

Stat 6304

Computational Statistics

Statistical Packages

Some Statistical Packages

- BMDP
- GLIM
- HIL
- JMP
- LISREL
- MATLAB
- MINITAB

Some Statistical Packages

- R
- S-PLUS
- SAS
- SPSS
- STATA
- STATISTICA
- STATXACT
- ... and many more

- BMDP is a comprehensive library of statistical routines from simple data description to advanced multivariate analysis, and is backed by extensive documentation.
- Each individual BMDP sub-program is based on the most competitive algorithms available and has been rigorously field-tested.
- BMDP has been known for the quality of its programs such as Survival Analysis, Logistic Regression, Time Series, ANOVA and many more.
- The BMDP vendor was purchased by SPSS Inc. of Chicago in 1995. SPSS Inc. has stopped all development work on BMDP, choosing to incorporate some of its capabilities into other products, primarily SY-STAT, instead of providing further updates to the BMDP product.
- BMDP is now developed by *Statistical Solutions* and the latest version (BMDP 2009) features a new modern user-interface with all the statistical functionality of the classic program, running in the latest MS Windows environments.

- LISREL is software for confirmatory factor analysis and structural equation modeling.
- LISREL is particularly designed to accommodate models that include latent variables, measurement errors in both dependent and independent variables, reciprocal causation, simultaneity, and interdependence.
- Vendor information: Scientific Software International
<http://www.ssicentral.com/>

- Matlab is an interactive, matrix-based language for technical computing, which allows easy implementation of statistical algorithms and numerical simulations.
- Highlights of Matlab include the number of toolboxes (collections of programs to address specific sets of problems) available.
- The disadvantage of Matlab is that it uses a significant amount of memory and it is relatively slow in executing if-statements, and for- and while-loops that can not be vectorized.
- Vendor information: The MathWorks, Inc.
<http://www.mathworks.com/>

- MINITAB is an easy-to-use, general purpose software package for statistical analysis,
- MINITAB covers the basic range of statistical analyzes and high-resolution graphics.
- Minitab can handle up to 150 million cells per worksheet. One can have at most 4,000 columns and at most 10 million rows.
- MINITAB used widely by statisticians, especially in teaching, it has a relatively simple command structure which allows you to perform complicated analyzes using small numbers of commands.
- Vendor information: Minitab Inc.
<http://www.minitab.com/>

- SAS is an extremely powerful, comprehensive and flexible package for the integrated management and statistical analysis of data - it is marketed as an “information delivery system”.
- SAS is widely used in research, the public sector, finance, industry and commerce, SAS covers the vast majority of statistical analyzes.
- Highlights of SAS include its data manipulation capabilities, and the breadth of statistical procedures available.

- SAS is largely seen as a programmer's package, but there are interactive point-and-click interfaces available to SAS for UNIX under X-windows (equivalent to those under Microsoft Windows)
- Most users find SAS a more difficult package to use than SPSS or STATA, but SAS covers more analyzes than SPSS and is better at handling complicated data than either SPSS or STATA.
- The disadvantage of SAS is its idiosyncratic syntax. Some users may find SAS a difficult package to get into because of its language.

- SAS statements must be ordered in logical steps:
 - DATA steps are used to input and manipulate data
 - PROC steps are used to analyze data.
 - Data must be read into a temporary or permanent SAS data set before they can be analyzed.
- Vendor information: SAS Institute, Inc.
<http://www.sas.com/>

SPSS

- SPSS is an acronym for “Statistics Package for the Social Sciences” originally and modified to “Statistical Product and Service Solutions” later.
- It is now officially named “IBM SPSS Statistics”.
- SPSS is a powerful, comprehensive, and flexible statistical and information analysis system.
- SPSS can take data from almost any type of file to generate tabulated reports, plots of distributions and trends, descriptive statistics, and complex statistical analyzes.
- SPSS is a general purpose statistical package with a SPSS is widely used package in the social sciences, particularly in sociology and psychology.
- Vendor information: SPSS Inc.
<http://www.spss.com/>

- STATA is a very powerful, flexible general-purpose package for data analysis and data management, with excellent graphics capabilities and a graphic editor.
- STATA covers a wide range of statistical techniques and is programmable, allowing the user to add new commands.
- STATA processes data efficiently by reading all data into memory. This makes the package different to SAS and SPSS. (However, this is a disadvantage of STATA, large amount of (RAM) memory required to manage datasets.)

- STATA is not point-and-click, but most users find it easy to use. The command language is quite easy, and it covers a wider range of analyzes than SPSS.
- The package is repeatedly updated every two months or so using STATA STBs, and STATA tends to have the most up-to-date analysis procedures.
- Vendor information: Stata Corporation
<http://www.stata.com/>

- StatXact is the most complete, up-to-date software package for exact nonparametric statistical inference on continuous or categorical data.
- StatXact covers mainstream nonparametric procedure, using a full-fledged GUI implementation under Windows.
- Allows user to construct their own exact procedures, or to combine two or more procedures in an optimal fashion through a suitable choice of scores.
- StatXact is powerful when it comes to handling small-sample data.
- Vendor information: Cytel Software
<http://www.cytel.com>