nag_robust_trimmed_1var (g07ddc)

1. Purpose

nag_robust_trimmed_1var (g07ddc) calculates the trimmed and Winsorized means of a sample and estimates of the variances of the two means.

2. Specification

```
#include <nag.h>
#include <nagg07.h>

void nag_robust_trimmed_1var(Integer n,  double x[], double alpha,
             double *tmean, double *wmean, double *tvar, double *wvar,
             Integer *k, double sx[], NagError *fail)
```

3. Description

nag_robust_trimmed_1var calculates the α-trimmed mean and α-Winsorized mean for a given α, as described below.

Let x_i, for $i = 1, 2, \ldots, n$ represent the n sample observations sorted into ascending order. Let $k = [\alpha n]$ where $[y]$ represents the integer part of y.

Then the trimmed mean is defined as;

$$\bar{x}_t = \frac{1}{n-2k} \sum_{i=k+1}^{n-k} x_i,$$

and the Winsorized mean is defined as;

$$\bar{x}_w = \frac{1}{n} \sum_{i=k+1}^{n-k} x_i + (kx_{k+1}) + (kx_{n-k}).$$

nag_robust_trimmed_1var then calculates the Winsorized variance about the trimmed and Winsorized means respectively and divides by n to obtain estimates of the variances of the above two means.

Thus we have

$$\text{Estimate of } \operatorname{var}(\bar{x}_t) = \frac{1}{n^2} \sum_{i=k+1}^{n-k} (x_i - \bar{x}_t)^2 + k(x_{k+1} - \bar{x}_t)^2 + k(x_{n-k} - \bar{x}_t)^2$$

and

$$\text{Estimate of } \operatorname{var}(\bar{x}_w) = \frac{1}{n^2} \sum_{i=k+1}^{n-k} (x_i - \bar{x}_w)^2 + k(x_{k+1} - \bar{x}_w)^2 + k(x_{n-k} - \bar{x}_w)^2.$$

4. Parameters

n

Input: the number of observations, n.
Constraint: **n** ≥ 2.

x[n]

Input: the sample observations, x_i, for $i = 1, 2, \ldots, n$.

alpha

Input: the proportion of observations to be trimmed at each end of the sorted sample, α.
Constraint: $0.0 \leq$ **alpha** < 0.5.

tmean
Output: the α-trimmed mean, $\bar{x}_t$.

wmean
Output: the α-Winsorized mean, $\bar{x}_w$.

tvar
Output: contains an estimate of the variance of the trimmed mean.

wvar
Output: contains an estimate of the variance of the Winsorized mean.

k
Output: contains the number of observations trimmed at each end, k.

sx[n]
Output: contains the sample observations sorted into ascending order.

fail
The NAG error parameter, see the Essential Introduction to the NAG C Library.

5. Error Indications and Warnings

NE_INT_ARG_LT
On entry, **n** must not be less than 2: **n** = ⟨*value*⟩.

NE_REAL_ARG_LT
On entry, **alpha** must not be less than 0.0: **alpha** = ⟨*value*⟩.

NE_REAL_ARG_GE
On entry, **alpha** must not be greater than or equal to 0.5: **alpha** = ⟨*value*⟩.

NE_INTERNAL_ERROR
An internal error has occurred in this function. Check the function call and any array sizes. If the call is correct then please consult NAG for assistance.

6. Further Comments

The time taken by the routine is proportional to n.

6.1. Accuracy

The results should be accurate to within a small multiple of ***machine precision***.

6.2. References

Hampel F R, Ronchetti E M, Rousseeuw P J and Stahel W A (1986) *Robust statistics. The approach based on influence functions.* Wiley
Huber P J (1981) *Robust statistics.* Wiley

7. See Also

None.

8. Example

The following program finds the α-trimmed mean and α-Winsorized mean for a sample of 16 observations where $\alpha = 0.15$. The estimates of the variances of the above two means are also calculated.

8.1. Program Text

```
/* nag_robust_trimmed_1var(g07ddc) Example Program.
 *
 * Copyright 1996 Numerical Algorithms Group.
 *
 * Mark 4, 1996.
 *
 */

#include <nag.h>
#include <stdio.h>
#include <nag_stdlib.h>
#include <nagg07.h>

#define NMAX 1000
main()
{

  /* Local variables */
  double tvar, wvar;
  double alpha, x[1000], tmean, wmean, propn, sx[1000];

  Integer i, k;
  Integer n;

  Vprintf("g07ddc Example Program Results\n\n");
  /*     Skip heading in data file */
  Vscanf("%*[^\n] ");

  Vscanf("%ld ", &n);
  for (i = 1; i <= n; ++i)
    Vscanf("%lf ", &x[i - 1]);
  Vscanf("%lf ", &alpha);

  g07ddc(n, x, alpha, &tmean, &wmean, &tvar, &wvar, &k, sx, NAGERR_DEFAULT);

  propn = (double) k / n;
  propn = 100. - propn * 200.;
  Vprintf("Statistics from middle %6.2f%% of data\n\n", propn);
  Vprintf("               Trimmed-mean = %11.4f\n", tmean);
  Vprintf("   Variance of Trimmed-mean = %11.4f\n\n", tvar);
  Vprintf("            Winsorized-mean = %11.4f\n", wmean);
  Vprintf("Variance of Winsorized-mean = %11.4f\n", wvar);
  exit(EXIT_SUCCESS);
}
```

8.2. Program Data

```
g07ddc Example Program Data
16
26.0 12.0 9.0 2.0 5.0 6.0 8.0 14.0 7.0 3.0 1.0 11.0 10.0 4.0 17.0 21.0
0.15
```

8.3. Program Results

```
g07ddc Example Program Results

Statistics from middle  75.00% of data

               Trimmed-mean =      8.8333
   Variance of Trimmed-mean =      1.5434

            Winsorized-mean =      9.1250
Variance of Winsorized-mean =      1.5381
```