, item = 5)
BIC(fit1, item = 5)
logLik(fit1, item = 5)


# prediction
predict(fit1, item = 5,
        match = 0, group = 0)
predict(fit1, item = 5,
        match = 0, group = 1)
```

```r
(fit1 <- difNLR(DataDIF, groupDIF,
                focal.name = 1,
                model = "4PL",
                type = "all"))

# coefficients
round(coef(fit1), 3)


# fit measures
AIC(fit1, item = 5)
BIC(fit1, item = 5)
logLik(fit1, item = 5)


# prediction
predict(fit1, item = 5,
        match = 0, group = 0)
predict(fit1, item = 5,
        match = 0, group = 1)
```

```
       I5
0.7851739

       I5
0.5624883
```

```r
(fit1 <- difNLR(DataDIF, groupDIF,
                focal.name = 1,
                model = "4PL",
                type = "all"))

# coefficients
round(coef(fit1), 3)


# fit measures
AIC(fit1, item = 5)
BIC(fit1, item = 5)
logLik(fit1, item = 5)


# prediction
predict(fit1, item = 5,
        match = 0, group = 0)
predict(fit1, item = 5,
        match = 0, group = 1)


# plotting ICC
plot(fit1, item = 5)
```

```r
(fit1 <- difNLR(DataDIF, groupDIF,
                focal.name = 1,
                model = "4PL",
                type = "all"))

# coefficients
round(coef(fit1), 3)


# fit measures
AIC(fit1, item = 5)
BIC(fit1, item = 5)
logLik(fit1, item = 5)


# prediction
predict(fit1, item = 5,
        match = 0, group = 0)
predict(fit1, item = 5,
        match = 0, group = 1)


# plotting ICC
plot(fit1, item = 5)
```



DDF detection - ddfMLR( ) function

What needs to be specified

```
ddfMLR(Data, key, group, focal.name, ...)
```

Example (Drabinová & Martinková, 2018)

Generated data with 2 DDF items (1 and 6)

```
summary(sapply(DataDDF, as.factor))
I1      I2      I3      I4      I5
0:363   0:361   0:105   0:641   0:104
1:434   1:437   1:345   1:260   1:299
2:203   2:202   2:550   2: 99   2:597

