Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 12, Number 2, February 2007 ISSN 1531-7714

Determining the Number of Factors to Retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis

Rubén Daniel Ledesma
CONICET / Universidad Nacional de Mar del Plata, Mar del Plata (Argentina)

Pedro Valero-Mora
Universidad de Valencia, Valencia (Spain).

Parallel Analysis is a Monte Carlo simulation technique that aids researchers in determining the number of factors to retain in Principal Component and Exploratory Factor Analysis. This method provides a superior alternative to other techniques that are commonly used for the same purpose, such as the Scree test or the Kaiser's eigenvalue-greater-than-one rule. Nevertheless, Parallel Analysis is not well known among researchers, in part because it is not included as an analysis option in the most popular statistical packages. This paper describes and illustrates how to apply Parallel Analysis with an easy-to-use computer program called ViSta-PARAN. ViSta-PARAN is a user-friendly application that can compute and interpret Parallel Analysis. Its user interface is fully graphic and includes a dialog box to specify parameters, and specialized graphics to visualize the analysis output.

Exploratory Factor Analysis (EFA) and Principal Component Analysis (PCA) are multivariate statistical techniques widely used in social and behavioral sciences. In a previous paper in *Practical Assessment, Research and Evaluation*, Costello and Osborne (2005) discuss common practices in studies using these techniques, and provide researchers with a compilation of "best practices" in EFA. One of the topics these authors discuss is the *number of factors to retain,* which is the most important decision to make after factor extraction. Mistakes at this stage, such as extracting too few or too many factors, may lead to erroneous conclusions in the analysis. Unfortunately, the most popular statistical programs do not provide users with the most accurate methods to solve this problem, as Costello and Osborne (2005) note. This is the case with Parallel Analysis (PA), which is one of the methods most recommended to deal with the number-of-factors-to-retain problem, but is not available in commonly used statistical packages. In view of this, the present paper: 1) briefly reviews some methods to deal with the number-of-factor-to-retain problem, 2) provides a general description of the PA technique, 3) introduces the *ViSta-PARAN* software, a new and freely available computer program to carry out PA; and 4) illustrates how to apply PA by using *ViSta-PARAN*.

THE NUMBER-OF-FACTORS-TO-RETAIN PROBLEM

Various authors have commented on the importance of deciding how many factors or components to retain when applying EFA and PCA (e.g., Fabrigar, Wegener, MacCallum, & Strahan, 1999; Hayton, Allen, & Scarpello, 2004). Hayton et al. (2004) states three reasons why this decision is so important. Firstly, it can affect EFA results more than other decisions, such as selecting an extraction method or the factor rotation method, since there is evidence of the relative robustness of EFA with regards to these matters (Zwick & Velicer, 1986). Secondly, the EFA requires that a balance be struck between "reducing" and adequately "representing" the correlations that exist in a group of variables; therefore, it's very usefulness depends on distinguishing important factors from trivial ones. Lastly, an error in terms of selecting the number of factors can significantly alter the solution and the interpretation of EFA results. Underextraction can lead to the loss of relevant information and a substantial distortion in the solution; for example, in the variables loading. On the other hand, overextraction although less grave, can lead to factors with few substantial loading, which can be difficult to interpret and/or replicate (Zwick & Velicer, 1986). To sum up, both underextraction and overextraction have consequences that adversely impact the EFA's efficiency and meaning.

Nonetheless, it is also important to recognize that in EFA, previous theory and research play an important role, such that the researcher should consider both substantive and statistical issues when deciding on the number of factors (Fabrigar et al., 1999). Of course, when theory is a guiding force, the better option may be the Confirmatory Factor Analysis approach.

Methods for determining the number of factors to retain

Given the importance of this decision, different methods have been proposed to determine the number of factors to retain. Further, various studies have been undertaken to evaluate the individual or comparative efficiency of these methods (Horn & Engstrom, 1979**;** Zwick and Vellicer, 1986; Hubbard & Allen, 1987; Lautenschlager, 1989; Buja and Eyuboglu, 1992). These studies are generally concerned with an evaluation of the ability of these methods to determine the number of non-trivial factors in data generated by simulation. A brief review of the principal focus areas and results of some of these studies follows. Special consideration is given to the work of Zwick and Vellicer (1986) who compared different methods under different conditions.

