Package ‘CMC’

February 15, 2013

Type Package

Title Cronbach-Mesbah Curve

Version 1.0

Date 2010-03-13

Author Michela Cameletti and Valeria Caviezel

Maintainer Michela Cameletti <michela.cameletti@unibg.it>

Description Calculation and plot of the stepwise Cronbach-Mesbah Curve

License GPL (>= 2)

LazyLoad yes

LazyData yes

Repository CRAN

Date/Publication 2012-10-29 13:13:38

NeedsCompilation no

R topics documented:

CMC-package 2
alpha.cronbach 3
alpha.curve 5
cain 7

Index **9**

`CMC-package` *Cronbach-Mesbach Curve*

Description

This package calculates and plots the step-by-step Cronbach-Mesbach curve, that is a method, based on the Cronbach alpha coefficient of reliability, for checking the unidimensionality of a measurement scale.

Details

Package:	CMC
Type:	Package
Version:	1.0
Date:	2010-03-13
License:	GPL (>= 2)
LazyLoad:	yes

Author(s)

Michela Cameletti and Valeria Caviezel

Maintainer: Michela Cameletti < `michela.cameletti@unibg.it` >

References

Curt, F., Mesbah, M., Lellouch, J. and Dellatolas, G. (1997) *Handedness scale how many and which items?* Laterality, 2, 137–154.

Hamon, A. and Mesbah, M. (2002) *Questionnaire reliability under the Rasch model*. Statistical Methods for Quality of Life Studies: Design, Measurement and Analysis. Mesbah, M., Cole, B.F. and Lee, M.L.T. (Eds.), Kluwer Academic Publishing, Boston, 155–168.

Mesbah, M. (2010) *Statistical quality of life*. Method and Applications of Statistics in the Life and Health Sciences. Balakrishnan, N. (Editor), Wiley, 839–864.

Nordmann, J., Mesbah, M., Berdeaux, G. (2005) *Scoring of Visual Field Measured through Humphrey Perimetry: Principal Component Varimax Rotation Followed by Validated Cluster Analysis*. Investigative Ophthalmology & Visual Science 46, 3169–3176.

`alpha.cronbach` *Cronbach reliability coefficient alpha*

Description

The function computes the Cronbach reliability alpha coefficient denoted with α.

Usage

```
alpha.cronbach(x)
```

Arguments

`x`	an object of class data.frame or matrix with n subjects in the rows and k items in the columns.

Details

Let $X_1, \ldots, X_k$ be a set of items composing a test and measuring the same underlying unidimensional latent trait. Moreover, let X_{ij} be the observed score (response) of a subject i $(i = 1, \ldots, n)$ on an item j $(j = 1, \ldots, k)$. Following the classical test theory, X_{ij} can be written as

$$X_{ij} = \tau_{ij} + \epsilon_{ij}$$

where τ_{ij}, the true score, and ϵ_{ij}, the error score, are two unknown random variables generally assumed to be independent (or at least not correlated). In particular, the true score is given by

$$\tau_{ij} = \mu_j + a_i$$

where μ_j is a fixed effect and a_i is a random effect with zero mean and variance σ_a^2, whereas ϵ_{ij} is a random effect with zero mean and variance σ_ϵ^2. Moreover, ϵ_{ij} and a_i are not correlated and for all $j = 1, \ldots, k$ and for $t \neq s$, (a_t, ϵ_{tj}) and (a_s, ϵ_{sj}) are independent.

The *reliability* ρ of any item is defined as the ratio of two variances: the variance of the true (unobserved) measure and the variance of the observed measure. Under the parallel model (see Lord and Novick, 1968), it can be shown that

$$\rho = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_\epsilon^2}$$

where σ_a^2 corresponds to the between-subject variability while σ_ϵ^2 is the variance of the measurement error. It is possible to prove that ρ is also the constant correlation between any two items. The reliability of the sum of k items is given by the well-known **Spearman-Brown** formula:

$$\tilde{\rho} = \frac{k\rho}{k\rho + (1-\rho)}.$$

The maximum likelihood estimate of $\tilde{\rho}$, under the assumption of Normal distribution for the error, is known as the **Cronbach alpha coefficient**, denoted with α.