I6      I7      I8      I9      I10     group
0:454   0:354   0:214   0:556   0:315   0:500
1:323   1:377   1:224   1:294   1:334   1:500
2:172   2:212   2:291   2:115   2:241
3: 51   3: 57   3:271   3: 35   3:110
```

Drabinová, A. & Martinková, P. (2018) difNLR: Generalized logistic regression models for DIF and DDF detection. Submitted.

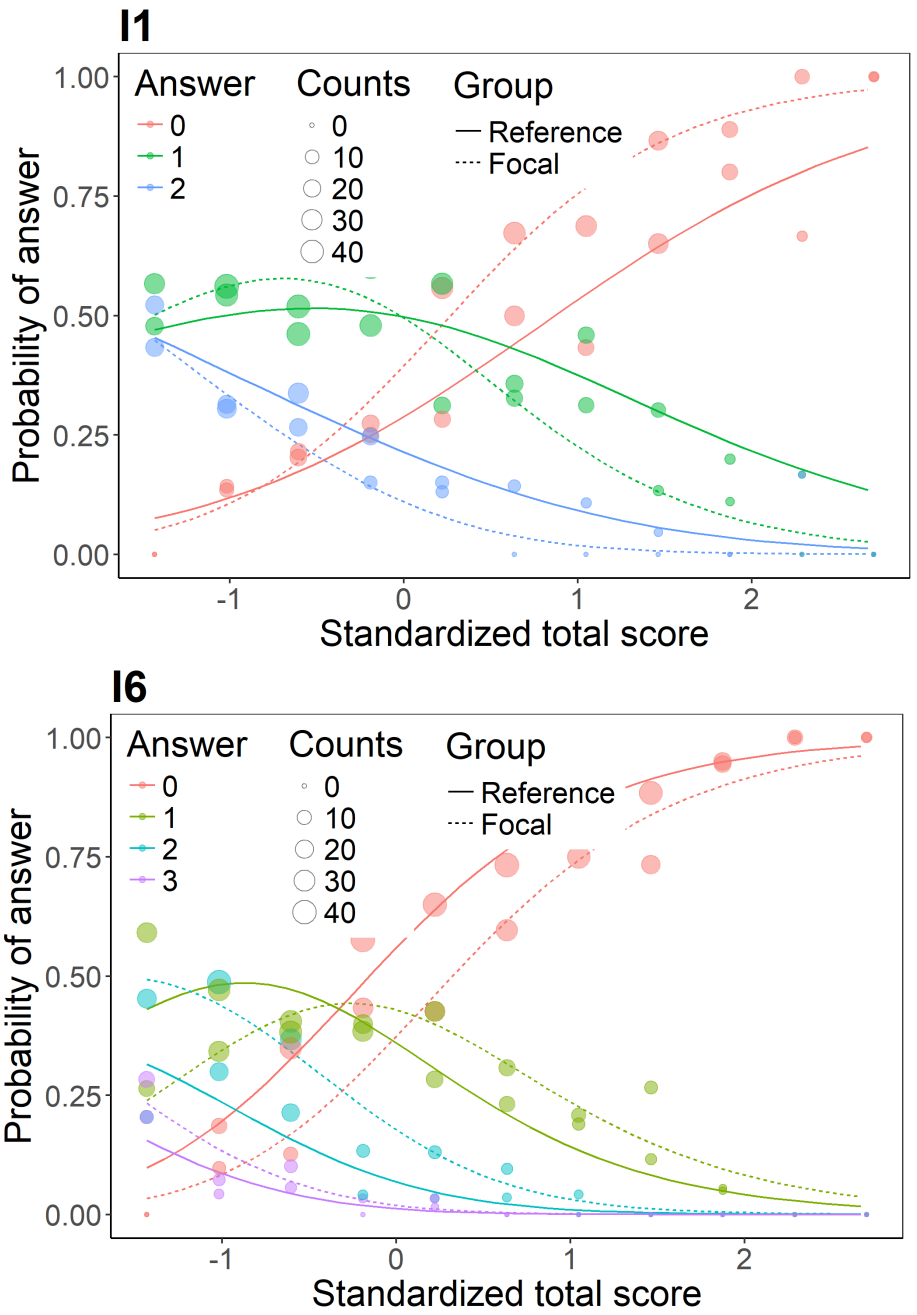
```
(fit2 <- ddfMLR(DataDDF,
                group = "group",
                focal.name = 1,
                key = rep(0, 10)))
```

```r
(fit2 <- ddfMLR(DataDDF,
                group = "group",
                focal.name = 1,
                key = rep(0, 10)))
```

```
Detection of both types of Differential Distractor
Functioning using Multinomial Log-linear Regression model

Likelihood-ratio chi-square statistics

Item purification was not applied
No p-value adjustment for multiple comparisons

    Chisq-value P-value
I1  29.5508      0.0000 ***
I2   1.1136      0.8921
I3   1.0362      0.9043
I4   4.1345      0.3881
I5   7.4608      0.1134
I6  47.0701      0.0000 ***
I7   1.3285      0.9701
I8   2.3629      0.8835
I9  10.4472      0.1070
I10  3.5602      0.7359

Sign. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Items detected as DDF items:
 I1
 I6
```

```r
(fit2 <- ddfMLR(DataDDF,
                group = "group",
                focal.name = 1,
                key = rep(0, 10)))


# plotting ICC
plot(fit2, item = fit2$DDFitems)
```

```r
(fit2 <- ddfMLR(DataDDF,
                group = "group",
                focal.name = 1,
                key = rep(0, 10)))


# plotting ICC
plot(fit2, item = fit2$DDFitems)
```



Features and Troubleshooting

Initial values and estimation methods

`difNLR()` function offers

- Two **estimation methods** (NLS, MLE)

