



HAL
open science

The R package CMC to calculate the Cronbach-Mesbah curve

Michela Cameletta, Valeria Caviezel

► **To cite this version:**

Michela Cameletta, Valeria Caviezel. The R package CMC to calculate the Cronbach-Mesbah curve. *Annales de l'ISUP*, 2013, 57 (1-2), pp.59-68. <hal-03615258>

HAL Id: hal-03615258

<https://hal.science/hal-03615258v1>

Submitted on 21 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

The R package CMC to calculate the Cronbach-Mesbah curve

Michela Cameletti ^a, Valeria Caviezel

Department of Mathematics, Statistics, Computer Science and Applications, Bergamo (Italy)

Abstract. This paper describes the R package CMC that calculates and plots the step-by-step Cronbach Mesbah curve. This is a simple but effective method, based on the Cronbach coefficient of reliability that can be used for checking the unidimensionality of a measurement scale adopted in a psychological or social science survey.

So far this graphical tool has been implemented only in an SAS macro called ANAQOL and therefore its dissemination has been somewhat limited to SAS users. The main aim of the CMC package is to extend the use of this method to a wider scientific community. To pursue this goal, the package is written in the open-source R language, which is rapidly evolving as the lingua franca of data analysis and statistical computing in many scientific fields. This paper illustrates the Cronbach-Mesbah curve via simulated datasets from the Rasch model.

Keywords: Cronbach reliability coefficient, Rasch model, R software.

1. Introduction

In many areas of psychology, health and other behavioral science disciplines, it is often not possible to measure the trait of primary interest directly. In such cases the researcher has to examine the traits indirectly by collecting information on variables that can be measured or observed directly, and which can also realistically be assumed to be indicators, in some sense, of the concept of real interest. A latent trait cannot be measured directly but is related to a number of measurable or manifest variables. Since the measurement of a variable requires an instrument, a latent trait is generally assessed by a questionnaire consisting of a number of dichotomous or polytomous items (observed variables).

From the definition of the Institute for Objective Measurement (IOM, www.rasch.org): "*measurement is the repetition of a unit amount that maintains its size, within an allowable range of error, no matter which instrument, intended to measure the variable of interest, is used and no matter who or what relevant person or thing is measured*". An objective measurement, as defined by IOM, should satisfy certain properties such as order, unidimensionality, local independence, linearity and specific objectivity. Among these properties, unidimensionality can be theoretically defined as the property of an instrument in measuring a single characteristic of the measured object (see [1]). Even if the measured object is characterized by many attributes, it is essential to focus on of just one of the measurement. When a total score contains more features of a subject, it is difficult to have precise information regarding the ability of the subject. In practice, when a measurement instrument is created, we aim to come as close as possible to this ideal in order to obtain a scale that is useful and relevant in the case study. In the human sciences it is still difficult to isolate a single attribute to measure. In any case, to satisfy the property of unidimensionality, there must be a dominant latent trait. If a questionnaire is characterized by unidimensionality, then the difference between two subjects in responding to a set of items depends only on the single latent trait.

In this paper we introduce a simple but effective graphical method to check the unidimensionality of a scale measurement in the framework of the classical test theory. It is the step-by-step Cronbach-Mesbah Curve proposed by Mesbah and described in [2], [3] and [4] and based on the Cronbach coefficient of reliability. In particular, the main goal of this paper is to present a package, called CMC, for calculating and plotting the Cronbach-Mesbah curve. As the package is written in the open-source R language, which is rapidly evolving as the lingua franca of data analysis, the

package is expected to receive a wide dissemination in the scientific fields of the social sciences. The package is available from the Comprehensive R Archive Network at <http://cran.r-project.org/package=CMC>. Up till now this graphical tool has been implemented only in an SAS macro called ANAQOL [5] and therefore its dissemination has been somewhat limited to SAS users. This paper illustrates the Cronbach-Mesbah curve via simulated datasets from the Rasch model [6] in order to check the curve performance when the unidimensionality hypothesis is violated.