The formula for computing α is given by

$$\alpha = \frac{k}{k-1}\left[1 - \frac{\sum_{j=1}^{n} s_j^2}{s_{TOT}^2}\right]$$

where $s_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_{ij} - \bar{X}_j)^2$, $s_{TOT}^2 = \frac{1}{nk-1}\sum_{i=1}^{n}\sum_{j=1}^{k}(X_{ij} - \bar{X})^2$, $\bar{X}j = \frac{1}{n}\sum_{i=1}^{n} X_{ij}$ and $\bar{X} = \frac{1}{nk}\sum_{i=1}^{n}\sum_{j=1}^{k} X_{ij}$.

Value

The function returns the value α of the Cronbach reliability coefficient computed as described above. The coefficient takes values in the interval [0,1]. If the actual variation amongst the subjects is very small, then the reliability of the test measured by α tends to be small. On the other hand, if the variance amongst the subject is large, the reliability tends to be large.

Warning

No missing values are admitted.

Author(s)

Michela Cameletti and Valeria Caviezel

References

Cronbach, L.J. (1951) *Coefficient alpha and the internal structure of tests*. Psychometrika, 16, 297–334.

Lord, F.M. and Novick, M.R. (1968) *Statistical Theories of Mental Test Scores*. Addison-Wesley Publishing Company, 87–95.

See Also

See Also `alpha.curve` and `cain`

Examples

```
data(cain)
out = alpha.cronbach(cain)
out
```

alpha.curve *Step-by-step Cronbach-Mesbah Curve*

Description

The function calculates and plots the Cronbach-Mesbah Curve for a given data set.

Usage

```
alpha.curve(x)
```

Arguments

x	an object of class data.frame or matrix with n subjects in the rows and k items in the columns.

Details

There is a direct connection between the Cronbach alpha coefficient α (see alpha.cronbach) and the percentage of variance explained by the first component in the Principal Component Analysis (PCA) on k items. The PCA is usually based on the analysis of the roots of the correlation matrix R of k variables which, under the hypothesis of a parallel model (see Lord and Novick, 1968) is:

$$\text{R} = \begin{array}{cccc} 1 & \rho & ... & \rho \\ \rho & 1 & ... & \rho \\ ... & ... & ... & ... \\ \rho & \rho & ... & 1 \end{array}$$

This matrix has only two different roots. The greater root is $\lambda_1 = 1 + \rho(k-1)$ and the other roots are $\lambda_2 = \ldots = \lambda_k = 1 - \rho = \frac{k-\lambda_1}{k-1}$. Thus, using the Spearman-Brown formula, we can express the reliability of the sum of the k items as follows:

$$\tilde{\rho} = \frac{k}{k-1}\left[1 - \frac{1}{\lambda_1}\right].$$

This indicates that there is a monotonic relationship between $\tilde{\rho}$, estimated by α, and the first root λ_1, which in practice is estimated using the observed correlation matrix and thus gives the percentage of variance of the first principal component. Then, α is considered as a measure of unidimensionality.

In particular, to assess the unidimensionality of a set of items, it is possible to plot a curve, called step-by-step **Cronbach-Mesbah** curve, which reports the number of items (from 2 to k) on the x-axis and the corresponding maximum α coefficient on the y-axis obtained through the following steps:

1. first of all the Cronbach coefficient $\alpha = \tilde{\alpha}^0$ is computed using all the k items.
2. One at a time, the i-th item ($i = 1, \ldots, k$) is left out and the Cronbach coefficient, denoted by α_{-i}, is computed using the remaining $(k-1)$ items. All the coefficients are collected in a set

given by

$$\tilde{\alpha}^1 = (\alpha_{-1}, \ldots, \alpha_{-j}, \ldots, \alpha_{-k})$$

where the apex refers to the number of item removed at each time. Then, the maximum of $\tilde{\alpha}^1$ is detected and the corresponding item is taken out. For example, if α_{-j} is the maximum of $\tilde{\alpha}^1$, the j-th item is removed definitely from the scale.

3. The procedure of step 2 is repeated conditionally on the item removed previously. Supposing that item j was removed, the remaining items are left out one at a time and the corresponding Cronbach coefficient is calculated. This gives rise to the following set of $(k-1)$ coefficients

$$\tilde{\alpha}^2 = \left(\alpha_{-(1,j)}, \ldots, \alpha_{-(j-1,j)}, \alpha_{-(j+1,j)}, \ldots, \alpha_{-(k,j)}\right).$$

The item corresponding to the maximum of $\tilde{\alpha}^2$ is then removed definitely. For example, if α_{-1} is the maximum of $\tilde{\alpha}^2$, the first item is removed definitely from the scale together with the j-th item removed at step 2.