```
difNLR(..., method = "nls")
difNLR(..., method = "likelihood")
```

- Automatic initial values

In case of convergence issues

- Recalculation of initial values based on bootstraped samples

```
difNLR(..., initboot = TRUE, nrBo = 20)

Starting values were calculated based on bootstraped samples.
```

Multiple comparison corrections

Both functions offer **corrections for multiple comparison**

```
difNLR(..., p.adjust.method = "none")
...
No p-value adjustment for multiple comparisons

     Chisq-value P-value
I1   6.2044      0.1844
I2   0.2802      0.9911
I3   2.7038      0.6086
I4   5.8271      0.2124
I5  48.0052      0.0000 ***
I6   7.2060      0.1254
I7   3.2390      0.5187
I8  16.8991      0.0020 **
I9   2.1595      0.7064
I10  4.6866      0.3210
I11 69.5328      0.0000 ***
I12  8.1931      0.0848 .
I13  2.5850      0.6295
I14  2.9478      0.5666
I15 20.6589      0.0004 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Detection thresholds: 9.4877 (significance level: 0.05)

Items detected as DIF items:
 I5
 I8
 I11
 I15
```

Multiple comparison corrections

Both functions offer **corrections for multiple comparison**

```
difNLR(..., p.adjust.method = "holm")
...
Multiple comparisons made with Holm adjustment of p-values

     Chisq-value P-value Adj. P-value
I1    6.2044      0.1844  1.0000
I2    0.2802      0.9911  1.0000
I3    2.7038      0.6086  1.0000
I4    5.8271      0.2124  1.0000
I5   48.0052      0.0000  0.0000      ***
I6    7.2060      0.1254  1.0000
I7    3.2390      0.5187  1.0000
I8   16.8991      0.0020  0.0243      *
I9    2.1595      0.7064  1.0000
I10   4.6866      0.3210  1.0000
I11 69.5328      0.0000  0.0000      ***
I12   8.1931      0.0848  0.9323
I13   2.5850      0.6295  1.0000
I14   2.9478      0.5666  1.0000
I15 20.6589      0.0004  0.0048      **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Detection thresholds: 9.4877 (significance level: 0.05)

Items detected as DIF items:
 I5
 I8
 I11
 I15
```

Item purification

`difNLR()` and `ddfMLR()` functions offer **item purification**

```
fit <- difNLR(..., purify = TRUE, nrIter = 10)