In the next section we illustrate all the steps to construct the Cronbach-Mesbah curve, starting from the hypothesis of the parallel model, while in section 3 the R package CMC is presented. In section 4 we illustrate the package with several simulated data from a dichotomous Rasch model with and without unidimensionality property. The paper ends with some concluding remarks.

2. Cronbach-Mesbah Curve

To study the unidimensionality of a set of observable items X_1, \dots, X_k , we introduce the parallel model where X_{ij} is the observed score (response) of subject i ($i = 1, \dots, n$) on item j ($j = 1, \dots, k$). Under the hypothesis of the parallel model [7], X_{ij} can be written as

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

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). 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_e^2 . Moreover, ε_{ij} and a_i are not correlated and for all $j = 1, \dots, k$ and (a_t, ε_{tj}) and (a_s, ε_{sj}) are independent for $t \neq s$ ($t, s = 1, \dots, n$). It is then easy to show that $E(X_{ij}) = \mu_j$, $Var(X_{ij}) = \sigma_a^2 + \sigma_e^2$ and $Cov(X_{ij}, X_{iz}) = \sigma_a^2$ ($j, z = 1, \dots, k$).

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 assumption that the observed items follow the parallel model, i.e. all two by two item correlations are equal, it can be shown that

$$\rho = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$$

where σ_a^2 corresponds to the between-subject variability and σ_e^2 is the variance of the measurement error. Moreover, 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 following 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 ε , is known as the Cronbach reliability coefficient [8], denoted with α .

The formula for computing α is given by:

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

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 $\bar{X} = \frac{1}{nk} \sum_{i=1}^n \sum_{j=1}^k X_{ij}$.

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

We can expect the same behavior for the Spearman-Brown formula $\tilde{\rho}$ and for its estimate α . In particular, as $\tilde{\rho}$ is an increasing function of the item number k , if every item follows the parallel model, then the reliability of a sum increases as k increases. Therefore, if α does not increase or even decreases as k increases, it means that the added item does not belong to the same dimension, i.e. does not follow the parallel model.

It is possible to show that there is a direct connection between the Cronbach reliability coefficient α and the percentage of variance explained by the first component of the principal component analysis (PCA) performed on the 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, is given by:

$$R = \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \dots & \dots & \dots & \dots \\ \rho & \rho & \dots & 1 \end{bmatrix}.$$

This matrix has only two different roots. The greater root is $\lambda_1 = 1 + \rho(k - 1)$ and the other roots are $\lambda_2 = \dots = \lambda_k = 1 - \rho = \frac{k - \lambda_1}{k - 1}$. 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 that gives the percentage of variance explained by the first principal component. In this case, 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 the 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 via 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 v -th item ($v = 1, \dots, k$) is left out and the Cronbach coefficient, denoted by α_{-v} , is computed using the remaining $(k - 1)$ items. All the coefficients are collected in a set given by:

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

where the apex refers to the fact that only one item has been removed. Then, the maximum of $\tilde{\alpha}^1$ is detected and the corresponding item is removed. For example, if α_{-j} is the maximum of $\tilde{\alpha}^1$, the j -th item is removed definitively 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 using $(k - 2)$ items. This gives rise to the following set of $(k - 1)$ coefficients:

$$\tilde{\alpha}^2 = (\alpha_{-(1,j)}, \dots, \alpha_{-(j-1,j)}, \alpha_{-(j+1,j)}, \dots, \alpha_{-(k,j)})$$

where the apex refers to the fact that 2 items have been removed. The item corresponding to the maximum of $\tilde{\alpha}^2$ is then removed definitively. For example, if $\alpha_{-(1,j)}$ is the maximum of $\tilde{\alpha}^2$, the first item is removed definitively 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 value of the Cronbach coefficient α . 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 an increase in the reliability of the total score. Thus, a decrease of the Cronbach-Mesbah curve, after

adding a variable, would suggest that the added variable does 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 measuring the same latent trait and the bank of items is characterized by unidimensionality.

