PARALLEL ANALYSIS

OVERVIEW

Parallel analysis (PA) is a method used to determine the number of components or factors to retain in Principal Component Analysis (PCA) and Exploratory Factor Analysis (EFA). Simulation studies have indicated that PA is the best technique among others such as the Scree plot of eigenvalues or Kaiser’s eigenvalue-greater-than-one rule for determining the number of components or factors. PA involves the following steps: (a) observed eigenvalues are extracted from the original data set; (b) simulated random data sets are generated that are parallel in structure to the observed data in terms of sample size and the number of variables; (c) the eigenvalues are extracted from the parallel datasets, and the mean and 95th percentile of the distribution of eigenvalues are determined; (d) a factor is retained if its observed eigenvalue is larger than the mean (or 95th percentile) of the corresponding eigenvalue from the parallel data sets. The motivation behind this approach is that the variance extracted from the real data should reasonably be greater than variance extracted from random data sets.

ARTICLES AND CHAPTERS

Description of Method


Provides an overview of PA, reviews the articles utilizing PCA or Factor Analysis (FA) from 1987 to 1993 in Ecology fields, and applies PA to published analyses by using SAS.


Provides an introduction of PA and shows that the distributions of eigenvalues generated from PA are insensitive to the underlying distribution of data used to simulate the random data. The author suggests retaining factors whose observed eigenvalues were larger than the 95th or 99th percentile of the distribution of corresponding expected eigenvalues.

Provides an introduction of two techniques of deciding the number of factors—PA and Velicer’s minimum average partial (MAP) test and provides SPSS and SAS syntax for conducting PA and MAP test.


Provides an overview of factor retention criteria (e.g., Kaiser’s minimum eigenvalue greater than 1 criterion, Cattell’s scree test, Bartlett’s chi-square test) and step by step tutorial of conducting PA by using SPSS.


Provides a practical introduction to using R for latent variable modeling analyses, including EFA and PA (Chapter 2). The author presents step-by-step guidance and provides syntax for conducting PA in R.

Applications


Uses PA as a technique to decide the number of factors in the subsequent EFA and CFA for three depression measures.


Uses PA as one of methods to decide how many components and factors to retain for the Integration of Stressful Life Experiences Scale.


Uses PA as one of techniques to determine the optimal number of factors for BASC–2 Behavioral and Emotional Screening System Student Form.


Uses PA to decide the number of factors extracted for the subsequent EFA for items of Wrist and Hand upper-extremity Fugl-Meyer (W/H UE FM).
Special Issues and Extensions


Liu, O. L., & Rijmen, F. (2008). A modified procedure for parallel analysis of ordered categorical data. *Behavior Research Methods, 40*(2), 556-562. Provides a modified procedure for PA to analyze ordinal data and deal with missing data. The SAS syntax for implementing this modified procedure is also provided.


SOFTWARE

R packages

paran: Performs parallel analysis and provides numerical and graphical evaluation of components or factors.
  • link to pdf

random.polychor.pa: Performs parallel analysis with polychoric correlation matrix.
  • link to pdf

fa.parallel: Performs parallel analysis and provides scree plot of eigenvalues of components or factors. Parallel analysis with polychoric or tetrachoric correlation could be specified.
  • link to pdf

Other statistical packages

Mplus syntax for parallel analysis

SPSS, SAS, MATLAB, and R Programs
Webpage provides updates in MATLAB and R for O’Connor’s 2000 paper.
  • link

SAS macro for parallel analysis
  • link to pdf

STATA
Resources from the Institute for Digital Research and Education (IDRE), Statistical Consulting Group, UCLA)
  • link

SELECTED INTERNET RESOURCES TO GET YOU STARTED

YouTube demonstration
Demonstration of PA using Brian O'Connor's SPSS syntax - (Youtube-www/how2stats.net).
  • YouTube link