Package 'vampyr'

October 12, 2022

Type Package

Title Factor Analysis Controlling the Effects of Response Bias

Version 1.1.1

Date 2021-06-17

Author David Navarro-Gonzalez, Andreu Vigil-Colet, Pere J. Ferrando, Urbano Lorenzo-Seva, Jorge N. Tendeiro

Maintainer David Navarro-Gonzalez <david.navarro@urv.cat>

Description Vampirize the response biases from a dataset! Performs factor analysis controlling the effects of social desirability and acquiescence using the method described in Ferrando, Lorenzo-Seva & Chico (2009) <doi:10.1080/10705510902751374>.

Depends R (>= 2.10)

Imports stats, psych, EFA.MRFA, GPArotation, PCovR, moments, corpcor, lavaan, fungible, semPlot

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

NeedsCompilation no

Repository CRAN

Date/Publication 2021-06-17 08:30:02 UTC

R topics documented:

`vampyr-package` *Factor Analysis Controlling the Effects of Response Bias*

Description

Package for performing Factor Analysis (FA) implementing the method developed by Ferrando, Lorenzo-Seva & Chico (2009). It allows the user to control the effects of two response bias: Social Desirability (SD) and Acquiescence (AC) by extracting the variance due to these factors before extracting the content ones. In other words, vampirize the response biases by substracting their impact on the factor structure.

Value

`ControlResponseBias`
: Performs FA controlling response biases.

Author(s)

David Navarro-Gonzalez

Andreu Vigil-Colet

Pere Joan Ferrando

Urbano Lorenzo-Seva

Jorge Tendeiro

References

Ferrando, P. J., Lorenzo-Seva, U., & Chico, E. (2009). A General factor-Analytic procedure for assessing response bias in questionnaire measures. Structural Equation Modeling, 16, 364-381. doi: 10.1080/10705510902751374

`ControlResponseBias` *Performs Factor Analysis controlling response bias*

Description

Performs Factor Analysis (FA) implementing the method developed by Ferrando, Lorenzo-Seva & Chico (2009). It allows the user to control the effects of two response bias: Social Desirability (SD) and Acquiescence (AC) by extracting the variance due to these factors before extracting the content ones.

Usage

```
ControlResponseBias(x, content_factors, SD_items, unbalanced_items, contSD = FALSE,
   contAC = FALSE, corr = "Pearson", rotat = "promin", target, factor_scores = FALSE,
    PA = FALSE, path = FALSE, display = TRUE)
```

Arguments

x	Raw sample scores or a covariance/correlation matrix directly. The SD items (if applicable) have to be included in the same dataset as the content ones.
content_factors	The number of content factors to be retained. At least, each factor has to be defined by 3 items, so the maximum number of content factors will be the number of items / 3.
SD_items	When controlling Social Desirability, a vector containing the Social Desirability markers has to be provided. At least 4 SD markers are necessary to perform the SD control procedure. Providing a valid input argument turns automatically contSD to TRUE.
unbalanced_items	When controlling AC, if the content items are not fully balanced between positive worded items and negative worded items, the user has to determine which items will not be used in the calculation of the centroid. For more information, check the details section. Providing a valid input argument turns automatically contAC to TRUE
contSD	Logical variable determining if the method for controlling Social Desirability will be used. If SD_items are provided, it automatically switch to TRUE.
contAC	Logical variable determining if the method for controlling Acquiescence will be used. If unbalanced_items are provided, it automatically switch to TRUE. However, it is possible to control Acquiescence without providing unbalanced_items when the scale is fully balanced.
corr	Determines the type of matrices to be used in the factor analysis. "Pearson": Computes Pearson correlation matrices (linear model); "Polychoric": Computes Polychoric/Tetrachoric correlation matrices (graded model). If the matrix is not positive definite, the smoothing procedure proposed by Bentler and Yuan (2011) is used.
rotat	Determines if the factor loading matrix will be rotated, and if it dues, which method will be used. By default, "promin" rotation wil lbe used. The rotation methods available include all the available options for the rotation of the loading matrix included in the following packages: `stats` (R Core Team, 2019), `GPArotation` (Bernaards, Jennrich, 2005) and `PCovR` (Vervloet, Kiers, Van den Noortgate & Ceulemans, 2015). If the user is not interested in rotating the factors, provide the "none" option. The argument has to be provided with the same syntactic structure as in the original package.
target	The target for the procrustes factor analysis rotation (optional).
factor_scores	Logical variable determining if the factor scores will be computed. If TRUE, Fully-Informative Prior EAP scores (Ferrando & Lorenzo-Seva, 2016) will be computed for continuous and graded models. When using Polychoric matrix, the precision of the scores is computed, and stored in $precision_matrix in a three dimensional matrix.
PA	Logical variable determining if the Optimal Implementation of Parallel Analysis (Timmerman & Lorenzo-Seva, 2011) will be computed for estimating the number of factors to be retained.

