List of Statistical Software Available at Warwick

Lisa Lavender, Simon French and Andrew Mead

(Contact: l.lavender@warwick.ac.uk)

Introduction

This list of statistical software has been compiled to help researchers identify statistical tools that may help in their projects. It is a live document and we would be grateful of updates, both advising of new software availability or change in licences and of the removal of any from Warwick systems[1].

It should be recognised that many of the systems share functionality and the use of any alternative system is a matter of preference. Thus we do not offer any opinions on the quality or relative value of different systems.

The full list of software licences at the University can be found at:
http://www2.warwick.ac.uk/services/its/servicessupport/software/list/

Notes on the following table

1. The list of software below was compiled during April 2013. It will be updated in the light of *ad hoc* comments as they are received.

2. We have grouped the software packages according to whether general statistical functionality is provided or whether the focus is on more specific uses. We recognise that this is a subjective classification and others, particularly the software developers, might classify the packages differently.

3. Statistical software packages often find favour in particular disciplines because they have been or were originally written for work in a particular discipline. SPSS is an obvious case in point. The initials derive from *Statistical Package for the Social Sciences*, but for many years have been known simply as SPSS. However, since it is a general purpose package, it may be used elsewhere – one of us has used its clustering algorithms in analysing software for atmospheric dispersion! Nonetheless it is still used mainly in the social sciences. Because of this, we indicate the history and origins of packages as we know them to indicate fit with disciplines.

4. There is a useful list of software on Wikipedia: see
http://en.wikipedia.org/wiki/Statistical_software.

---

[1]

General Purpose Statistical Packages

Broadly, these packages offer the user and easy way of applying standard estimation, confidence interval and hypothesis testing techniques to general data, usually input in the form of a spreadsheet.

Software	Functionality	Background	Licence availability restrictions links to distributor
Genstat	The package provides access to a wide range of statistical procedures, with the design and analysis of experiments a major strength, and facilities also available for statistical modelling, time series, spatial analysis and many other specialist areas. The menu-based interface enables inexperienced users to access most of these facilities a select from various sensible options, while more experienced users can use the underlying programming language to develop more complex analyses. Includes a good graphical capability and easy transfer of data to and from other packages, including Excel. A focus of the development team is in encouraging "good statistical practise", rather than the simple application of statistical methods!	GenStat was developed within the agricultural research community in the UK and Australasia but is now much more widely used in a range of research environments, and is increasingly being used in academia. An initial focus on designed experiments has now expanded to cover most main-stream statistical methods. A recent innovation is the "Teaching and Learning" version that provides a "cut-down version of the software for use in schools and taught university courses.	A multi-user license for the full version of GenStat is currently purchased by the School of Life Sciences, with stand-alone licenses owned elsewhere within the University. A license for the "Teaching and Learning" version was purchased (no cost) by the Faculty of Science. Access is available through the IT group in the School of Life Sciences. This version is only avialble to be used for taught course students, and should not be used for research applications. GenStat is marketed by VSN International (http://www.vsni.co.uk/) . their website contains a range of additional resources including access to taught courses using the software, and for teaching.
Minitab Statistical analysis software	The package covers most statistical procedures to some extent, with a strong emphasis on the design and analysis of experiments, and on statistical process control, both with a focus on industrial applications. Because perhaps of its pedigree as a tool for teaching statistics, its help functionality is particularly strong, offering help in recognising the appropriateness of particular statistical methods to your data.	Minitab was developed at Penn State University and has been one of the standard packages for teaching statistics for over three decades. It has also developed into a reasonably full general purpose statistical package that may be used in research.	Site licence purchased by IT Services. Further information and download details: http://www2.warwick.ac.uk/services/its/servicessupport/software/list/minitab Minitab is marketed by Minitab Inc (http://www.minitab.com/). Their website contains a range of teaching resources, some chargeable, including a list of text-books which reference Minitab, many providing examples in Minitab code.