K1 - Kaiser's eigenvalue-greater-than-one rule

The *K1* method proposed by Kaiser (1960) is perhaps the best know and most utilized in practice (Fabrigar et. al, 1999). According to this rule, only the factors that have eigenvalues greater than one are retained for interpretation. Despite the simplicity of this method, many authors agree that it is problematic and inefficient when it comes to determining the number of factors. Fabrigan et. al (1999) notes three problems with using this method. First, the method was proposed for the PCA case - eigenvalues of the correlation matrix with unities at the diagonal -, and it is not a valid rule in the EFA case - eigenvalues of the correlation matrix with communality estimates at the diagonal -. Secondly, as with all mechanical rules, this method can lead to arbitrary decisions; for instance, it doesn't make much sense to regard a factor with an eigenvalue of 1.01 as 'major' and one with an eigenvalue of .99 as 'trivial'. Lastly, in various simulation studies with PCA and EFA, this method has demonstrated tendency to substantially overestimate the number of factors, and, in some cases, even underestimate them (Zwick & Velicer, 1986). In fact, there is a problematic relation between the number of factors retained and the number of variables. Kaiser himself reported that the number of components retained by *K1* is commonly between one-third and one-fifth or one-sixth the number of variables included in the correlation matrix (Zwick & Velicer, 1986). In summary, despite *K1*'s widespread use, experts agree that it has deficiencies and that its use is not recommended.

Cattell's Scree test

Another popular approach is based on the Cattell's *Scree test* (Cattell, 1966), which involves the visual exploration of a graphical representation of the eigenvalues. In this method, the eigenvalues are presented in descending order and linked with a line. Afterwards, the graph is examined to determine the point at which the last significant drop or break takes place—in other words, where the line levels off. The logic behind this method is that this point divides the important or major factors from the minor or trivial factors.

This method has been criticized for its subjectivity, since there is not an objective definition of the cutoff point between the important and trivial factors. Indeed, some cases may present various drops and possible cutoff points, such that the graph may be ambiguous and difficult to interpret. Zwick and Velicer (1986) indicate that, when analyzing how examiners interpret the Scree test, the results can be very varied, depending on the training received by the examiners and also the nature of the solution. Nonetheless, they also mention that the Scree test can be more accurate and less variable than the *K1* method. For instance, factors that affect the *K1* method, such as the number of variables, do not appreciably affect the Scree test. Lastly, Zwick and Velicer (1986) also note that the Scree test has a tendency to overestimate, and conclude that, given the existence of better methods, its use is also not recommended.

Velicer's MAP test - Minimum Average Partial

Velicer (1976) proposes the MAP test (Minimum Average Partial), a method based on the application of PCA and in the subsequent analysis of partial correlation matrices. This rule employs the EFA concept of 'common' factors to determine how many components to extract. The method seeks to determine what components are common, and is proposed as a rule to find the best factor solution, rather than to find the cutoff point for the number of factor. Due to the way it is calculated, one of the properties of this solution is that it does not retain factors that have too few loading; when using this rule, "at least two variables will have high loadings on each retained component" (Zwick and Velicer, 1986). In terms of its ability to select factors, MAP has proven superior to the techniques described in the previous paragraphs (Wood, Tataryn, & Gorsuch, 1996; Zwick and Velicer, 1986). Zwick and Velicer (1986) found that, through their simulation study, the MAP method was quite accurate under many conditions, although under certain conditions it may reveal a tendency to underestimate the number of factors. Specifically, when there are low variables loading and low number of variables per component, the MAP method consistently underestimated the number of major components.

Horn's Parallel Analysis

Horn (1965) proposes PA, a method based on the generation of random variables, to determine the number of factors to retain. PA, as it was introduced early on by Horn (1965), compares the observed eigenvalues extracted from the correlation matrix to be analyzed with those obtained from uncorrelated normal variables. From a computational point of view, PA implies a Monte Carlo simulation process, since 'expected' eigenvalues are obtained by simulating normal random samples that parallel the observed data in terms of sample size and number of variables. When this technique was put forward, a factor was considered significant if the associated eigenvalue was bigger than the mean of those obtained from the random uncorrelated data. Currently, it is recommended to use the eigenvalue that corresponds to a given percentile, such as the 95th of the distribution of eigenvalues derived from the random data (Cota, Longman, Holden, Fekken, & Xinaris, 1993; Glorfeld, 1995). Additionally, when using PA in factor analysis, the procedure is essentially the same, except that the diagonal of the correlation matrix is replaced by squared multiple correlations, which is the first step in approximating variables communalities in EFA.