fit$difPur
       I1 I2 I3 I4 I5 I6 I7 I8 I9 I10 I11 I12
Step0   0  0  0  0  1  0  0  1  0   0   1   0
Step1   1  0  0  0  1  0  0  1  0   0   1   0
Step2   0  0  0  0  1  0  0  1  0   0   1   0
Step3   1  0  0  0  1  0  0  1  0   0   1   0
Step4   0  0  0  0  1  0  0  1  0   0   1   0
Step5   1  0  0  0  1  0  0  1  0   0   1   0
Step6   0  0  0  0  1  0  0  1  0   0   1   0
Step7   1  0  0  0  1  0  0  1  0   0   1   0
Step8   0  0  0  0  1  0  0  1  0   0   1   0
Step9   1  0  0  0  1  0  0  1  0   0   1   0
Step10  0  0  0  0  1  0  0  1  0   0   1   0
```

Learn more

ShinyItemAnalysis

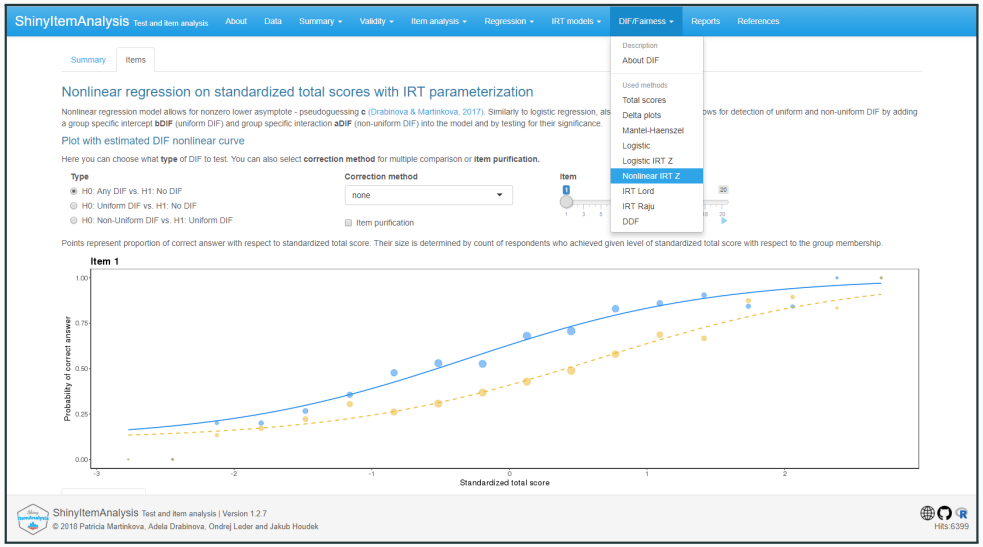
`difNLR` package exploited by `ShinyItemAnalysis`

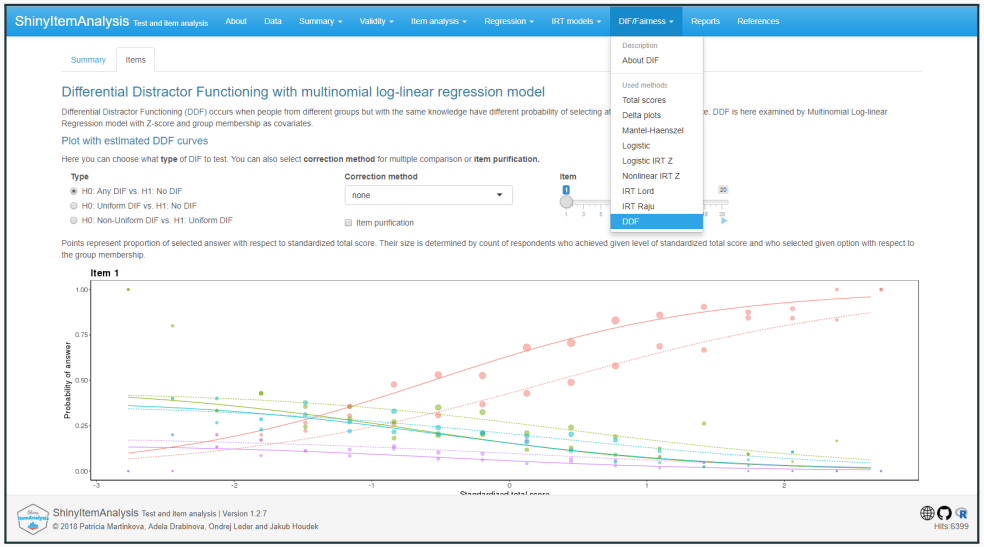
- R package and **`shiny`** application
- Available on CRAN and also online at

https://shiny.cs.cas.cz/ShinyItemAnalysis/

and

https://cemp.shinyapps.io/ShinyItemAnalysis/





Summary

`difNLR`

- User-friendly R package for DIF/DDF detection
- Provides wide range of generalized logistic models
- Good properties in terms of power and rejection rate (Drabinová & Martinková, 2017)
- Offers features common in DIF/DDF detection
- Accessible via shiny application `ShinyItemAnalysis`

Drabinová, A. & Martinková, P. (2017) Detection of Differential Item Functioning with Nonlinear Regression: A Non-IRT Approach Accounting for Guessing. *Journal of Educational Measurement, 54*(4), 498-517.

drabinova@cs.cas.cz
www.cs.cas.cz/drabinova/

This research was supported by Czech Science Foundation grant GJ15-15856Y