Ferrando, Pere J.; Masip-Cabrera, Antoni; Navarro-González, David; Lorenzo-Seva,
Urbano
The Psychometric Toolbox: an Excel package for use in measurement and psychometrics
courses

The Psychometric Toolbox: an Excel package for use in measurement and psychometrics courses

Pere J. Ferrando[*], Antoni Masip-Cabrera, David Navarro-González, and Urbano Lorenzo-Seva

Research Centre for Behavioral Assessment

'Universitat Rovira i Virgili', Spain

The Psychometric Toolbox (PT) is a user-friendly, non-commercial package mainly intended to be used for instructional purposes in introductory courses of educational and psychological measurement, psychometrics and statistics. The PT package is organized in six separate modules or sub-programs: Data preprocessor (descriptive analyses and data transformations); Guttman scaling (item analysis and individual scoring based on Guttman's model); Classical Item Analysis (CTT-based item analysis and reliability estimation); Item Factor Analysis (FA-Based Item Analysis and omega reliability estimation); Scoring and Norming (linear and nonlinear score transformations and normative tables); and Item Response Theory-Basic analysis (item calibration and test scoring based on the one and two parameter logistic models). Because they have been designed for instructional use, these modules are intended to (a) be very easy to use, (b) provide clear and well explained results, and (c) make use of graphical displays whenever possible.

The Psychometric Toolbox (PT) is designed as a user-friendly, non-commercial package intended to be used for instructional purposes in introductory courses of educational and psychological measurement, psychometrics and statistics. While excellent programs of this type already exist – for example, TAP (Brooks & Johanson, 2003), which focuses on

classical test theory, and IRT-Lab (Penfield, 2003), which specializes in Item Response Theory – the main contributions of PT are its wide scope and versatility, which make it possible to carry out Guttman analysis, classical item and test analysis, factor analysis, and IRT analysis. Furthermore, a variety of procedures for scoring and norming are available. Finally, PT includes a powerful data preprocessor which allows many statistical treatments, sampling procedures, and data transformations to be performed.

Program description

The PT package is organized in six separate modules: (1) Data preprocessing, (2) Guttman scaling, (3) Classical item analysis, (4) Item factor analysis, (5) Scoring and Norming, and (6) IRT analysis. Because they have been designed for instructional use, these modules are intended (a) to be very easy to use, (b) to provide clear results, and (c) to make extensive use of graphical displays. As well as the basic functions expected in programs of this type, some not-so-usual features are now provided as example. The preprocessor module allows the user to replace missing values, perform split-half test partitions, randomly split the sample for cross-validation purposes, perform Jacknife re-sampling, and create automatically parallel forms. The item factor analysis computes the omega coefficient with its confidence intervals and displays one and two canonical solutions in order to detect violations to unidimensionality (e.g. Ferrando & Lorenzo-Seva, 2013). The Scoring and Norming program provides confidence intervals for the transformed scores and creates the test normative table on the basis of the user's specifications.

PT is a collection of Excel macros and does not need a specific installation (as long as Microsoft Excel is already installed in the computer). We developed them in Microsoft Office Excel 2010, and the programming language used is Visual Basic of Applications. The functions implemented are made available to the user in a specific ribbon, and the contents of the document must be enabled by the user if the implemented macros are to be executed. We check our modules before publishing them on the web page, so that we are sure that they are free of any malicious code. With the datasets that we have analyzed so far, we have found no limitations of disk or memory space. With regards to the accuracy and correctness of the PT results, we have checked them to those which can also be obtained from commercial programs (SPSS-descriptives, SPSS-reliability, SPSS-factor, and BILOG-MG3) and found all of them to be accurate at least to the second decimal place.

Limitations and future extensions

At present, the PT package is only intended for binary and graded-response items analyzed on a scale-by-scale basis and fitted with dominance models. In the near future, we plan to extend the package capabilities so that (a) other items formats (mainly multiple-choice) could be assessed, and (b) other models, including multidimensional models could be also fitted. Improvements in the functioning of the existing modules as well as the output they provide are also on the way.

Availability and accompanying documentation

The program and accompanying material are available free of charge from the website:

http://psico.fcep.urv.cat/utilitats/PsychometricTools/index.html.

The documentation consists of: (a) a summary of the functioning of each module with the procedures available, (b) a complete user's guide, which is intended for instructors and includes a theoretical background and detailed step-by-step analyses of empirical data, and (c) the data-files used as examples in the guide.

RESUMEN

Psychometric Toolbox: un paquete Excel diseñado para cursos en medida y psicometría. El "Psychometric Toolbox" (PT) es un paquete no comercial de fácil utilización, diseñado principalmente para ser utilizado con fines educativos en cursos introductorios de medición educativa y psicológica, psicometría y estadística. El paquete PT está organizado en seis módulos o subprogramas separados: Data preprocessor (análisis descriptivos y transformaciones de datos); Guttman scaling (análisis de ítems y puntuación individual basada en el modelo de Guttman); Classical Item Analysis (análisis de ítems basados en la TCT y estimación de la fiabilidad); Item Factor Analysis (Análisis factorial de ítems y estimación del coeficiente de fiabilidad omega); Scoring and Norming (transformaciones lineales y no lineales y elaboración de tablas normativas); y Item Response Theory-Basic analysis (calibración de ítems y puntaciones basadas en los modelos logísticos de uno y dos parámetros). Debido a que han sido diseñados para uso educativo, estos módulos han sido desarrollados para (a) ser muy fáciles de usar, (b) proporcionar resultados claros y bien explicados, y (c) hacer uso de presentaciones gráficas siempre que sea posible.

REFERENCES

Brooks, G. P., & Johanson, G. A. (2003). TAP: Test Analysis Program. *Applied Psychological Measurement*, *27*, 303-304.

Ferrando, P. J., & Lorenzo-Seva, U. (2013). *Unrestricted item factor analysis and some relations with item response theory*. *Technical Report*. Department of Psychology, Universitat Rovira i Virgili, Tarragona. *http://psico.fcep.urv.es/utilitats/factor*.

Penfield, R. D. (2003). IRT-Lab: Software for research and pedagogy in item response theory. *Applied Psychological Measurement*, *27*, 301-302.