Various studies indicate that PA is an appropriate method to determine the number of factors (Humphreys & Montanelli, 1975; Zwick & Velicer, 1986). Zwick & Velicer (1986) found that, among the methods analyzed, PA is the most accurate, showing the least variability and sensitivity to different factors. Glorfeld (1995) concurs with

this appraisal and states that, if the findings on the functioning of the different methods are reviewed, one would find few reasons to choose another method over PA. Similarly, several academic journals editorialize in favor of this position as well—for instance, the journal *Educational and Psychological Measurement* (Thompson & Daniel, 1996). Finally, it can be said there is sufficient agreement in considering PA the best available alternative for solving the number-of-factors-to-retain problem in EFA and PCA.

MORE ON PARALLEL ANALYSIS

Like all Monte Carlo methods, PA requires an intensive computational process. In fact, some authors have indicated that one of the main drawbacks of PA is that it can be quite costly in terms of computation time (Montanelli & Humphreys, 1976). A quick solution to this problem has been to use regression formulas which approximate the expected eigenvalues for a given observed data matrix (Lautenschlager, Lance, & Flaherty, 1989; Longman, Cota, Holden, & Fekken, 1989a; Keeling, 2000). Still another approach has been to provide tables of simulated eigenvalues, such as those published by Buja & Eyuboglu (1992). These tables incorporate estimated eigenvalues for a large range of different sample sizes and number of variables.

Even though using regression formulas or statistical tables may simplify the use of PA, we believe that both approaches are unnecessary today because of the availability and the computational power of personal computers. For example, Table 1 gives the times observed to complete PA for Principal Component Analysis (PCA) and Principal Axis Factor Analysis (FA) by using the normal data simulation approach in our new ViSta-PARAN program. Different combinations of number of variables (p) and sample size (n) are presented. We can see that, with a practical application and a standard personal computer, the computational cost of PA is not at all excessive.

Table 1: Time to complete PA under different conditions. Mean time observed in seconds for different combinations of p variables and n observations, for PCA and FA models. Estimates based on five tests of 100 replications each ran on a 950-MHz PC

n, p	PCA	FA
75, 5	2.46	3.08
150, 10	3.28	4.38
300, 20	6.08	7.30

A PA based on permutations of the data, as has been proposed by Buja & Eyuboglu (1992), can take more time. In this approach, expected eigenvalues are computed form the multivariate permutations of the observed data, rather than from simulated normal data. This method has the advantage of not needing to keep the assumption of multivariate normality of the classical PA since the null reference set is conditional on the observed data. Nonetheless, the authors also indicate that PA possesses a certain degree of robustness to handle deviations from normal assumptions, meaning that the permutation approach does not likely offer practical advantages.

Lastly, it has also been suggested that tetrachoric or polychoric correlations instead of Pearson correlations may be more appropriate

when working with item-level data in the Likert or binary format. There is evidence that PA performs well with respect to the latter case (Weng & Cheng, 2005).

Parallel Analysis and Statistical Software

Although there is consensus among statisticians that PA provides an appropriate rule for factor selection, the application of this method in psychological and social research has been limited (e.g., Fabrigar et al, 1999). In contrast, some methods that have marked limitations are used extensively by researchers. One of these methods is the Kaiser's eigenvalue-greater-than-one rule. This method is the default option in many statistical packages, which might explain its popularity to a certain extent. On the other hand, software that offers a PA option is not widely known among researchers.

