TRICUBE

PURPOSE

Carry out a tricube transformation (also called a triweight transformation).

DESCRIPTION

The tricube transformation is used in robust analysis. For many applications, it combines the properties of resistance with relatively high efficiency. Resistance means that changes in a small part of the data do not cause large changes in the estimate. The mean is an example of a non-resistant estimate while the median is an example of a resistant estimate. Efficiency is a measure of how well the estimate performs for data from a given distribution. For example, the mean is a 100% efficient estimator for normally distributed data. However, it has poor efficiency for heavy tailed distributions. A desirable property for robust estimators is that they maintain high efficiency under a variety of distributions. The tricube transformation of a variable has this property for many applications.

The tricube transformation of the variable X is defined as follows:

1. m = the median absolute value of X
2. cutoff = 6 * m
3. $T(X) = (1- (X/cutoff)^3)^3$ if abs(X) < cutoff
   $T(X) = 0$ if abs(X) >= cutoff

This transformation defines a set of weights that are used in a subsequent calculation. Data is down-weighted depending on its distance from the median absolute value and extreme outliers are given zero weight.

SYNTAX

LET <v1> = TRICUBE <v2> <SUBSET/EXCEPT/FOR qualification>
where <v2> is the input variable;
<v1> is an output variable where the tricube-transformed values are saved;
and where the <SUBSET/EXCEPT/FOR qualification> is optional.

EXAMPLES

```
LET W = TRICUBE RES
LET W2 = TRICUBE RES SUBSET TAG > 2
```

NOTE

The tricube transformation is most commonly applied to the residuals from a (linear or nonlinear) fit. The tricube transformation has the advantage that it allows the analyst to carry out subsequent weighted linear or non-linear fits that are robust and resistive to outliers in the data. The following is a typical sequence using TRICUBE:

```
FIT Y = A+B*EXP(-C*X)
LET W = TRICUBE RES
WEIGHTS W
FIT Y = A+B*EXP(-C*X)
```

An unweighted fit is performed first. Then the tricube transformation is applied to the residuals. These tricube transformed residuals are used as the weights in a weighted fit.

Robust fits based on the tricube are part of a class of techniques called M-estimators of regression. The fitting technique above is also called iteratively reweighted least squares. See the documentation for the WEIGHTS command (in Volume I) for examples of other weighting functions.

Be aware that this type of robust regression protects against outliers in the dependent variable (i.e., the Y). However, it is still susceptible to outliers in the independent variables. These outliers are called high leverage points. See the documentation for the FIT command (in Volume I) for details on identifying high leverage points (at least for linear fits).

The tricube transformation can be continued for additional iterations of the fitting until some convergence criterion is reached. For example, you can compute the difference between the residuals in two successive steps, sum these differences, and then stop when this sum is below some cutoff value. Although DATAPLOT does not do this automatically, it is straightforward to write a macro to do this.

DEFAULT

None

SYNONYMS

TRIWEIGHT for TRICUBE

RELATED COMMANDS

FIT	=	Carries out a least squares fit.
RES	=	A variable containing the residuals from a FIT.
WEIGHTS	=	Specifies the weights variable.
BIWEIGHT	=	Carry out a biweight transformation

REFERENCE

"Graphical Methods for Data Analysis," Chambers, Cleveland, Kleiner, and Tukey, Wadsworth, 1983 (page 122).

APPLICATIONS

Robust Fitting

IMPLEMENTATION DATE

88/7

PROGRAM

```
LET X = DATA 1 2 3 4 5 6 7 8 9 10
LET Y = DATA 2 4 60 7 9 12 14 15 18 20
FIT Y X; LET PRED2 = PRED
CHARACTER CIRCLE BLANK BLANK; LINE BLANK SOLID DASH
SEGMENT 1 COORDINATES 65 85 70 85; SEGMENT 1 PATTERN SOLID
SEGMENT 2 COORDINATES 65 81 70 81; SEGMENT 2 PATTERN DASH
LEGEND 1 ORIGINAL FIT; LEGEND 1 COORDIANTES 72 84
LEGEND 2 TRICUBE FIT; LEGEND 2 COORDIANTES 72 80
LET B = TRICUBE RES
WEIGHTS B
FIT Y X
PLOT Y PRED2 PRED VS X
```