Statistical Computing and Dynamic Graphics Using Lisp-Stat

Luke Tierney
School of Statistics
University of Minnesota

August 27, 1998

Abstract

Lisp-Stat is an environment for general statistical computing and for using and developing dynamic graphical methods. The Lisp-Stat system contains a set of functions for performing basic statistical calculations and constructing dynamic statistical graphs such as linked scatterplots and rotatable three-dimensional plots. In addition, new numerical and graphical tools can be added to the system using the Lisp programming language and a simple object-oriented programming system.

The course will begin with an overview of the capabilities of the system and then introduce basic elements of Lisp programming and concepts of object-oriented programming. These ideas will then be used on several examples to illustrate how to add new graphical and numerical methods to the system. The examples may include a simple set of tools for fitting generalized linear models and an implementation of the Grand Tour for viewing data sets in four or more dimensions.

Several extended breaks will be scheduled and a small number of microcomputers or workstations will be made available during these breaks to allow participants to experiment with the ideas presented.

What is Lisp-Stat

Lisp-Stat is an extensible environment for statistical computing and dynamic graphics.

Some of its features are

- a variety of numerical statistical operations
- a variety of interactive and dynamic graphical methods
- a very high level programming language (Lisp) that can be used to
  - simplify combinations of calculations
  - adapt methods to specific problems
  - add new capabilities to the system
- an extensible graphics system through
  - access to graphical interface tools (Menus and Dialogs)
  - standard graphs as building blocks
  - use of object-oriented programming

Some things Lisp-Stat does not do:

- incorporate a large collection of tools for specialized analyses

  Such tools can be implemented within the system or, if C or Fortran implementations are already available, they can be accessed through dynamic or static loading.

- provide extensive support for presentation graphics

  Other systems on workstations and PC's already provide a wide range of tools for preparing presentation graphics.

Implementation and Portability:

- Lisp-Stat is a general specification
- XLISP-STAT is a first implementation
- XLISP-STAT runs on
  - Apple Macintosh
  - MS Windows
  - Commodore Amiga (J. Lindsey)
  - *Sun* workstations using *SunView*
  - BSD UNIX workstations running *X11*
- A Kyoto Common Lisp-based implementation may be available soon

Obtaining XLISP-STAT

XLISP-STAT source code and executables are available free of charge.

Source code for the UNIX, Macintosh, and MS Windows versions is available by anonymous *ftp* from `umnstat.stat.umn.edu` and by email from `statlib`.

> If you have access to the *internet* but have never used *ftp*, there is probably someone at your site who can help you.

Executables for Macintosh and MS Windows are also available for anonymous *ftp* from `umnstat.stat.umn.edu` and some other sites.

If you do not have access to *ftp*, you can get a copy of the Macintosh or MS Windows executables by sending disks and a self-addressed, stamped mailer to

Luke Tierney
School of Statistics
University of Minnesota
Minneapolis, MN 55455

The number of disks required is

Two 800K disks for the Macintosh version

One $3\frac{1}{4}$" high-density disk for the MS Windows version

Documentation

A tutorial introduction is available as a technical report:

Tierney, L., (1989), "XLISP-STAT: A statistical environment based on the XLISP language," U of M Tech. Rep. 528.

The LaTeX source for this report is available for *ftp* from `umnstat.stat.umn.edu` or by email from `statlib`.

More complete documentation is available as a book:

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

The technical report corresponds to Chapter 2 of the book.

Some Historical Notes

- Several extensible statistical environments have been based on high level languages:
  - The New *S* language
  - APL-based systems (Anscombe)
  - Tools for the *Gauss* system
- Several researchers have based systems on the Lisp language:
  - McDonald and Pedersen
  - Stuetzle
  - Oldford and Peters
  - Buja and Hurley
- A strong argument in favor of Lisp is that Lisp provides excellent support for *experimental programming*.

Course outline

1. Overview of Lisp-Stat
2. A Tutorial Introduction
3. Some Lisp Programming
4. Objects
5. Outline of the Graphics System
6. Some Dynamic Graphics Examples

Course Objectives

After completing the course, you should be able to

- use Lisp-Stat with your own data for basic statistical calculations and graphics.
- develop some simple numerical and graphical tools for your own applications
- understand the basics of Lisp programming and object-oriented programming

You should also be well prepared to learn more about Lisp-Stat by working through the book.