3. The CMC package

The code for computing the Cronbach-Mesbah Curve has been organized in a library developed using the open-source R programming environment [9]. The result is an R package called *CMC* (from the acronym of Cronbach-Mesbah Curve) which is composed by two functions. The first one, called `alpha.cronbach`, computes the Cronbach reliability coefficient α given by Eq. (1) for an object of class `data.frame` (or `matrix`) with n subjects (rows), k items (columns) and no missing values. Just to illustrate the procedure, with reference to the dataset `cain` included in the *CMC* package, the Cronbach coefficient is promptly calculated typing the following instructions:

```
library("CMC")
data("cain")
alpha.cronbach(cain)
[1] 0.8316127
```

The second function included in the package is called `alpha.curve` and computes the Cronbach-Mesbah curve following the procedure described in Section 2. The function takes as input an object of class `data.frame` (or `matrix`) with n subjects (rows), k items (columns) and no missing values. The output of the function is given by 2 objects:

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 coefficients. The values of the columns are 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 $(\max \tilde{\alpha}^{k-2}, \dots, \max \tilde{\alpha}^1, \tilde{\alpha}^0)$. Finally, the last column (*Removed.Item*) reports the name of the item removed at each step, that is $(\operatorname{argmax} \tilde{\alpha}^{k-2}, \dots, \operatorname{argmax} \tilde{\alpha}^1, --)$. Note that the last element contains no name because this corresponds to the case of the complete Cronbach coefficient $\tilde{\alpha}^0$ when all the k items are used.
2. The plot of the corresponding Cronbach-Mesbah curve created using the columns *N.Item* and *Alpha.Max* of the `data.frame` described above. Moreover, the names of the removed items are also reported in the graph.

With regard to the `cain` dataset, after giving the instruction:

```
>alpha.curve(cain)
```

we obtain a table with three columns and the corresponding Cronbach-Mesbah curve reported, respectively, in Table I and Figure 1.

Table I: The output returned by the `alpha.curve` function for the `cain` dataset

| | <i>N.Item</i> | <i>Alpha.Max</i> | <i>Removed.Item</i> |
|----|---------------|------------------|---------------------|
| 1 | 2 | 0.6715459 | 120 |
| 2 | 3 | 0.6988724 | 116 |
| 3 | 4 | 0.7351930 | 113 |
| 4 | 5 | 0.7510661 | 114 |
| 5 | 6 | 0.7676828 | 19 |
| 6 | 7 | 0.7827610 | 17 |
| 7 | 8 | 0.7973100 | 125 |
| 8 | 9 | 0.8103813 | 111 |
| 9 | 10 | 0.8193292 | 123 |
| 10 | 11 | 0.8251892 | 122 |
| 11 | 12 | 0.8316127 | -- |

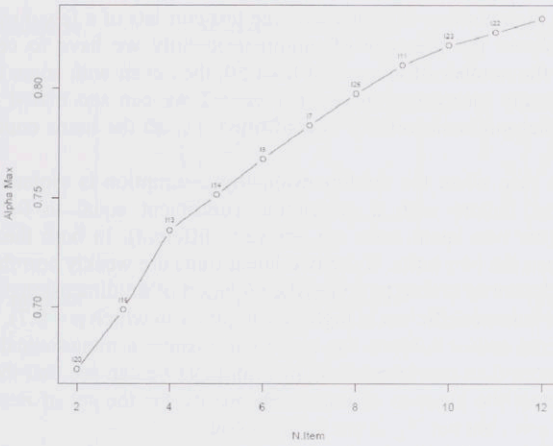


Figure 1: Cronbach-Mesbah curve for the `cain` dataset of CMC package

4. Simulation study

In this section we check the performance of the Cronbach-Mesbah curve with simulation studies. To simulate the data we use some R code that generates data according to the dichotomous Rasch model, together with some deviations from them.