Today, there are some stand-alone programs for PA (Longman, Cota, Holden, & Fekken, 1989b; Kaufman & Dunlap, 2000; Watkins, 2000) as well as some specialized macros for SPSS and SAS users (O'Connor, 2000). Of course, these stand-alone programs do not provide a full factor analysis; they only report the PA output. Also, it is important to note that in some cases they have computational restrictions. For example, the Watkins program only provides PA for the PCA case (Watkins, 2000). In the case of the SPSS and SAS macros, they can be used with well-known statistical software and in combination with a variety of factor analysis methods. But running and using a macro is not a user-friendly experience, especially for those users who are only familiar with standard graphical user interfaces (GUIs).

This paper describes a new program that has better capabilities and is easier to use than the previously published software. First, unlike some stand-alone software (see p.e. Watkins, 2000), this program is integrated into a multi-purpose statistical package called ViSta, "The Visual Statistics System" (Young, 2003). ViSta offers researchers an array of tools to edit, transform and analyze statistical data. Second, compared to the PA programs that are integrated into general statistical packages (like SPSS or SAS), our program has the following advantages: a) it is integrated into a free, non-commercial, open and expandable statistical software package; b) it works as a completely interfaced plug-in analysis, which means that PA can be performed using all of ViSta's graphical user interface tools, including a dialog box for parameter specification; and c) it includes specialized graphics to visualize the PA output, which can be used to better understand the results of the analysis. These features provide researchers with a new, user-friendly way to carry out PA.

THE VISTA-PARAN PROGRAM: AN EXAMPLE

ViSta-PARAN is a *plug-in* of ViSta "The Visual Statistics System" (Young, 1996). ViSta is a statistical system focused on exploratory data analysis and statistical visualization methods (Young, Valero-Mora, and Friendly, 2006). It is almost entirely programmed with the LispStat statistical language (Tierney, 1990). ViSta is free, open and expandable, and provides basic, intermediate and advanced statistical methods.

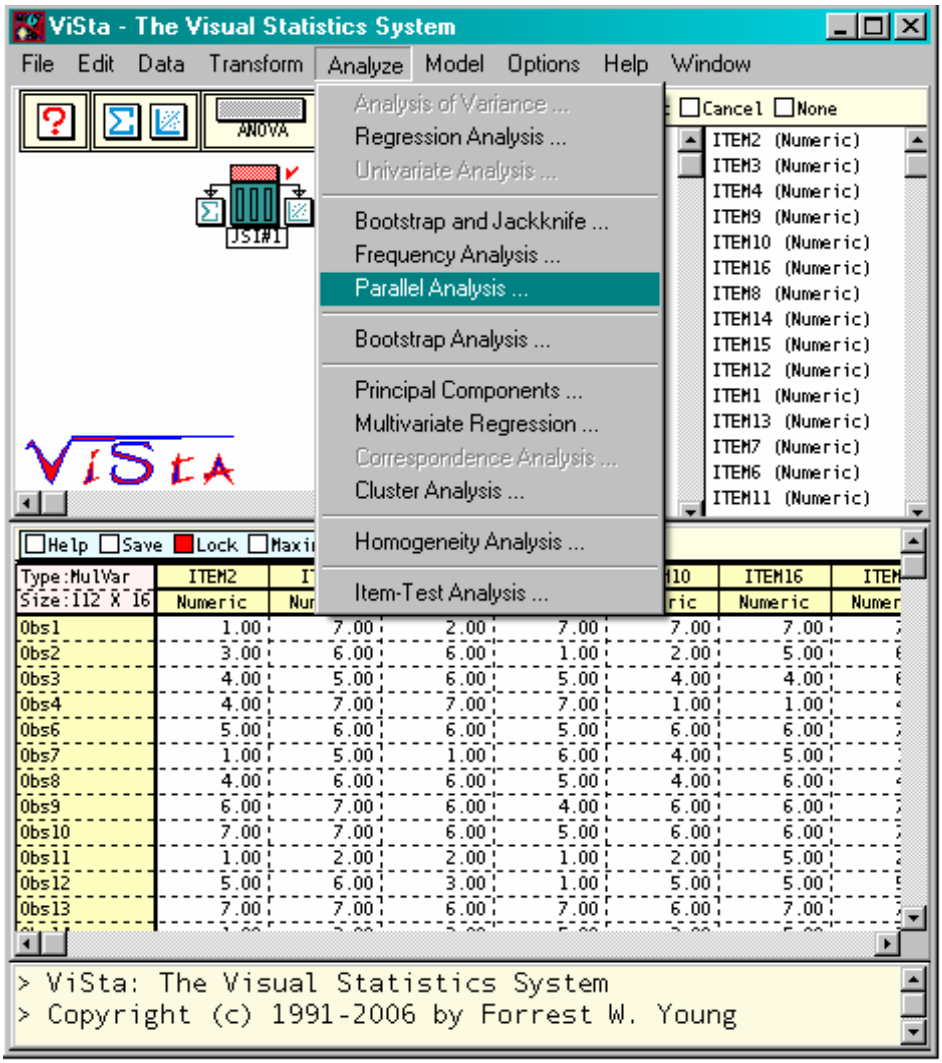
Figure 1 shows a screen shot of the ViSta program and an open data file. The example contains data from a job satisfaction study in a public health institution (Terrony, 2002). The dataset includes 16 items from a job satisfaction inventory (JSI), each with a rating based on a seven-point scale, ranging from strong agreement to strong disagreement. The data was generated from a sample of health professionals in a public hospital in Mar del Plata, Argentina. We will use this dataset to illustrate how to use the *ViSta-PARAN* program.

