Teaching Linear Regression in the PC Lab

Karsten Schmidt

Faculty of Business and Economics
Schmalkalden University of Applied Sciences
Germany

kschmidt@fh-sm.de

Workshop Proposal for the TI-Nspire & Derive Strand

ABSTRACT

The Bachelor program of the Faculty of Business and Economics at Schmalkalden University of Applied Sciences includes a good deal of statistics education as each student must take two courses in statistics. The first course (*Introductory Statistics*) is given in the traditional way, i.e. it takes place in a classroom equipped with blackboard and overhead projector; the only technology used are pocket calculators. The second course (*Intermediate Statistics*) was redesigned in recent years. The course is now held in the PC lab throughout the semester and its title was changed to *Computer-assisted Statistics*. In addition to statistical software (*SPSS*), a Computer Algebra System (*Derive*) is used throughout the course. Students are then already familiar with *Derive* as this software was used one semester earlier in a (required) course in matrix algebra. The faculty acquired a special *Derive* license that permits its use on the students' own PCs, such that all students have permanent access to *Derive*, be it in class, at home, or during the exam.

This presentation will be about new possibilities emerging from the fact that the course is taught in the PC lab, concentrating on topics from linear regression (which in fact is the main topic of the course). *SPSS* and *Derive* are used both for numerical and graphical analyses. Since the creation of graphs is more challenging in *Derive* than in *SPSS*, a utility file was developed to facilitate such tasks. This file includes functions for the preparation of a 2D scatterplot with the associated least squares regression line for any two-variable, or simple, linear regression model, as well as functions for a 3D scatterplot with the associated least squares regression plane for any three-variable model. In addition, the so-called system of normal equations, which plays a major role in the derivation of the Ordinary Least Squares (OLS) estimator, is a system of two straight lines in the case of simple regression, which again can be viewed in a 2D plot: the OLS estimator is the point (a vector in R^2) where the two lines intersect. In the case of a three-variable model, the OLS estimator is the point (a vector in R^3) where the three planes, defined by the system of normal equations, intersect. It is also interesting to look at the graph of the sum of squared residuals (SSR) which is the objective function that is minimized by the OLS estimator. The SSR function is a convex function of the two (three) elements of the parameter vector in simple (three-variable) linear regression. In the case of simple regression, a 3D plot reveals that the OLS estimator is the point where the three-dimensional SSR function has its minimum.

We will also take a quick look at some files and applications from the World Wide Web which are interesting and helpful in teaching linear regression.

Keywords
Linear Regression, Least Squares Method, PC lab, Graphics