myJavaStat: an Environment for Teaching Statistics

E. James Harner
Hengyi Xue
West Virginia University, Department of Statistics
423 Hodges Hall
Morgantown, WV 26506, USA
jharner@stat.wvu.edu

myJavaStat is a Java application for teaching algebra-based statistics courses. It consists of a series of modules organized by chapters to accompany the textbook: Statistics: the Craft of Data Collection, Description, and Inference (http://www.8-mobius.com), but it can be used independently of any textbook. The software is available for free download at: http://javastat.stat.wvu.edu. myJavaStat consists of the Data, Five-Step, and Grapher modules as well as many smaller modules.

The Data Module

The Data module is a highly interactive data analysis environment. It builds on data tables that are dynamically linked to the analysis and graphical views. Observations can be masked, colored, and labeled in either the table or plot views. The analyses are based on a number of plot types in which the X and Y variables (axes) can either be categorical or numerical. All analyses and plots can be dynamically conditioned on one or more categorical and numerical variables.

The Five-Step Module

The Five-Step module is a powerful tool for simulating problems that are built on user-specified or standard probability models. In Step 1 the user species a discrete probability model or the user can select from standard probability models such as the normal, binomial, Poisson, etc. In Step 2 the user selects the sampling method and sample size. In the next step the statistic of interest and optionally the event of interest are given. The simulation is done in Step 4. Finally, in Step 5 the results are presented analytically and graphically.

Very realistic problems, such as epidemics, can be modeled. The Five-Step Method allows students to study problems that are (often) easy to explain, but difficult to solve analytically. It leads naturally to topics such as permutation tests.

The Grapher Module

The Grapher module is an computer-algebra and modeling environment. It is organized around functions and models. Multiple functions can be plotted, manipulated, and transformed. These operations are done dynamically. The user can zoom in and has the ability to solve algebra problems graphically. Models are like functions except that they are connected to data. The data table, which can be generated by plotting points, by entering data, or by importing files, is similar to the table in the Data module. The data can be edited and transformed. Standard models as well as non-linear models can be fitted, assessed, and dynamically manipulated. Functions and models can appear on the plot at the same time if desired.

Planned Enhancements

myJavaStat is a standalone Java applications that has the ability to communicate with a remote server. Currently, data in XML format can be imported from a server. We plan to link myJavaStat into our Web-based environment called IDEAL (http://ideal.stat.wvu.edu). This will then give myJavaStat extensive course management features, including examples, exercises, and quizzes. IDEAL includes an extensive reporting system as well as database-driven mail and chat components.

We plan to build a back-end assessment system using R, which will be called from a Java servlet. It will retrieve data from the database and then build cognitive assessments models using hierarchical Bayesian methodologies. Feedback will be provided to individual students on what they know and what they don't know and it will be updated weekly. Summary information on what students know will be given to the instructor.

REFERENCES

Harner, E. James. (1999). "A 3-tier Web-based Adaptive Statistical Learning Environment," ISI 99:52nd Session, pp. 1718 in Book 2: Contributed Papers, Helsinki.

Harner, E. James and Hengyi Xue. (2000). "JavaStat: A Client/Server Statistical Computing Environment," in Proceedings in Computational Statistics 2000: Short Communications and Posters, ed. by Wim Jansen and Jelke G. Bethehem. Statistics Netherlands: 199-200.

Harner, E. James, Hengyi Xue, Lingyi Zheng, and Jun Tan. (2001). "A Java/XML Distributed Environment for Statistical Computing," in the Bulletin of the International Statistical Institute, Book 3, 53rd Session, Seoul, Korea: 199-200.

Xue, Hengyi and E. James Harner. (2001). "JavaStat: A Distributed Statistical Computing Environment," in the DSC 2001 Proceedings of the 2nd International Workshop on Distributed Statistical Computing, ed. by K. Hornik and F. Leisch, ISSN 1609-395X, Technische Universitt Wien, Vienna, Austria. http://www.ci.tuwien.ac.at/Conferences/DSC-2001/Proceedings/

RÉSUMÉ

myJavaStat c'est une Java ambience pour enseignement préliminaire statistiques utilisant simulation et problème solvant. Les principal module êtes Données (pour problème solvant), Cinq–Trace (pour simulation probabilité modèles), et Graphique (pour faits algèbre et données travailler comme mannequin). myJavaStat c'est sous active développement à font le travailler à une serveur et une verso-la fin base de données pour trajet l'administration et étudiante évaluation.