The *ViSta-PARAN* program may be invoked by selecting *Parallel Analysis* from the ViSta's Analysis menu options (see Figure 1). Doing so will bring up a pop-up window which prompts the user to specify the following (see Figure 2):

1) *Model*: ViSta-PARAN offers PA for two models: Principal Component Analysis and Principal Axis Factor Analysis. In the latter case, the diagonal of the correlation matrix is replaced by squared multiple correlations, which is the first step for approximating a principal axis factor solution.

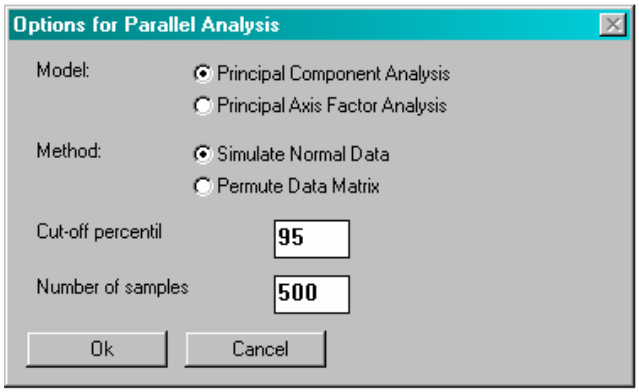
2) *Method*: ViSta-PARAN offers two PA computation methods: normal data simulation and data permutation. As we mentioned previously, the Data Permutation option leads to a multivariate permutation test that uses random permutations of the observed data instead of normal simulated data.

Figure 1: Partial view of ViSta environment with an example of a data matrix for PA.



The user can also specify the number of samples to be simulated (default=200). In cases where a random data eigenvalue is close to the observed eigenvalue, it is advisable to run the analysis again using a larger number of replications in order to obtain a more reliable solution. In Figure 2, this number has been changed to 500. The cutoff percentile can also be specified (default = 95th); this represents the point at which estimated eigenvalues are taken as a comparison baseline.

Figure 2: ViSta-PARAN dialog box for interactive model and parameter specification.



ViSta-PARAN offers two types of output: numerical and visual. Figure 3 shows the numerical report for the example. The report contains two parts: the first part displays general information about the data that have been analyzed and the parameters chosen in the analysis; the second part displays output more specifically linked to PA, including observed versus estimated eigenvalues. In our example, the first and second eigenvalues extracted from the actual data are larger than the corresponding 95th percentile random data eigenvalue. This suggests that the first two components are those to be retained for interpretation.