This procedure is repeated until only 2 items remain. Note that at each step the removed item is the one which leaves the scale with its *maximum* α value. If we remove a poor item, the α coefficient will increase, whereas if we remove a good item α must decrease. More precisely, the Spearman-Brown formula shows that increasing the number of items leads to increase the reliability of the total score. Thus, a decrease of the Cronbach-Mesbah curve, after adding a variables, would suggest that the added variable do not constitute an unidimensional set with the other variables. On the other hand, if the step-by-step Cronbach-Mesbah curve increases monotonically, then all the items contribute to measure the same latent trait and the bank of items is characterized by unidimensionality.

Value

The functions returns:
1) an object of class data.frame with 3 columns. The first column *N.Item* contains the number of item used for computing the Cronbach α coefficients. It contains the values between 2 and k corresponding, respectively, to the case when only 2 items or all the items are used. The second column, *Alpha.Max*, refers to the maximums of the Cronbach coefficients calculated at each step of the procedure, that is $(\max \tilde{\alpha}^{k-2}, \ldots, \max \tilde{\alpha}^1, \max \tilde{\alpha}^0)$. Finally, the last column, *Removed.Item*, reports the name of the item removed at each step, that is $(arg \max \tilde{\alpha}^{k-2}, \ldots, arg \max \tilde{\alpha}^1, arg \max \tilde{\alpha}^0)$.
2) The corresponding Cronbach-Mesbah curve plot created using the first 2 columns of the data.frame described above. Note that also the names of the removed items are reported in the graph.

Warning

No missing values are admitted.

Author(s)

Michela Cameletti and Valeria Caviezel

References

Curt, F., Mesbah, M., Lellouch, J. and Dellatolas, G. (1997) *Handedness scale how many and which items?* Laterality, 2, 137–154.

Hamon, A. and Mesbah, M. (2002) *Questionnaire reliability under the Rasch model*. Statistical Methods for Quality of Life Studies: Design, Measurement and Analysis. Mesbah, M., Cole, B.F. and Lee, M.L.T. (Eds.), Kluwer Academic Publishing, Boston, 155–168.

Mesbah, M. (2010) *Statistical quality of life*. Method and Applications of Statistics in the Life and Health Sciences. Balakrishnan, N. (Editor), Wiley, 839–864.

Nordmann, J., Mesbah, M., Berdeaux, G. (2005) *Scoring of Visual Field Measured through Humphrey Perimetry: Principal Component Varimax Rotation Followed by Validated Cluster Analysis*. Investigative Ophthalmology & Visual Science 46, 3169–3176.

See Also

See Also `alpha.cronbach` and `cain`

Examples

```
data(cain)

out = alpha.curve(cain)
out
```

`cain` *Data concerning Computer Anxiety INdex (cain)*

Description

The data are selected from Chapter 6 of Bond and Fox (2007). In particular we consider only 12 items (I7, I9, I11, I13, I14, I16, I20, I22, I23, I24, I25, I26) in the set of the original 26 items. The data are indicative of a person's feelings of anxiety toward computers. Each item has six response options (1 = strongly agree, 2 = agree, 3 = slightly agree, 4 = slightly disagree, 5 = disagree, 6 = strongly disagree). The subjects are 371 7-years students.

Usage

```
data(cain)
```

Format

A data frame with 371 subjects and 12 items.

Source

Bond, T.G. and Fox, C.M. (2007). *Applying the Rasch model - Fundamental measurement in the human sciences*. Second ed. Lawrence Erlbaum Associates publishers, Mahwah, N.J.

See Also

See Also `alpha.cronbach` and `alpha.curve`

Examples

```
data(cain)
dim(cain)
str(cain)
```

Index

∗Topic **datasets**
    cain, 7
∗Topic **package**
    alpha.cronbach, 3
    alpha.curve, 5
    CMC-package, 2

alpha.cronbach, 3, *5*, *7, 8*
alpha.curve, *4*, 5, *8*

cain, *4*, *7*, 7
CMC *(*CMC-package*)*, 2
CMC-package, 2