

# Comment

Forrest W. Young

## 1. INTRODUCTION

This review, which covers the Lisp-Stat software as well as the Lisp-Stat book, is organized around the full title of the book, to which I have added the words *functional*, *extensible* and *free*. Fortunately, Luke Tierney didn't use all of these words in his title (which would have become something like "Lisp-Stat: A Functional and Object-Oriented Environment for Statistical Computing and Dynamic Graphics that Is Extensible and Free"), but this expanded title does incorporate what I see as all of the most important aspects of Lisp-Stat, and does make for a good way of organizing this review.

## 2. LISP

The key insight in Lisp-Stat is that it would be reasonable, even beneficial, to develop a statistical system based on the Lisp language. This insight was not Tierney's, but rather it belongs to one of the computational statisticians who have advocated such a development over the last few years (see McDonald and Pedersen, 1988; Oldford and Peters, 1988; Stuetzle, 1987; and Buja, Asimov, Hurley and McDonald, 1988). At first, Lisp seems like a very strange choice, especially for those of us who are only familiar with procedural languages such as Fortran, C or Pascal. In fact, Tierney asks very early in his book (page 3) the rhetorical question: "Why Lisp?" His answer: Lisp combines the strengths of a general purpose programming language with the needs in data analysis for interactive, experimental programming. He also notes that Lisp is a functional language that is object-oriented and that can be easily and elegantly extended (see below). There are few, if any, other general purpose languages that provide all of these characteristics.

At first, Lisp is a bit strange to use, since it is based entirely on prefix notation using a very simple and consistent syntax. However, the consistency and simplicity of the notation and syntax is soon seen to be one of the true strengths of the language. Indeed, the language and the statistical system are very easy to learn and to teach to novice

programmers and novice data analysts. The tutorial chapter (2) is good for getting started with Lisp-Stat and Chapters 3 and 4 on Lisp provide an introductory overview of Lisp that is sufficient for using Lisp-Stat. If one wants to do serious program development a book specifically on Lisp (Winston and Horn, 1989) will be needed.

Lisp-Stat is currently based on the XLISP implementation of Lisp, which is a subset of the industry-standard version of Lisp known as Common Lisp. XLISP has two drawbacks. The major drawback is that XLISP has no compiler, only an interpreter. While this is appropriate for "interactive, experimental programming," a compiler is better for Lisp-Stat code that is stable and that will be re-used many times. Tierney (personal communication) says that a new implementation of Lisp-Stat, based on Kyoto Common Lisp (KCL), which has a compiler, is being developed. The second drawback, which seems to be less critical, is that XLISP does not implement CLOS, the industry standard Common Lisp Object System. Since object-oriented programming is very important for dynamic graphics, Tierney has developed his own, nonstandard subsystem. While I have no experience with CLOS, Tierney's object system (TOS?) is easy to learn and use, and seems appropriate for dynamic statistical graphics and for other data analysis uses. The KCL implementation will include both object systems.

## 3. STAT

What Lisp lacks, and what Lisp-Stat adds, are the capabilities needed for statistics, data analysis and statistical graphics. Thus Lisp-Stat is Lisp extended to include vectorized arithmetic, functions for basic statistical computations (mean, standard-deviation, etc.), functions for maximization and maximum likelihood estimation, functions for linear algebra and matrix manipulation (transpose, svd, etc.), an interface to several windowing systems (X-Window, Amiga, Macintosh, but not yet Microsoft Windows) and tools for constructing graphs and dynamic graphs.

What this provides is an environment in which it is easy to perform basic statistical calculations, optimization, matrix algebra and dynamic graphics. *What this does not provide* (and here we have the major limitation of Lisp-Stat) is an environment in which it is easy to manipulate data (there is no spreadsheet, no datasets, etc.) or to perform

---

*Forrest W. Young is Professor of Psychometrics and Biostatistics at the University of North Carolina, BC 3270 Davie Hall, Chapel Hill, North Carolina 27599-3270.*