Software	Functionality	Background	Licence availability restrictions links to distributor
Origin Data analysis and graphing software application.	The package provides some basic statistical functionality, though this range appears to be extending with every release – the most recent version includes a range of tools for signal and image processing, but is primarily a graphical package offering a range of high quality plots and graphs of publication quality. It is strong on curve fitting.	Origin's background lies in the development of statistical graphics. Its purpose was originally to provide high quality statistical graphics for publication. Statistical functionality, initially relating to curve fitting and more recently expanding into other areas, has developed since.	Available to university members only. Further information and download details: http://www2.warwick.ac.uk/services/its/servicessupport/software/list/origin. Origin is marketed by OriginLab (http://www.originlab.com/). The site has some support including a set of video tutorials.
R	R is at its heart a statistical programming and data manipulation language, capable of being used to perform just about any statistical analysis. It is also very strong on graphics. R is used extensively across academic statistics for developing new statistical methodology. Being an open project, there are an enormous number of add-ins and packages available to perform specific analyses. Some are very well tested and validated, but others are there purely for others to test and experiment with. So it is wise to check the pedigree of any add-in that you use for analysis. Some of the add-ins provide graphical interfaces making particular analyses or suites of analyses accessible to inexperienced users (e.g. R Commander and Rattle), but the R package itself is designed for professional statisticians and researchers. The R Studio package provides a slightly more friendly interface for less-experienced users, and is certainly a good environment in which to teach the use of R.	'S' was developed as a statistical programming language by J.M. Chambers at Bell AT&T Labs. As a language it allows any statistical analysis to be written and run on virtually any data. 'R' is a free implementation of a dialect of S. (S-Plus is a commercial implementation: see entry on S-Plus). As a matter of local history, in 2007 Prof David Firth and Dr Heather Turner of Warwick Statistics won the John M Chambers Statistical Software Award for their Generalized Nonlinear Models (gnm) R package.	*Note that this is public domain software and, while there is reason to be confident in R itself, many add-ins are not tested and validated to commercial standards.* The R website with links for downloads, manuals, FAQs, book lists, etc. is at: http://www.r-project.org/ A library of contributed add-in packages is maintained (under the menu item packages) at: http://cran.r-project.org/. R Studio can be downloaded from http://www.rstudio.com/

Software	Functionality	Background	Licence availability restrictions links to distributor
SAS Statistical analysis software	SAS is an enormous suite of tools and programs, supported by a fourth generation programming language and advanced database and data warehousing tools. It provides advanced statistical tools which, together with its programming language, allow a professional statistician to develop and conduct virtually any analysis. It also provides very advanced data mining tools. Indeed, it is one of the suites of tools that are driving the analysis of 'big data' within the public and private sectors, lying at the heart of the analytics movement. In short, there is very little that the full SAS suite cannot do, but it has a steep learning curve.	SAS has its origins in agricultural statistics programs developed at North Carolina State University in the 1960s. Set up as a private company, the SAS Institute is now one of the world's biggest providers of statistical and operational research software with profits nearing $3billion. Its software is used throughout the (commercial) research community, business and industry. The USA Food and Drug Administration (FDA) and a number of other research funders and regulatory bodies often require the use of SAS data analysis programs within their protocols	The university subscribes to the Academic Analysis suite and "SAS/OR" – an advanced analytics product. Made available through a CHEST site licence, including home use rights for staff and students. This software is for **research only.** Now available as a free download from ITS: http://www2.warwick.ac.uk/services/its/servicessupport/software/list/sas The SAS Institute website (http://www.sas.com/) provides a range of support, but tending to supporting advanced SAS users rather than beginners
S-Plus Statistical analysis software	S-Plus is a general purpose statistical package and statistical programming environment. Standard analyses are available as pre-programmed functionality. It is a very powerful statistical package which can provide virtually any analysis.	S was developed as a statistical programming language by J.M. Chambers at Bell AT&T Labs. As a language it allows any statistical analysis to be written and run on virtually any data. S-Plus is a commercial implementation of S. (R is a free implementation of a dialect of S: see entry on R). S-Plus has been developed and owned by a series of companies over the years and is now owned by Tibco, a company which focuses on business need for Analytics and data mining software. Currently S-Plus is being absorbed into Tibco's Spotfire package.	A site licence with Insightful allows home use by staff and students. Available as a free download from ITS: http://www2.warwick.ac.uk/services/its/servicessupport/software/list/splus Spotfire S+ is marketed through the following website: http://spotfire.tibco.com/
SPSS Statistical analysis software.	SPSS is a general purpose statistical package which provides a wide range of functionality, being particularly strong in descriptive statistics, regression, and multivariate analysis. There is a structural equation modelling add-in: see entry on AMOS below. There is a user friendly graphical interface and a very extensive help facility which not only describes the package but provides much useful material on the statistical methods themselves.	SPSS originally stood for statistical package for the social science. It was developed in the 1960s and has evolved over the years into a fully-fledged package that allows a very broad range of statistical analyses, across a wide range of disciplines, while it remains the most widely used package within the social sciences. It is now owned and sold but IBM.	Site licence that extends to home use for staff and students. Now available as a free download from ITS: http://www2.warwick.ac.uk/services/its/servicessupport/software/list/spss There are many textbooks which provide introductions to the use of SPSS within particular disciplines. Amazon lists a huge number.

Statistical Packages Designed for a Specific Set of Functions

These packages are designed to provide a limited set of functions, but often more advanced ones than are found in general purpose statistical packages: e.g. time series analysis or Bayesian modelling.