Georg Rasch [6] developed a mathematical model for constructing measures based on the probabilistic relationship between any *item's difficulty* and any *person's ability*. The use of a Rasch models enables predictions of how persons at each level of ability are expected to perform regarding each item. The model predicts the conditional probability of an ordered binary outcome, given the *person's ability* and the *item's difficulty*. This probability is expressed in logistic form. If “correct” answers are coded as 1 and “incorrect” answers are coded as 0, the model expresses the probability of obtaining a correct answer as a function of the size of the difference between the ability of the respondent and the difficulty of the item. This probability is given by:

$$P(X_{ij} = 1 | \vartheta_i, \delta_j) = \frac{\exp\{\vartheta_i - \delta_j\}}{1 + \exp\{\vartheta_i - \delta_j\}}$$

where ϑ_i ($i = 1, 2, \dots, n$) is the respondent parameter (*person's ability*) and δ_j ($j = 1, 2, \dots, k$) is the item parameter (*item's difficulty*).

The Rasch model is characterized by unidimensionality and additivity. Unidimensionality means that a single construct is being measured (i.e. the latent trait is a scalar). Rasch models produce measurements on an interval scale. This implies additivity on the scale and invariance over the entire continuum, if the data fit the model.

In this paper, following [10] and [11], we consider five values for the sample size n (100, 250, 500, 1000 and 4000 subjects) and five for the test length k (10, 15, 30, 50 and 75 dichotomous items). Concerning the unidimensional case, to simulate the *persons' ability* we consider a standard normal distribution, while for the *items' difficulty* a uniform distribution (-3, +3). Then we compute some Cronbach-Mesbah curves for all the 25 generated datasets. The most important ones are shown in Figure 2. If the number of items is less than 10, we can not find monotonically increasing curves

even if we consider a high number of subjects. If the test consists of a few dozen items, in order to obtain a curve that shows the presence of unidimensionality we have to consider at least 500 subjects, while when the number of items is at least 50, then even with some hundred subjects we can have a monotonically increasing curve. In Figure 2 we can see that if the curves increase monotonically, then the unidimensionality is confirmed, i.e. all the items contribute to measuring the same latent trait.

We also consider the case when the unidimensionality assumption is violated. We simulate data considering two latent factors with a correlation coefficient equal to 0.2 (a low degree of correlation) and 0.7 (the two latent traits are not very different). In both the cases the items are equally divided between the two traits. If the two latent traits are weakly correlated (see Figure 3 in which $\rho = 0.2$), then the curve is able to reveal the violation of unidimensionality among the items, but not when the correlation coefficient is high (see Figure 4 in which $\rho = 0.7$). Then if we consider a great number of items and/or subjects the curve can assume a monotonically increasing aspect even in presence of more than one latent trait. In summary, we can see that the Cronbach-Mesbah curves are able to detect the absence of unidimensionality for the set of items if the correlation coefficient is close to zero, but not if ρ is not far from one.

Finally we considered the case with simulated data considering two latent traits with a correlation coefficient equal to -0.2 and -0.7. In all simulations the Cronach-Mesbah curve is able to detect the absence of unidimensionality.