`path` Logical variable determining if path diagram will be plotted (FALSE by default). It uses the function `semPaths` internally, and for preventing displaying issues, it is limited to 40 total items and 5 content factors. Since it was designed just for providing a visual representation of the factor loadings, the residuals and intercepts have been omited, and the loadings lower than .3 in absolute value are not being plotted (in AC the threshold is .2).

`display` Determines if the output will be displayed in the console, TRUE by default. If it is TRUE, the output is printed in console and if it is FALSE, the output is returned silently to the output variable, without printing anything in console. Alternatively, a character array can be provided for printing specific sections of the output, the options being the following: "Summary", "Descriptives", "Adequacy", "PA" (Parallel Analysis),"GOF" (Goodness of Fit indices) and "Loadings".

Details

The function is based on the method proposed by Ferrando, Lorenzo-Seva & Chico (2009), and it is adapted from the MATLAB code of Psychological Test Toolbox program (Navarro-Gonzalez, Vigil-Colet, Ferrando & Lorenzo-Seva, 2019).

The FA procedure is structurated in 3 steps:

1. Extract the Social Desirability factor using the SD markers provided by the user. This items should be pure measures of SD and are assumed to be uncorrelated with the content.

2. Extract the Acquiescence factor using a balanced set of content items, containing the same number of positively keyed and negatively keyed items. If the scale is only partially balanced, some items of the predominant type (positively keyed or negatibely keyed) have to be excluded from the balanced set. The recommended exclusion criteria is the following: exclude the items with lower loadings on AC, using the "unbalanced_itemd" argument. If there is no a priori reasoning, we recommend selecting some items at random to be used as unbalanced in a first exploratory analisis and then, look at the AC loadings and select the ones with lower loadings for a definitive second analisis.

3. Extract the content factors using the residual matrix obtained after substracting the variance due to Social Desirability and Acquiescence.

It is not mandatory to control both of the response biases. The functions allows the user to select which bias should be controlled between SD, AC or both.

The function implements all the available options for the rotation of the loading matrix included in the following packages:

- `stats-package` (R Core Team, 2019).

- `GPArotation-package` (Bernaards, Jennrich, 2005).

- `PCovR-package` (Vervloet, Kiers, Van den Noortgate & Ceulemans, 2015).

The selection of rotation methods include some of the most knowns ones, like:

-`promin` (Lorenzo-Seva, 1999), available in the PCovR package.

-`varimax` (Kaiser, 1958), available in the stats package.

- `Oblimin` (Jackson, 2005), available in the GPArotation package.

If the user is not interested in any of the rotation methods available or prefer rotating the loading matrix externally, please consider that the response bias factors (SD and/or AC) should not be rotated.

The robust goodness fit of indices are computed using `lavaan` (Rosseel, 2012).

Value

`loadings`	Obtained loading matrix.
`Phi`	Correlation between factors, only available when an oblique rotation has been selected.
`Factor_scores`	Obtained factor scores for each participant in the bias and content factors. This variable is returned only if the argument factor_scores has set to TRUE.
`precision_matrix`	Precision values for each factor score estimate, including the 90 % CI, the Posterior SD (PSD) and the reliability of the estimation. Only available for graded model.
`comunalities`	Obtained communalities for each item.
`ECV`	Explained common variance for each factor.
`reduced`	Reduced matrix obtained once extracting all the factors.
`produced`	Produced final matrix.
`RMSEA`	Root Mean Square Error of Approximation.
`Chi`	Robust Mean-Scaled Chi Square of the model.
`TLI`	Non-Normed Fit Index (NNFI; Tucker & Lewis).
`CFI`	Comparative Fit Index.
`GFI`	Goodness of Fit Index.
`RMSR`	Root Mean Square Residuals.
`kelley`	Kelley's criterion.
`PA_Real_Data`	A vector containing the percentage of explained variance by the real data for each factor (If Parallel Analysis is required by using the argument PA = TRUE).
`PA_Mean_Random`	A vector containing the percentage of explained variance by the mean of random data for each factor (If Parallel Analysis is required by using the argument PA = TRUE).
`PA_Percentile_Random`	A vector containing the percentage of explained variance by the percentile of distribution of random data for each factor (If Parallel Analysis is required by using the argument PA = TRUE).
`N_factors_mean`	The number of factors to be retained suggested comparing the real data with the mean of the random data (If Parallel Analysis is required by using the argument PA = TRUE).
`N_factors_percentiles`	The number of factors to be retained suggested comparing the real data with the percentile of distribution of the random data (If Parallel Analysis is required by using the argument PA = TRUE).