Software	Functionality	Background	Licence availability restrictions links to distributor
AMOS Structural Equation Modelling Package	AMOS provides structural equation modelling within the SPSS environment. This enables the user to specify, estimate, assess and present models to showing hypothesized relationships among variables. Users can choose either the graphical user interface or non-graphical, programmatic interface.	See entry on SPSS	Site licence that extends to home use for staff and students. Now available as a free download from ITS: http://www2.warwick.ac.uk/services/its/servicessupport/software/list/spss
BUGS Bayesian Analysis	BUGS offers flexible software for the Bayesian analysis of complex statistical models using Markov chain Monte Carlo (MCMC) methods. It can be tailored to do a wide range of analyses, but its strength lies in Bayesian regression, hierarchical models and spatio-temporal modelling. A somewhat old introduction to the package is contained in: Lunn, D. J., Thomas, A., Best, N. and Spiegelhalter, D. (2000). "WinBUGS -- a Bayesian modelling framework: concepts, structure, and extensibility." *Statistics and Computing* **10**: 325--337. A more recent and extensive introduction is available in: Lunn, D. J., Jackson, C., Best, N., Thomas, A. and Spiegelhalter, D. (2013). *The BUGS Book: a Practical Introduction to Bayesian Analysis*. London, Chapman and Hall. An Excel front-end to BUGS is available: BugsXLA. See P. Woodward (2012). *Bayesian Analysis Made Simple: An Excel GUI for WinBUGS*. London, Chapman and Hall.	BUGS stands for Bayesian Updating by Gibbs Sampling. BUGS began as a project in Cambridge at the MRC Biostatistics Laboratory in the late 1980s. It is an open software project. Currently development is focused on OpenBUGS, which can be run as BRUGS within R (see entry on R).	*Note that this is public domain software and, while there is reason to be confident in BUGS itself, many add-ins are not tested and validated to commercial standards.* WinBUGs is available at: http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml. BRUGS may be downloaded at: http://cran.r-project.org/web/packages/BRugs/index.html. OpenBUGS may be downloaded at: http://www.openbugs.info/w/. RB The Excel GUI BugsXLA may be downloaded from: http://bugsxla.philwoodward.co.uk/

Software	Functionality	Background	Licence availability restrictions links to distributor
NVivo Qualitative data analysis software	NVivo supports a wide range of qualitiative research methodologies and accepts data in a wide variety of formats: e.g. audio files, videos, digital photos, Word, PDF, spreadsheets, rich text, plain text and web and social media data. It includes tools for network and organizational analysis, action or evidence-based research, discourse analysis, grounded theory, conversation analysis, ethnography, literature reviews, phenomenology, and mixed methods research.	NVivo has its origins in the Nudist software package developed in the early 1980s. Nudist stands for Non-Numerical Unstructured Data Indexing Searching and Theorizing and was one of the first programs to support the management and analysis of unstructured qualitative data. NVivo has taken over and extended the functionality of Nudist and is the flagship software of QSR International.	Supplied by QSR International through a CHEST site licence including home use by both staff and students. Can be downloaded by University members only from: http://www2.warwick.ac.uk/services/its/servicessupport/software/nvivo. NVivo is marketed though QSR's website: http://www.qsrinternational.com/products_nvivo.aspx, where there is a range of support material including video tutorials.
Qualtrics Survey software	Qualtrics provides a means of developing and managing questionnaire surveys via the web. The questionnaires are easily constructed and deployed and a variety of formats of questions are possible. The results may be analysed simply withing Qualtrics with graphical output in a variety of formats. Alternatively data may be output as a spreadsheet or other format for further or more advanced analysis in another package, commonly SPSS.	Qualtrics is a private company founded about 10 years ago. Their software is the market leader in academic business research.	Site licence purchased by WBS, for use by staff and students in WBS only. Further details: http://www2.warwick.ac.uk/fac/soc/wbs/central/issu/help/kb/software/non_microsoft_apps/qualtrics The company website is at: www.qualtrics.com. Note that the company offer a free account to create simple questionnaires.

General Purpose Packages which have some Statistical Functionality

These packages are *not* basically designed for statistical analysis, but all have some statistical functionality and some, being mathematical tools, pretty much allow you to write any statistical analysis from scratch. But we would generally advise using packages that were designed primarily for statistical analysis since their functionality will generally have been tested and verified for statistical purposes. If you write your own calculations in a mathematical calculation tool, you risk miscoding the statistical technique and drawing poor or completely wrong inferences.