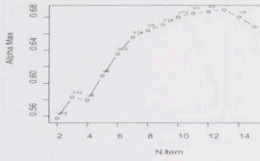
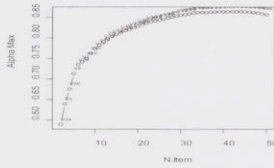
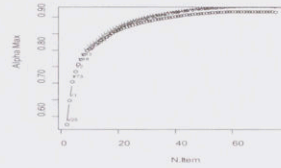
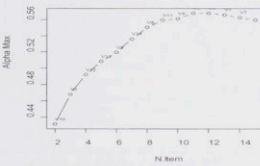
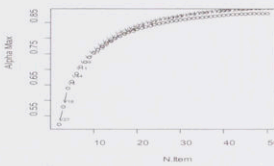
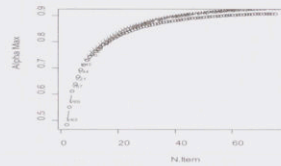
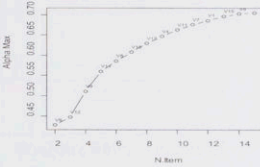
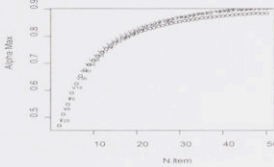
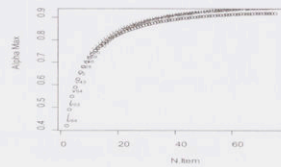
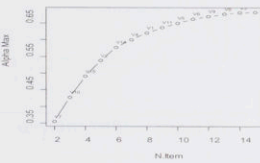
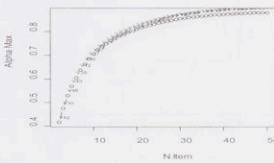
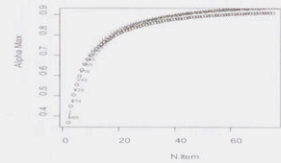
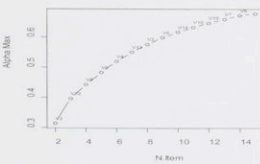
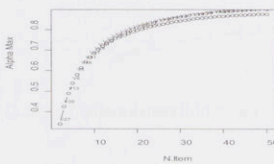
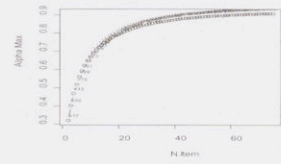
15 items – 100 subjects*50 items – 100 subjects**75 items – 100 subjects**15 items – 250 subjects**50 items – 250 subjects**75 items – 250 subjects**15 items – 500 subjects**50 items – 500 subjects**75 items – 500 subjects**15 items – 1000 subjects**50 items – 1000 subjects**75 items – 1000 subjects**15 items – 4000 subjects**50 items – 4000 subjects**75 items – 4000 subjects*

Figure 2: Cronbach-Mesbah curves in the case of unidimensionality with simulate data