Author(s)

David Navarro-Gonzalez

Andreu Vigil-Colet

Pere Joan Ferrando

Urbano Lorenzo-Seva

Jorge Tendeiro

References

Bentler, P. M., & Yuan, K. H. (2011). Positive definiteness via off-diagonal scaling of a symmetric indefinite matrix. Psychometrika, 76(1), 119-123. doi: 10.1007/s1133601091913

Bernaards, C. A. and Jennrich, R. I. (2005) Gradient Projection Algorithms and Software for Arbitrary Rotation Criteria in Factor Analysis. Educational and Psychological Measurement, 65, 676-696. doi: 10.1177/0013164404272507

Ferrando, P. J., & Lorenzo-Seva, U. (2016). A note on improving EAP trait estimation in oblique factor-analytic and item response theory models. Psicologica, 37, 235-247.

Ferrando, P. J., Lorenzo-Seva, U., & Chico, E. (2009). A General factor-Analytic procedure for assessing response bias in questionnaire measures. Structural Equation Modeling, 16, 364-381. doi: 10.1080/10705510902751374

Jackson, J. E. (2005). Oblimin Rotation. Encyclopedia of Biostatistics. 6. doi: 10.1002/0470011815.b2a13060

Kaiser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. Psychometrika, 23, 187-200. doi: 10.1007/BF02289233

Lorenzo-Seva, U. (1999). Promin: A Method for Oblique Factor Rotation. Multivariate Behavioral Research, 34(3), 347-365. doi: 10.1207/S15327906MBR3403_3

Navarro-Gonzalez, D., Vigil-Colet, A., Ferrando, P.J., & Lorenzo-Seva, U. (2019). Psychological Test Toolbox: A New Tool to Compute Factor Analysis Controlling Response Bias. Journal of Statistical Software, 91(6). doi: 10.18637/jss.v091.i06

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48(2).doi: 10.18637/jss.v048.i02

Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality Assessment of Ordered Polytomous Items with Parallel Analysis. Psychological Methods, 16, 209-220. doi: 10.1037/a0023353

Vervloet, M., Kiers, H.A.L., Van den Noortgate, W., & Ceulemans, E. (2013). PCovR: An R Package for Principal Covariates Regression. Journal of Statistical Software, 65(8). doi: 10.18637/jss.v065.i08

Examples

```
## This is an example using the vampyr_example dataset, which contains 4 SD markers (the first
## 4 items, defined using the "SD_items" argument) and 6 Physical aggression items extracted
## from the IDAQ questionnaire. Since the 6 content items are balanced (3 positively worded and
## 3 negatively worded), there is no need to use the "unbalaced_items" argument.
```

```
ControlResponseBias(vampyr_example, content_factors = 1, SD_items = c(1,2,3,4),
    corr = "Pearson", contAC = TRUE, rotat = "none")

## For speeding purposes, some options have been disabled. Pearson correlation matrix is
## being used, but Polychoric matrix may be advisable. Parallel Analysis ('PA' argument)
## is disabled, as well as the factor scores and the path diagram.
```

`vampyr_example` *vampyr_example database*

Description

A database to be used as example in the functions included on `controlSDAQ` package. It contains the answers of 300 participants to the Physical aggression subscale of IDAQ questionnaire (Ruiz-Pamies, Lorenzo-Seva, Morales-Vives, Cosi, Vigil-Colet, 2014), which was developed for assessing Physical, Verbal and Indirect aggression. The original questionnaire contains 27 Likert-items, ranging from 1 to 5, and this dataset only contains the 4 SD markers and the 6 Physical aggression items.

Usage

```
data("vampyr_example")
```

Format

A data frame with 300 observations and 10 variables, where 6 items measuring Physical aggression and 4 Social Desirability markers. Items number 1, 2, 3 and 4 are SD markers (pure SD measures), and the remaining 6 items measure Physical aggression. Items number 5, 7 and 8 are positively worded, and items number 6, 9 and 10 are negatively worded.

Details

The original sample contains 750 participants, and the following database only contains 300 selected at random.

In the original dataset, which contains 27 items, the SD items correspond to items number 2, 8, 13 and 21, and the Phyisical aggression items are (in order) items number 1, 6, 17, 19, 20 and 25.

Source

More information about the questionnaire can be found at:

http://psico.fcep.urv.cat/tests/idaq/en/descripcion.html

References

Ruiz-Pamies, M., Lorenzo-Seva, U., Morales-Vives, F., Cosi, S., & Vigil-Colet, A. (2014). I-DAQ: a new test to assess direct and indirect aggression free of response bias. The Spanish Journal of Psychology, 17, E41. doi: 10.1017/sjp.2014.43

Examples

```
data(vampyr_example)
```

Index