

## Euredit outlier detection function: bacon

### 1 Purpose

**bacon** detects outliers using modified BACON algorithm.

### 2 Specification

```
#include <euredit_sys.h>
```

```
void bacon (long n, long m, long *sx, double *data, double miss_val, long c,
            double alpha, long method, double *indwt, double *swts, long info[2])
```

### 3 Parameters

**n**

*Input:* the number of observations in the data.

*Constraint:*  $n \geq 1$ .

**m**

*Input:* the number of variables in the data.

*Constraint:*  $m \geq 1$ .

**sx[m]**

*Input:* if  $sx[i] > 0$  the  $i$ th variable is included in the analysis otherwise it is excluded from the analysis.

**data[n\*m]**

*Input:* the data stored by row.

**miss\_val**

*Input:* the flag for missing values in the data.

**c**

*Input:* specifies the size of the initial subset of the data as  $c*m$ .

*Constraint:*  $c \geq 1$ .

**alpha**

*Input:* the rejection level.

*Constraint:*  $0 < c < 1$ .

**method**

*Input:* if **method** = 1 then the EM algorithm is used to estimate missing values and the Mahalanobis metric is used; otherwise records with missing values are excluded from the the set of good observations and a Euclidean distance is used.

**indwt[n]**

*Output:* indicative weights. If  $indwt[i] = 1$  the  $i$ th observation is accepted, if  $indwt[i] = 0$  the  $i$ th observation is rejected.

**swts[n]**

*Input:* **swts**[ $i$ ] contains the weight on the  $i$ th observation. If **swts** is NULL, all weights are taken to be 1.

**info**

*Output:* information on the success of the function call.

**info**[0] = 0: the function successfully completed its task.

**info**[0] =  $i$ : the specification of the  $i$ th formal parameter was incorrect,  $i = 1, 2, \dots, 11$ .

**info**[0] = 11: too few complete observations for value of  $c$ .

**info**[0] = 19: failure to compute  $\chi^2$  value.

**info**[0] = 20: failure in calculating covariance matrix.

**info**[0] = 25: failure in EM algorithm.  
**info**[0] = 30: covariance matrix not positive definite.  
**info**[0] = 99: the function failed to allocate enough memory.  
**info**[1] contains additional information for system debugging.