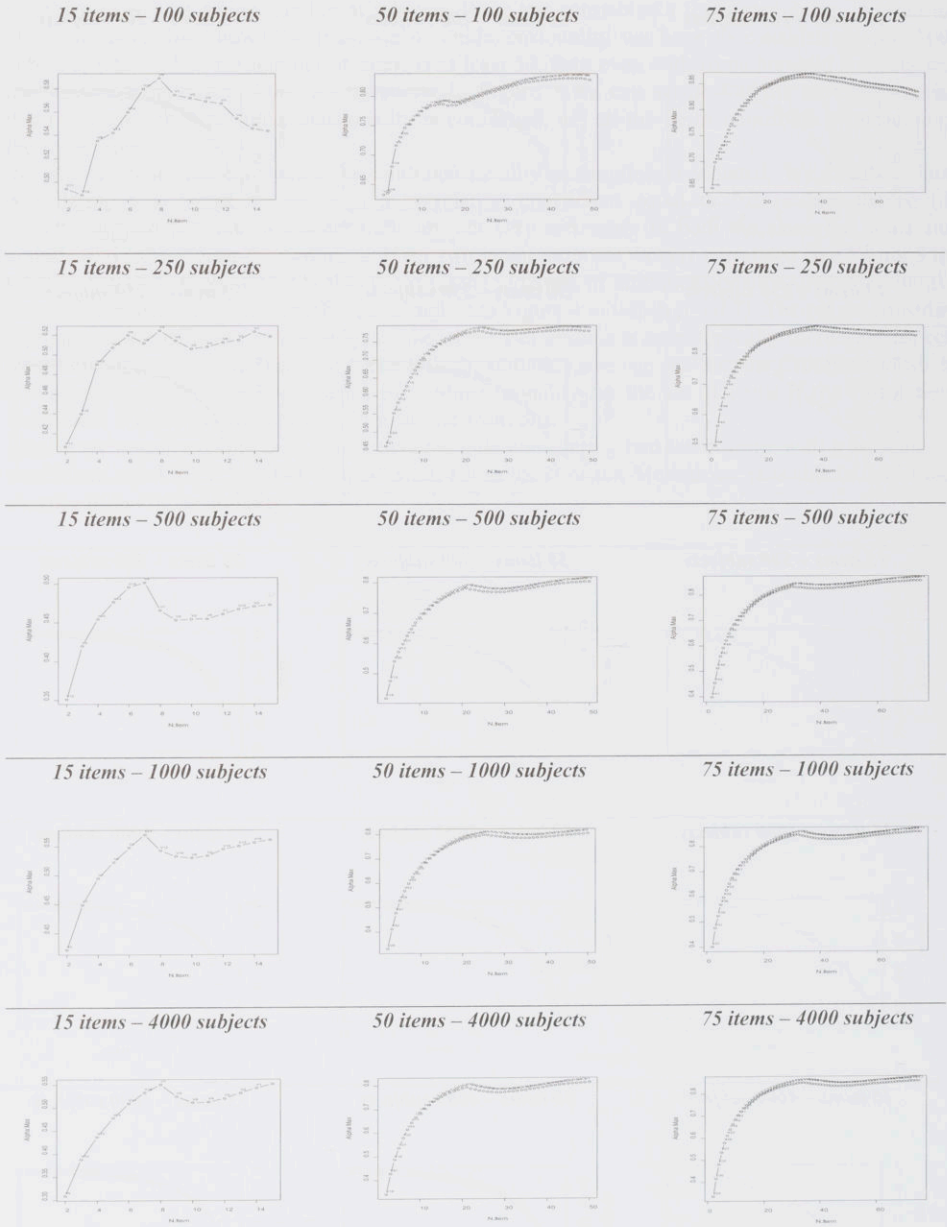
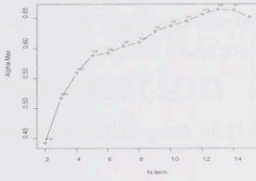
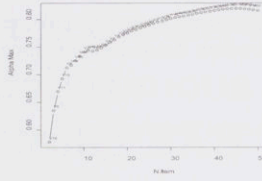
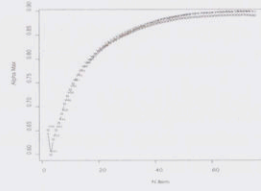
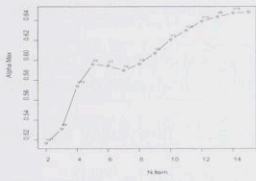
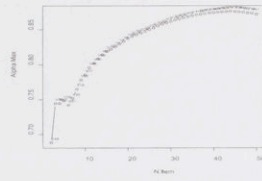
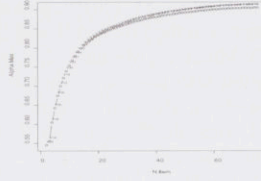
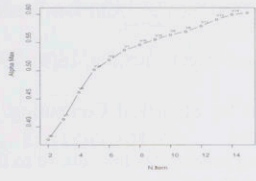
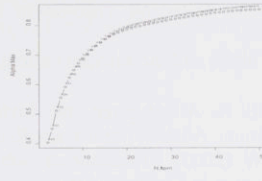
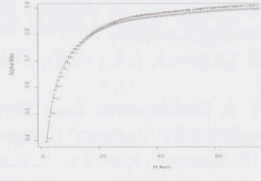
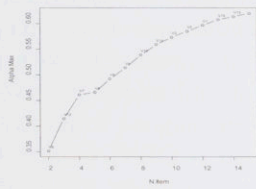
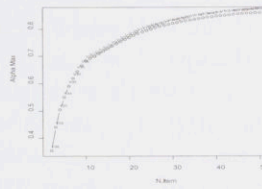
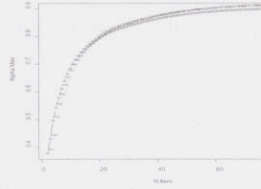
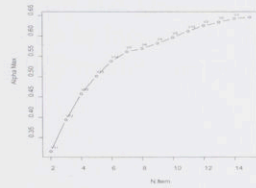
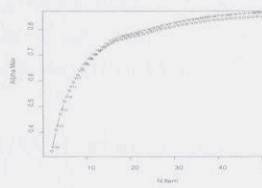
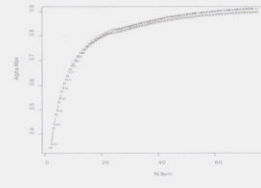


