LOWESS: A Program for Smoothing Scatterplots by Robust Locally Weighted Regression

The visual information on a scatterplot can be greatly enhanced with little additional cost by plotting smoothed points. Suppose the points of the scatterplot are $(x_i, y_i), i = 1, \ldots, n$. The smoothed points are $(x_i, \hat{y}_i)$ where $\hat{y}_i$, the fitted value at x_i, portrays the location of Y given $X = x_i$. Plotting the smoothed points, which form a nonparametric regression of Y on X, frequently allows the perception of effects on the scatterplot that are otherwise difficult to detect.

Robust locally weighted regression (Cleveland 1979) is a method for smoothing scatterplots in which the fitted value at x_k is the value of a line fit to the data using weighted least squares where the weight for (x_i, y_i) is large if x_i is close to x_k and small if x_i is not close to x_k. A robust fitting procedure guards against outliers distorting the smoothed points. Information on obtaining FORTRAN routines for robust locally weighted regression is available from the Computing Information Library at Bell Laboratories.

The routine LOWESS, which is directly called by the user, consists of 107 lines of code. LOWESS calls a support routine LOWEST, which has 79 lines of code, and requires a routine that sorts an array and a routine that sorts an array and passively sorts a second array according to the first. Two sort routines with 69 and 83 lines of code are provided for users who do not have sort routines on their systems. The user may elect to receive, along with the program documentation and listings, either a tape or a punched deck of the programs.

William S. Cleveland
Bell Laboratories
600 Mountain Ave.
Murray Hill, NJ 07974

REFERENCE

CLEVELAND, WILLIAM S. (1979), "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, 74, 829–836.

Lattice Designs: Analysis of Variance and Randomized Layouts

Two FORTRAN programs have been developed at the University of Illinois Agronomy Statistical Laboratory for application to lattice square, rectangular and square lattice, and randomized complete block experimental designs. One program (RANDLAT) constructs randomized experimental layouts for such experiments; the other (RCBANOV) performs analysis of variance on data collected from them. Both programs can accommodate $k \times k$ lattice squares with $(k + 1)/2$ or fewer replicates for $k = 3, 4, 5, 7, 8, 9, 11$, or 13, and one or two repetitions of double or triple $k(k + 1)$ rectangular lattices and $k \times k$ square lattices for $3 \leq k \leq 13$, as well as randomized complete block designs with $2 \leq t \leq 999$ treatments and $2 \leq r \leq 10$ replicates. Earlier versions of the programs are described in Carmer (1965) and Carmer, Seif, and Jacob (1963).

RANDLAT provides the user with a randomized assignment of treatments to experimental units in accordance with the restrictions on randomization appropriate to the specific design. For an individual experiment a "derandomized" layout is constructed first; for the various lattice designs the "derandomized" layouts are taken from Cochran and Cox (1957), Clem and Federer (1950), or Robinson and Watson (1949). Randomization procedures appropriate to the particular design are then applied; for the lattices these may be summarized as follows.

1. Assignment of treatment numbers at random,
2. Randomization of blocks or rows of treatments within replicates, and
3.(a) Randomization of columns of treatments within replicates if the design is a lattice square, or (b) randomization of treatments within blocks if the design is a rectangular or square lattice.

Copies of the printed output may be used for both office files and field or laboratory notebooks. Besides relieving the user of a tedious randomization procedure, use of the program eliminates copying errors and encourages use of a different randomization each time a particular design is constructed.

The analysis of variance program, RCBANOV, has excellent capabilities for transformation of input variables before analysis. Analysis of variance computations correspond to those presented in Cochran and Cox (1957). In addition to the analysis of variance table for each of up to 10 dependent variables, the output includes unadjusted and adjusted means, average effective error variance, F test of adjusted treat-