Features of an internet-based learning package for advanced statistics

Ulrike Genschel, Sebastian Paris Scholz, Ursula Gather
University of Dortmund, Department of Statistics
Vogelpothsweg 87
44221 Dortmund ,Germany
genschel@statistik.uni-dortmund.de

The discovery of the internet as a useful teaching medium has lead to an elaborate accumulation of interactive educational material for statistics. The collection available on the internet consists of, for example, numerous JAVA applets, graphical visualizations, exercises and self-tests (e.g., multiple choice questions), software-type applications, as well as simple supplemental reading material. It appears that the notion of interactive learning is as diverse as above-mentioned learning tools (which will be seen in a later example).

The vast majority of this internet-based learning material is intended for introductory level statistics (i.e., basic statistical concepts required in the curricula of most non-statistics subjects). Interactive internet materials that go beyond the statistical basics are much more difficult to find. It seems that the more complex a statistical concept becomes, the less interactivity is found in available applets or other applications, if they exist at all.

As an example, consider existing JAVA applets that enable students to self-explore simple linear regression. Often students can move single data points and directly observe the corresponding change in the regression line. There are similar applets illustrating the behaviors of the arithmetic mean and the median upon adding large outlying observations to a data set. There is a large amount of interactivity to teach these important, but basic, concepts in statistics. On the other hand, JAVA applets found on the internet for teaching (more sophisticated) statistical testing theory are typically constructed so that, roughly speaking, the applet evaluates a test statistic (i.e., p-value) after the students provide some input (e.g. data).

That is, most JAVA applets for statistical tests function like a statistical software-package: data are entered in a window and then a test is performed yielding a result (with the benefit that the student can perform tests without knowledge of any statistical software). However, these applet constructions do not help to actually foster a student appreciation or understanding (even intuitively) of how a statistical tests works or what it does; in some ways, such applets may be considered as counterproductive to statistical education.

Obviously, it is not always easy, or even possible, to develop an interactive applet for a given statistical concept. From our own experiences in developing interactive learning material for advanced topics in statistics, it seems that the following difficulties are the ones most often encountered:

1. Prerequisites should be made clear and available. At the same time, it is difficult to build an advanced learning environment that is self-contained, as it ideally should be. Although prerequisite learning material (for meeting a further educational goal) is often included in the learning system, it can never be quite complete as the objective learning material becomes advanced. One way to encounter this problem is to create a list of background statistical concepts needed before proceeding with more demanding lesson objectives. To include prerequisite learning material, a hyperlinked structure has shown to be very helpful and should be made available to student.

2. Interactive graphics or visualization of statistical concepts at an advanced level can be challenging to create.

We will finally present an interactive Java applet trying to meet the proposed standards. The

goal of the applet is to illustrate the complex notion of statistical risk functions in decision theory. In the applet, students can choose among different decision and loss functions, various distributional classes as well as several types of estimators, to graphically illustrate the abstract concept of statistical risk. The Java applet is constructed in such a manner that the user receives the opportunity to review each element in the basic framework of decision theory and to practice the steps needed to perform the computation and visualisation of such statistical risks. The handling of risk function and the interpretation of needed components are given in detail as well.

REFERENCES

Cramer, E., Cramer, K. and Kamps, U. (2002) e-stat: A Web-based Learning Environment in Applied Statistics. In Proceedings of the 15th Conference on Computational Statistics (eds W. Hrdle and B. Rnz), 309-314. Berlin.

Lane, D. (1993-2000). Rice Virtual Lab in Statistics. Applets are available at http://www.ruf.rice.edu/∼lane/rvls.html.

Lee, J.J., Kang, G. and Han, K.S. (2002). Computer Aided Teaching for Statistics in Internet Age. Computational Statistics 17, 355-365.