nag_summary_stats_1var (g01aac)

1. Purpose

nag_summary_stats_1var (g01aac) calculates the mean, standard deviation, coefficients of skewness and kurtosis, and the maximum and minimum values for a set of ungrouped data. Weighting may be used.

2. Specification

```
#include <nag.h>
#include <nagg01.h>

void nag_summary_stats_1var(Integer n, double x[], double wt[],
     Integer nvalid[], double xmean[], double xsd[], double xskew[],
     double xkurt[], double xmin[], double xmax[], double wsum[],
     NagError *fail)
```

3. Description

The data consist of a single sample of n observations, denoted by x_i, with corresponding weights, w_i, for $i = 1, 2, \ldots, n$.

If no specific weighting is required, then the array w need not be defined.

The quantities computed are as follows.

(a) The sum of the weights $W = \sum_{i=1}^{n} w_i$.

(b) Mean $\bar{x} = \dfrac{\sum_{i=1}^{n} w_i x_i}{W}$.

(c) Standard deviation $s = \sqrt{\dfrac{\sum_{i=1}^{n} w_i (x_i - \bar{x})^2}{d}}$ where $d = W - \left(\sum_{i=1}^{n} w_i^2\right) \Big/ W$.

(d) Coefficient of skewness $\gamma_1 = \dfrac{\sum_{i=1}^{n} w_i (x_i - \bar{x})^3}{s^3}$.

(e) Coefficient of kurtosis $\gamma_2 = \dfrac{\sum_{i=1}^{n} w_i (x_i - \bar{x})^4}{s^4} - 3$.

(f) Maximum and minimum elements of the sample.

(g) The number of valid observations for which $w_i > 0$, if the weighting is supplied, or n. Suppose m observations are valid, then the quantities in (c), (d) and (e) will be computed if $m \geq 2$, and will be based on $m - 1$ degrees of freedom. The other quantities are evaluated provided $m \geq 1$.

4. Parameters

n

Input: the number of observations, n.
Constraint: **n** ≥ 1.

x[n]

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

wt[n]
Input: if weighted estimates are required, then **wt** must contain the weights w_i, for $i = 1, 2, \ldots, n$. Otherwise, **wt** need not be defined and the corresponding argument must be set to the null pointer, ((double *)0).

nvalid
Output: the number m of valid observations – see Section 3(g) above.

xmean
Output: the mean, $\bar{x}$.

xsd
Output: the standard deviation, s.

xskew
Output: the coefficient of skewness, γ_1.

xkurt
Output: the coefficient of kurtosis, γ_2.

xmin
Output: the smallest value in the sample.

xmax
Output: the largest value in the sample.

wsum
Output: the sum of the weights in the array **wt**, that is $\sum_{i=1}^{n} w_i$. This will be n if weighted estimates are not used.

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

5. Error Indications and Warnings

NE_INT_ARG_LE
On entry, n must not be less than or equal to 0: **n** = ⟨*value*⟩.

NE_CASES_ONE
The number of valid cases is one. In this case, standard deviation and coefficients of skewness and of kurtosis cannot be calculated.

NE_CASES_ZERO
The number of valid cases is zero.

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

6. Further Comments

The time taken by the function is approximately proportional to n.

6.1. Accuracy

A single pass updating algorithm is used, which is believed to be stable.

7. See Also

nag_median_1var (g07dac)

8. Example

In the program below, nprob determines the number of data sets to be analysed. For each analysis, a set of observations and, optionally, weights is read and printed. After calling the function, the calculated quantities are printed. In the example, there is one set of data with 24 unweighted data values.