Figure 3: Image of the ViSta-PA's report.

```
PARALLEL ANALYSIS REPORT

MODEL: PA-JSI

VARIABLES: (ITEM2 ITEM3 ITEM4 ITEM9 ITEM10 ITEM16 ITEM8

Model: (Principal Component of correlation matrix)
Method: (Normal Data Simulation)
Number of simulated samples: 500
Eigenvalues at percentile: 95.0

              Observed       Mean    Perc95.0
 Eigenvalue1   6.13695    1.71302     1.85063
 Eigenvalue2   2.29122    1.54855     1.65371
 Eigenvalue3   1.02222    1.42764     1.50655
 Eigenvalue4   0.84595    1.32215     1.39679
 Eigenvalue5   0.83410    1.22945     1.29750
 Eigenvalue6   0.74199    1.14753     1.21075
 Eigenvalue7   0.61704    1.06960     1.12518
 Eigenvalue8   0.54557    0.99706     1.05313
 Eigenvalue9   0.51858    0.92562     0.98555
 Eigenvalue10  0.43912    0.85710     0.91177
 Eigenvalue11  0.42196    0.79260     0.84542
 Eigenvalue12  0.40146    0.72979     0.78517
 Eigenvalue13  0.38426    0.66495     0.71736
 Eigenvalue14  0.32548    0.59897     0.65972
 Eigenvalue15  0.29474    0.52772     0.59684
 Eigenvalue16  0.17935    0.44826     0.52301
```

The visual output of ViSta-PARAN is shown in Figure 4. This output consists of two graphics: a "*Scree Parallel"* plot and a "*Scree Simulation"* plot. The Scree Parallel plot graphs the observed and estimated eigenvalues, shown in Figure 4 in red (observed eigenvalues), green (95th percentile random data eigenvalues) and gray (mean of the random data eigenvalues). The point at which the

lines intersect indicates the number of factors that must be retained according to the PA criterion. This occurs at the second principal component, as was stated in the report. In a sense, the first graph is a summary of the second, which provides greater detail on the simulation process. The Scree simulation plot shows the Scree plot of the observed eigenvalues and all the Scree plots resulting from the simulated data (shown as a broad blue stripe). The latter, due to their large quantity, tend to overlap and appear as one broad blue stripe. Comparing the red line with the blue stripe, we have a visual "significance test" for each of the eigenvalues. Furthermore, the user can add various other features to this second graph, such as averages for the eigenvalues, a boxplot to explore the distribution of the simulated eigenvalues, etc. (see buttons below the graph). Taken together, the *Scree Parallel* plot and the *Scree Simulation* Plots provide a graphical analogy for both the computational process and the solution of the PA application.

Figure 4: View of the ViSta-PA's graphics: *"Scree Parallel"* plot and *"Scree Simulation"* plot

