Interactive graphics for elementary statistical education

Katharina Cramer
University of Oldenburg, Department of Mathematics
26111 Oldenburg, Germany
e-stat@uni-oldenburg.de

Udo Kamps
University of Oldenburg, Department of Mathematics
26111 Oldenburg, Germany
kamps@mathematik.uni-oldenburg.de

1. The system EMILeA-stat

Within the "New Media in Education Funding Programme" the German Federal Ministry of Education and Research (bmb+f) supports the project "e-stat" (project period April 2001 - March 2004) in developing and providing a multimedia, web-based, and interactive learning and teaching environment in applied statistics called EMILeA-stat. The structure of EMILeA-stat, its scope and objectives are sketched. The talk includes an insight into the system with a focus on interactive visualizations.

Statistical and quantitative thinking and acting have become fundamental skills in several branches of natural sciences, life sciences, social sciences, economics, and engineering. Owing to these various applications and the necessity of using statistical methodology in so many fields, the processes of learning and teaching should be continuously improved, mainly by increasing motivational and graphical aspects. EMILeA-stat (see www.emilea.de) is developed as a system with a strong graphical concept and suitable for teaching statistics at schools, universities, and in further vocational training. Basic statistical contents are presented on three levels of abstraction in order to take into account that different types of users have different needs, because of their individual mathematical and theoretical backgrounds.

Furthermore, user-oriented views and scenarios, which are near to real world applications, are integrated into the learning and teaching environment EMILeA-stat. Thereby, different teaching concepts and instructional designs are supported and the system can be seen as an innovative high-quality contribution to the learning facilities of the present and the future.

2. Interactive visualizations

EMILeA-stat covers a variety of topics in statistics and probability. The theoretical statistical content is supported by interactive visualizations, such as Java-Applets, throughout. In the talk some examples concerning descriptive statistics and the main tools of presenting data are explained. Starting with elementary visualizations like pie charts, bar and line graphs, box plots, stem-and-leaf plots, histograms, and plots of the empirical distribution function, measures of location and scale like mean, standard deviation, quantiles, etc. are considered as well as measures of relationships between measurement or categorical variables.

Methods for the description of economic data like price index numbers and measures of concentration, e.g., Lorenz curves and Gini coefficients, are also illustrated by interactive visualizations. Furthermore, regression and time series analysis are explained by Java applets. Finally, procedures of multivariate data analysis like cluster analysis, principal component analysis, factor analysis, and multidimensional scaling supplement the toolbox of descriptive statistics and exploratory data analysis.

Each interactive visualization is available on three levels of abstraction. The level displayed is prescribed by the teacher but the user can of course, in case of interest, switch between these versions of the applet. The elementary level A offers a low degree of interactivity whereas on level C (advanced) the full range of functionality is accessible. The histogram applet, for example, provides many interactive features on level C: Included examples may be modified by adding new data points. They can be given numerically or by clicking with the right mouse button under the x-axis. Moreover, existing data values (points on the x-axis) can be moved to the right and left with the left mouse button. The axes are automatically rescaled and each bar can be split of into two bars by clicking with the mouse into it. The endpoints of the bars can also be shifted by the user. The number of bars or the length of the intervals can be specified. An optimal bandwidth may be chosen according to methods known in statistical literature. Finally, the numerical values for the intervals, frequencies, etc. are given in a table. Furthermore, the applet can also be used for representing own data. On level B this option is not included. Moreover, the bandwidth is always equidistant on the basic level. The most restricted version, the applet on level A, offers even less interactivity: It is only possible to manipulate the given data by adding points in the graphic and to choose the number of intervals.

The "user management" is standardized such that the frequent user should be able to work with a new applet easily. Instructions accompanying the graphics indicate the concept of learning by discovery enabled via the available interactivity.

REFERENCES

Cramer, E., Cramer, K., and Kamps, U. (2002). e-stat: A web-based learning environment in applied statistics. In COMPSTAT Proceedings in Computational Statistics (eds. W. Härdle and B. Rönz), 309-314. Physica-Verlag. Heidelberg.

Härdle, W. and Rönz, B. (eds.) (2002). COMPSTAT Proceedings in Computational Statistics. Physica-Verlag, Heidelberg.

Möbus, C., Albers, B., Hartmann, St., and Zurborg, J. (2002). Intelligent WBT: Specification and architecture of the distributed, multimedia e-learning system e-STAT. In COMPSTAT Proceedings in Computational Statistics (eds. W. Härdle and B. Rönz), 401-406. Physica-Verlag. Heidelberg.

Möbus, C., Albers, B., Hartmann, St., Thole, H.J., and Zurborg, J. (2002). Towards a specification of distributed and intelligent web based training systems. In Intelligent Tutoring Systems, Proceedings of the 6th International Conference, ITS2002 (eds. St.A. Cerri, G. Gouarderes, and F. Paraguacu), 291-300. Springer. Berlin.