8.1. Program Text

```
/* nag_summary_stats_1var(g01aac) Example Program
 *
 * Copyright 1990 Numerical Algorithms Group.
 *
 * Mark 1, 1990.
 */

#include <nag.h>
#include <stdio.h>
#include <nag_stdlib.h>
#include <nagg01.h>

#define NMAX 30

main()
{
  double xsd, xskew, xkurt, wsum, xmean, xmax, xmin;
  double wt[NMAX], x[NMAX];
  Integer i, weight, j, n, nprob, nvalid;
  static NagError fail;

  /*  Skip heading in data file */
  Vscanf("%*[^\n]");
  Vprintf("g01aac Example Program Results\n");
  Vscanf("%ld", &nprob);
  for (j=1; j<=nprob; j++)
    {
      Vscanf("%ld %ld", &n, &weight);
      Vprintf("Problem  %5ld\n", j);
      Vprintf("Number of cases %ld\n", n);
      if (n<=0 || n>NMAX)
        {
          Vfprintf(stderr, "Error: n is out of range: n = %5ld\n", n);
          exit(EXIT_FAILURE);
        }
      else
        {
          for (i=0; i<n; i++)
            Vscanf("%lf", &x[i]);
          Vprintf("Data as input -\n");
          for (i=0; i<n; i++)
            Vprintf("%12.1f%c", x[i], (i%5==4 || i==n-1) ? '\n' : ' ');
          if (weight)
            {
              Vprintf("Weights as input -\n");
              for (i=0; i<n; i++)
                Vscanf("%lf", &wt[i]);
              for (i=0; i<n; i++)
                Vprintf("%12.1f%c", wt[i], (i%5==4 || i==n-1) ? '\n' : ' ');
              g01aac(n, x, wt, &nvalid, &xmean, &xsd, &xskew, &xkurt, &xmin,
                     &xmax, &wsum, &fail);
            }
          else
            g01aac(n, x, (double *)0, &nvalid, &xmean, &xsd, &xskew, &xkurt,
                   &xmin, &xmax, &wsum, &fail);

          if (fail.code==NE_NOERROR)
            {
              Vprintf("\n");
              Vprintf("Successful call of g01aac\n");
              Vprintf("No. of valid cases %5ld\n", nvalid);
              Vprintf("Mean           %13.1f\n", xmean);
              Vprintf("Std devn       %13.1f\n", xsd);
              Vprintf("Skewness       %13.1f\n", xskew);
              Vprintf("Kurtosis       %13.1f\n", xkurt);
              Vprintf("Minimum        %13.1f\n", xmin);
              Vprintf("Maximum        %13.1f\n", xmax);
              Vprintf("Sum of weights %13.1f\n", wsum);
```

```
            }
          else
            {
              Vprintf("Unsuccessful call of g01aac\n");
              if (fail.code==NE_CASES_ONE)
                {
                  Vprintf("No. of valid cases %5ld\n", nvalid);
                  Vprintf("Mean           %13.1f\n", xmean);
                  Vprintf("Minimum        %13.1f\n", xmin);
                  Vprintf("Maximum        %13.1f\n", xmax);
                  Vprintf("Sum of weights %13.1f\n", wsum);
                  Vprintf("Std devn and coeffts of skewness\n");
                  Vprintf("and kurtosis not defined\n");
                }
              else
                Vprintf("%s \n", fail.message);
              exit(EXIT_FAILURE);
            }
        }
    }
  exit(EXIT_SUCCESS);
}
```

8.2. Program Data

```
g01aac Example Program Data
 1
24 0
193.0  215.0  112.0  161.0  92.0   140.0  38.0   33.0   279.0  249.0
473.0  339.0  60.0   130.0  20.0   50.0   257.0  284.0  447.0  52.0
67.0   61.0   150.0  2200.0
```

8.3. Program Results

```
g01aac Example Program Results
Problem      1
Number of cases 24
Data as input -
       193.0        215.0        112.0        161.0         92.0
       140.0         38.0         33.0        279.0        249.0
       473.0        339.0         60.0        130.0         20.0
        50.0        257.0        284.0        447.0         52.0
        67.0         61.0        150.0       2200.0

Successful call of g01aac
No. of valid cases    24
Mean                 254.3
Std devn             433.5
Skewness               3.9
Kurtosis              14.7
Minimum               20.0
Maximum             2200.0
Sum of weights        24.0
```