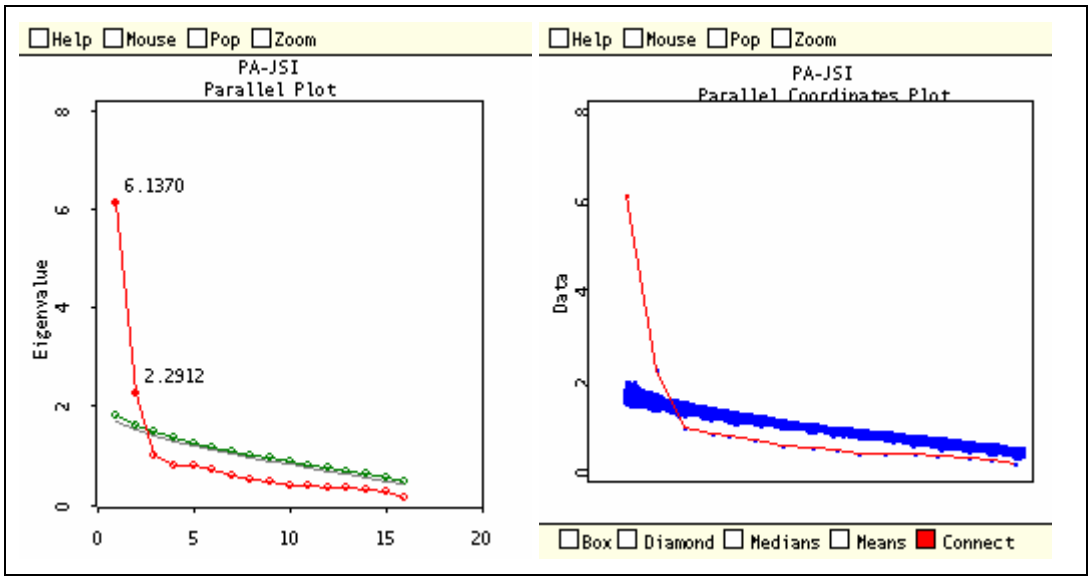


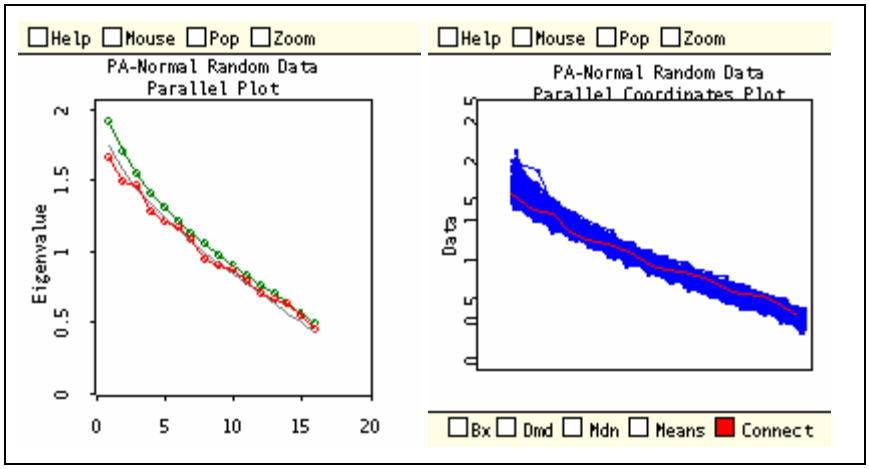
Figure 5 can help one better understand these graphs. In this case, PA has been applied to a matrix with dimensions equal to those of the real data matrix, but based on randomly generated uncorrelated normal variables. In this case, the expectation is that no eigenvalues will supersede the Pa criteria, which can be clearly seen in the graph. In the *Scree Parallel*, it can be seen that the observed eigenvalues are all found below the cutoff line estimated using the simulated data. While in the *Scree Simulation*, we can see that the red line lies within the blue stripe of expected eigenvalues.

FINAL REMARKS

PA is one of the most recommendable rules for factor selection in EFA and PCA, but it is not available as an analysis option in the most commonly used statistical software packages. In view of that, this paper describes a new, interactive and easy-to-use computer program capable of carrying out PA – the *ViSta-PARAN* program.

ViSta-PARAN has been developed with the LipStat statistical programming language. It is integrated as an analysis module in ViSta *"The Visual Statistics System,"* a free, open and expandable statistical system focused on multivariate statistical visualization methods. *ViSta-PARAN* provides parametric and non-parametric PA for Principal Component and Principal Axis factor analysis. The results can be easily interpreted by using reports or specialized graphics.

Figure 5: Example of *"Scree Parallel"* plot and *"Scree Simulation"* plot for a data matrix of random uncorrelated normal variables.



We hope this new tool increases the use of valid rules to deal with the number-of-factors-to-retain problem. Additionally, this program is a useful tool for educational purposes, as a way of introducing the PA method in statistics courses. The *ViSta-PARAN* program may be obtained without cost at: http://www.mdp.edu.ar/psicologia/vista/. It runs on Windows 95 or later and requires ViSta 6.4 (http://forrest.psych.unc.edu/research/index.html) or ViSta 7 (http://www.uv.es/visualstats/Book/)

REFERENCES

Buja, A. & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research, 27,* 509-540.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1,* 245-276.

Costello, A. B. & Osborne, J. W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation, 10(7).* Available online: http://pareonline.net/getvn.asp?v=10&n=7

Cota, A.A., Longman, R.S., Holden, R.R., Fekken, G.C., & Xinaris, S. (1993). Interpolating 95th percentile eigenvalues from random data: An empirical example. *Educational & Psychological Measurement, 53,* 585-596.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *3*, 272-299.

Glorfeld, L.W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement, 55*, 377-393.

Hayton, J.C., Allen, D.G. & Scarpello,V. (2004) Factor Retention Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis. *Organizational Research Methods, 7*, 191-205

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179-185.

Horn, J. L. and Engstrom, R. (1979). Cattell's scree test in relation to Bartlett's chi-square test and other observations on the number of factors problem. *Multivariate Behavioral Research*, *14*, 283-300.