Software	Functionality	Background	Licence availability restrictions links to distributor
ArcGIS	ArcGIS is a software package focussed on the management, manipulation and visualisation of geographically referenced data. It includes a range of statistical analysis routines that enable the summary of geographically referenced data. there are a number of add-on packages that provide enhanced geostatistical functionality.	ArcGIS was developed by ESRI (the Environmental Systems Research Institute) in 1997 from a range of earlier packages focussed on different aspects of the management, manipulation and visualisation of geographically referenced data.	Supplied by ESRI via a CHEST site license, allowing staff and students to use on University and personal machines for University-related work. Runs on Windows only. Further information is available at http://www2.warwick.ac.uk/services/its/servicessupport/software/gis
Excel	Excel provides a number of basic statistical functions and analyses. There are add-in analytic tools packs, both published by Microsoft and supplied with the Excel /MS Office Installation and also supplied by third parties (including the StatPlus package that is the best method of accessing such statistical routines for Excel running on a Mac). Over the years there have been several warnings that some of the inbuilt functions and analyses have flaws, but these have died down and the latest versions of excel do provide an easy means of conducting simple statistical analyses. The various chart formats and graphical tools are very useful too. It should also be remembered that Excel is an excellent tool for managing the data before inputting into other statistical packages.	Microsoft Office and Excel have a long and well known history. They are the market leaders.	Downloads and licence details are at: http://www2.warwick.ac.uk/services/its/servicessupport/software/microsoft A free (limited functionality) version of StatPlus is available (for both Windows and Mac environments) from http://www.analystsoft.com/en/products/statplus/ (Windows) or http://www.analystsoft.com/en/products/statplusmacle/ (Mac). A full version is also available to be purchased from associated web-sites.
MATLAB Mathematical software	Matlab is a general purpose mathematical analysis and calculation package. It supports the development and evaluation of virtually any mathematical calculation, including, of course, statistical calculations. The matrix manipulation functions allow many statistical analyses to be developed from 'first principles', but the Statistics Toolbox provides access to a range of purpose-built functions for standard statistical analyses, including EDA, regression modelling, analysis of variance, multivariate statistics, design of experiments and statistical process control.	Matlab's development extends back to the 1970s. It is a tool that has been used extensively across mathematics and the physical sciences and is arguably the market leader for a general purpose mathematical calculation tool.	Supplied by the MathWorks through a total Academic Headcount Agreement which allows staff and student use at home and on University-owned machines. Further information and download details: http://www2.warwick.ac.uk/services/its/servicessupport/software/matlab

Software	Functionality	Background	Licence availability restrictions links to distributor
Mathematica includes data manipulation tools	Mathematica is a general purpose mathematical analysis and calculation package. It supports the development and evaluation of virtually any mathematical calculation, including, of course, statistical calculations. However, it is not a purpose-built statistical package. Most statistical analyses can be conducted 'from first principles'. Includes some built-in functions for multivariate statistical approaches, and claims to allow calculations of probabilities and expectations for over 100 different forms of distribution.	Mathematica was developed by Stephen Wolfram who had a vision of developing a mathematical calculation engine embedded in a live notebook interface. Essentially the user writes a notebook (think of a research paper or a student text), but the mathematics, graphics and data in the text are live. Change any element and the notebook recalculates.	Supplied by Wolfram Research. Site licence allows staff and students use at home and on University-owned machines. The software can be downloaded from the Warwick tree, with further details at: http://www2.warwick.ac.uk/services/its/servicessupport/software/list/mathematica. Mathematica is marketed by Wolfram Research: http://www.wolfram.com/. The website contains many support and training tools and a list of several textbooks that introduce Mathematica to a range of audiences and disciplines.
NAG Numerical and statistical components library	NAG is essentially a software library of numerical and mathermatical algorithms that can be called in from user-written programs. It contains some statistical functionality, but is much more a general purpose suite. If a user wishes to write a special purpose program which contains and supports a specific statistical analysis then NAG would provide a route forward. But for general statistical analyses. The use of either a general purpose statistical package or writing a specific analysis in R or S-Plus would be a more usual approach.	NAG originally stood for *Nottingham Algorithms Group and its history extends back into the late 1960s*. It is now known as and marketed by the *Numerical Algorithms Group*.	Site licence covers use by staff and students on University and home machines. Further information and download details: http://www2.warwick.ac.uk/services/its/servicessupport/software/list/nag
Oracle database object-relational database management system	Oracle is the market leading database system and as such has a number of statistical functions built into its reporting and querying tools. However, for the vast majority of statistical analyses it would not be the tool of choice. If complex database functions are needed to support a statistical investigation, then the SAS suite (see above) has both a powerful database system and an advanced analytic toolkit.	Oracle as a company was founded in 1977. Its database system is the current market leader, though it is perceived more as an enterprise system offering companies a host of functionalities from web to process management and e-commerce, all built around a very powerful database management system.	IT Services has a site licence for Oracle database enterprise. Please contact the ITS Service Desk to discuss your requirements.