Figure 3: Cronbach-Mesbah curves in the case of bidimensionality ($\rho = 0.2$) with simulate data

15 items – 100 subjects*50 items – 100 subjects**75 items – 100 subjects**15 items – 250 subjects**50 items – 250 subjects**75 items – 250 subjects**15 items – 500 subjects**50 items – 500 subjects**75 items – 500 subjects**15 items – 1000 subjects**50 items – 1000 subjects**75 items – 1000 subjects**15 items – 4000 subjects**50 items – 4000 subjects**75 items – 4000 subjects*Figure 4: Cronbach-Mesbah curves in the case of bidimensionality ($\rho = 0.7$) with simulate data

4. Concluding remarks

This paper described the R package CMC for the Cronbach Mesbah curve and used several simulated data for illustration.

To improve our work it would be interesting to consider the case when the items are polytomous and evaluate the behavior of the curve when the underlying dimensions are more than two.

Acknowledgments

The authors thank Mounir Mesbah (Université Pierre et Marie Curie - Paris VI) for many helpful comments and suggestions.

References

1. Penta, M., Arnould, C. & Decruynaere, C. (2005). Développer et interpréter une échelle de mesure. Application du modèle de Rasch. Mardaga.
2. Mesbah, M. (2010). Statistical quality of life. In N. Balakrishnan (ed.), *Method and Applications of Statistics in the Life and Health Sciences*, Wiley, 839-864.
3. Moret, L., Mesbah, M., Chwalow, J. & Lellouch, J. (1993). Validation interne d'une échelle de mesure: relation entre analyse en composantes principales, coefficient alpha de Cronbach et coefficient de corrélation intra-classe. *Revue d'Epidémiologie et de Santé Publique*, 41(2), 179-186.
4. Mesbah, M., Lellouch, J., Curt, F. & Dellatolas, G. (1997). Handedness scale: how many and which items? *Laterality*, 2, 137-154.
5. Hardouin, J.B. & Mesbah, M. (2007). The SAS Macro-Program AnaQol to Estimate the Parameters of Item Responses Theory Models. *Communications in Statistics - Simulation and Computation*, 36, 437-453.
6. Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Danish Institute for Educational Research.
7. Lord, F.M. & Novick, M.R. (1968). *Statistical Theories of Mental Test Scores*. Addison-Wesley Publishing Company.
8. Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
9. R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
10. Suárez-Falcón, J.C. & Glas, C.A.W. (2003). Evaluation of global testing procedures for item fit to the Rasch model. *British Journal of Mathematical and Statistical Psychology*, 56, 127-143.
11. Brentari, E., Golia, S., (2007). Unidimensionality in the Rasch model: how to detect and interpret. *Biometrics*, 67(3), 253-261.