Hubbard, R., & Allen, S.J. (1987). An empirical comparison of alternative methods for Principal Component Extraction. *Journal of Business Research, 15,* 173-190.

Humphreys, L. G. & Montanelli, R. G. (1975). An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behavioral Research*, *10*, 193-206.

Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141-151.

Kaufman, J.D. & Dunlap, W.P. (2000). Determining the number of factors to retain: A Windows-based FORTRAN-IMSL program for parallel analysis. *Behavior Research Methods, Instruments, & Computers, 32,* 389-395.

Keeling, K. B (2000). A Regression Equation for Determining the Dimensionality of Data. *Multivariate Behavioral Research, 35*, 457-468.

Lautenschlager, G.J., Lance, C.E., & Flaherty, V.L. (1989a). Parallel analysis criteria: Revised regression equations for estimating the latent roots of random data correlation matrices. *Educational and Psychological Measurement, 49,* 339-345.

Lautenschlager, G. J. (1989). A comparison of alternatives to conducting Monte Carlo analyses for determining parallel analysis criteria. *Multivariate Behavioral Research*, *24*, 365-395.

Longman, R.S., Cota, A.A., Holden, R.R. & Fekken, G.C. (1989a). A regression equation for the parallel analysis criterion in principal components analysis: Mean and 95th percentile eigenvalues. *Multivariate Behavioral Research, 24,* 59-69.

Longman, R.S., Cota, A.A., Holden, R.R. & Fekken, G.C. (1989b). PAM: A double precision FORTRAN routine for the parallel analysis method in principal components analysis. *Behavior Research Methods, Instruments, & Computers, 21,* 477-480.

Montanelli, R. G. & Humphreys, L. G. (1976). Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: A Monte Carlo study. *Psychometrika*, *41*, 341-348.

O'Connor, B.P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32,* 396-402.

Peres-Neto, P., Jackson, D., & Somers, K. (2005) How many principal components? stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics & Data Analysis 49,* 974 – 997

Terroni, N. (2002). *El nivel de satisfacción de los trabajadores de una institución de salud pública. Su relación con el ausentismo y la salud psicofísica.* Master Thesis in Social Psychology, Facultad de Psicología, Universidad Nacional de Mar del Plata.

Thompson, B. and Daniel, L. G. (1996) Factor Analytic Evidence for the Construct Validity of Scores: A Historical Overview and Some Guidelines. *Educational and Psychological Measurement, 56*, 2, 197 - 208

Tierney, L. (1990). *Lisp-Stat An Object-Oriented Environment for Statistical Computing and Dynamic Graphics*. New York: John Wiley & Sons.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41, 321-327

Weng, L. J. & Cheng, Ch. P. (2005) Parallel Analysis with Unidimensional Binary Data. *Educational and Psychological Measurement, 65,* 697-716

Watkins, M.W. (2000). *Monte Carlo PCA for parallel analysis* [computer software]. State College, PA: Ed & Psych Associates.

Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under- and overextraction on principal axis factor analysis with varimax rotation. *Psychological Methods*, *1*, 354-365.

Young, F.W. (2003). *ViSta "The Visual Statistics System".* [computer software] [on-line] Available: URL: http://forrest.psych.unc.edu/research/index.html

Young, F.W. (1996). *ViSta: Developing Statistical Objects.* The Visual Statistic project, Thurstone Psychometric Lab, Univ. N. Carolina, Chapel Hill, Research Memorandum Number 1996-1.

Young, F.W., Valero-Mora, P. & Friendly, M. (2006). *Visual Statistics Seeing Data with Dynamic Interactive Graphics.* NJ: Wiley and Sons.

Zwick, W.R. & Velicer, W.F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432-442.

Citation

Ledesma, Rubén Daniel and Pedro Valero-Mora (2007). Determining the Number of Factors to Retain in EFA: An easy-to-use computer program for carrying out Parallel Analysis. *Practical Assessment Research & Evaluation*, 12(2). Available online: http://pareonline.net/getvn.asp?v=12&n=2

Authors

Correspondence concerning this paper should be addressed to

Rubén Daniel Ledesma
Río Negro 3922
Mar del Plata (7600), Argentina

rdledesma@gmail.com

tel/fax 54 223 4752266