

NAG Toolbox for MATLAB

g10za

1 Purpose

g10za orders and weights data which is entered unsequentially, weighted or unweighted.

2 Syntax

```
[nord, xord, yord, wtord, rss, ifail] = g10za(weight, x, y, wt, 'n', n)
```

3 Description

Given a set of observations (x_i, y_i) , for $i = 1, 2, \dots, n$, with corresponding weights w_i , g10za rearranges the observations so that the x_i are in ascending order.

For any equal x_i in the ordered set, say $x_j = x_{j+1} = \dots = x_{j+k}$, a single observation x_j is returned with a corresponding y' and w' , calculated as

$$w' = \sum_{l=0}^k w_{i+l}$$

and

$$y' = \frac{\sum_{l=0}^k w_{i+l} y_{i+l}}{w'}.$$

Observations with zero weight are ignored. If no weights are supplied by you, then unit weights are assumed; that is $w_i = 1$, for $i = 1, 2, \dots, n$.

In addition, the within group sum of squares is computed for the tied observations using West's algorithm (see West 1979).

4 References

Draper N R and Smith H 1985 *Applied Regression Analysis* (2nd Edition) Wiley

West D H D 1979 Updating mean and variance estimates: An improved method *Comm. ACM* **22** 532–555

5 Parameters

5.1 Compulsory Input Parameters

1: **weight** – string

Indicates whether user-defined weights are to be used.

If **weight** = 'W', user-defined weights are to be used and must be supplied in **wt**.

If **weight** = 'U', the data is treated as unweighted.

Constraint: **weight** = 'W' or 'U'.

2: **x(n)** – double array

The values, x_i , for $i = 1, 2, \dots, n$.

3: **y(n)** – double array

The values, y_i , for $i = 1, 2, \dots, n$.

4: **wt(*)** – double array

Note: the dimension of the array **wt** must be at least 1 if **weight** = 'U' and at least **n** if **weight** = 'W'.

If **weight** = 'W', **wt** must contain n weights, w_i , for $i = 1, 2, \dots, n$.

If **weight** = 'U', **wt** is not referenced.

Constraint: if **weight** = 'W', $\mathbf{wt}(i) \geq 0.0$, for $i = 1, 2, \dots, n$, and at least one $\mathbf{wt}(i) > 0.0$ for some i .

5.2 Optional Input Parameters

1: **n** – int32 scalar

Default: The dimension of the arrays **x**, **y**, **xord**, **yord**, **wtord**. (An error is raised if these dimensions are not equal.)

n , the number of observations.

Constraint: $n \geq 1$.

5.3 Input Parameters Omitted from the MATLAB Interface

iwrk

5.4 Output Parameters

1: **nord** – int32 scalar

The number of distinct observations.

2: **xord(n)** – double array

The first **nord** elements contain the ordered and distinct x_i .

3: **yord(n)** – double array

The first **nord** elements contain the values y' corresponding to the values in **xord**.

4: **wtord(n)** – double array

The first **nord** elements contain the values w' corresponding to the values of **xord** and **yord**.

5: **rss** – double scalar

The within group sum of squares for tied observations.

6: **ifail** – int32 scalar

0 unless the function detects an error (see Section 6).

6 Error Indicators and Warnings

Errors or warnings detected by the function:

ifail = 1

On entry, **weight** $\neq$ 'W' or 'U',
or **n** < 1.

ifail = 2

On entry, **weight** = 'W' and at least one element of **wt** is < 0.0 , or all elements of **wt** are 0.0.

7 Accuracy

For a discussion on the accuracy of the algorithm for computing mean and variance see West 1979.

8 Further Comments

g10za may be used to compute the pure error sum of squares in simple linear regression along with g02da; see Draper and Smith 1985.

9 Example

```
weight = 'U';
x = [1;
     3;
     5;
     5;
     3;
     4;
     9;
     6;
     9;
     9];
y = [4;
     4;
     1;
     2;
     5;
     3;
     4;
     9;
     7;
     4];
wt = [0];
[nord, xord, yord, wtord, rss, ifail] = g10za(weight, x, y, wt)

nord =
        6
xord =
     1
     3
     4
     5
     6
     9
     6
     9
     9
     9
yord =
     4.0000
     4.5000
     3.0000
     1.5000
     9.0000
     5.0000
     9.0000
     4.0000
     7.0000
     4.0000
wtord =
     1
```

```
      2
      1
      2
      1
      3
      0
      0
      0
      0
rss =
      7
ifail =
      0
```
