NAG Toolbox

nag_stat_summary_onevar_combine (g01au)

1 Purpose

nag_stat_summary_onevar_combine (g01au) combines sets of summaries produced by nag_stat_summary_onevar (g01at).

2 Syntax

```
[pn, xmean, xsd, xskew, xkurt, xmin, xmax, rcomm, ifail] =
nag_stat_summary_onevar_combine(mrcomm, 'b', b)
[pn, xmean, xsd, xskew, xkurt, xmin, xmax, rcomm, ifail] = g01au(mrcomm, 'b', b)
```

Note: the interface to this routine has changed since earlier releases of the toolbox:

At Mark 25: **b** was made optional.

3 Description

Assume a dataset containing n observations, denoted by $x = \{x_i : i = 1, 2, \ldots, n\}$ and a set of weights, $w = \{w_i : i = 1, 2, \ldots, n\}$, has been split into b blocks, and each block summarised via a call to nag_stat_summary_onevar (g01at). Then nag_stat_summary_onevar_combine (g01au) takes the b communication arrays returned by nag_stat_summary_onevar (g01at) and returns the mean ($\bar{x}$), standard deviation (s_2), coefficients of skewness (s_3) and kurtosis (s_4), and the maximum and minimum values for the whole dataset.

For a definition of $\bar{x}, s_2, s_3$ and s_4 see Section 3 in nag_stat_summary_onevar (g01at).

4 References

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: **mrcomm**(**20**, **b**) – REAL (KIND=nag_wp) array

The jth column of **mrcomm** must contain the information returned in **rcomm** from one of the runs of nag_stat_summary_onevar (g01at).

5.2 Optional Input Parameters

1: **b** – INTEGER

Default: 1

b, the number of blocks the full dataset was split into.

Constraint: **b** ≥ 1.

5.3 Output Parameters

1: **pn** – INTEGER

The number of valid observations, that is the number of observations with $w_i > 0$, for $i = 1, 2, \ldots, n$.

2: **xmean** – REAL (KIND=nag_wp)

$\bar{x}$, the mean.

3: **xsd** – REAL (KIND=nag_wp)

s_2, the standard deviation.

4: **xskew** – REAL (KIND=nag_wp)

s_3, the coefficient of skewness.

5: **xkurt** – REAL (KIND=nag_wp)

s_4, the coefficient of kurtosis.

6: **xmin** – REAL (KIND=nag_wp)

The smallest value.

7: **xmax** – REAL (KIND=nag_wp)

The largest value.

8: **rcomm**(**20**) – REAL (KIND=nag_wp) array

An amalgamation of the information held in **mrcomm**. This is in the same format as **rcomm** from nag_stat_summary_onevar (g01at).

9: **ifail** – INTEGER

ifail $= 0$ unless the function detects an error (see Section 5).

6 Error Indicators and Warnings

Errors or warnings detected by the function:

ifail $= 11$

Constraint: **b** ≥ 1.

ifail $= 21$

On entry, **mrcomm** is not in the expected format.

ifail $= 31$ (*warning*)

On entry, the number of valid observations is zero.

ifail $= 51$ (*warning*)

On exit we were unable to calculate **xskew** or **xkurt**. A value of 0 has been returned.

ifail $= 52$ (*warning*)

On exit we were unable to calculate **xsd**, **xskew** or **xkurt**. A value of 0 has been returned.

ifail $= -99$

An unexpected error has been triggered by this routine. Please contact NAG.

ifail $= -399$

Your licence key may have expired or may not have been installed correctly.

ifail $= -999$

Dynamic memory allocation failed.

7 Accuracy

Not applicable.

8 Further Comments

The order that the b communication arrays are stored in **mrcomm** is arbitrary. Different orders can lead to slightly different results due to numerical accuracy of floating-point calculations.

Both nag_stat_summary_onevar_combine (g01au) and nag_stat_summary_onevar (g01at) consolidate results from multiple summaries. Whereas the former can only be used to combine summaries calculated sequentially, the latter combines summaries calculated in an arbitrary order allowing, for example, summaries calculated on different processing units to be combined.

9 Example

This example summarises some simulated data. The data is supplied in three blocks, the first consisting of 21 observations, the second 51 observations and the last 28 observations. Summaries are produced for each block of data separately and then an overall summary is produced.

9.1 Program Text

```
     function g01au_example

fprintf('g01au example results\n\n');

x1 = [-0.62; -1.92; -1.72; -6.35;  2.00;  7.65;  6.15;
       3.81;  4.87; -0.51;  6.88; -5.85; -0.72;  0.66;
       2.23; -1.61; -0.15; -1.15; -8.74; -3.94;  3.61];
wt1 = [4.91;  0.25;  3.90;  3.75;  1.17;  3.19;  2.66;
       0.02;  3.59;  3.63;  4.83;  3.72;  1.72;  0.78;
       4.74;  1.72;  3.94;  1.33;  0.51;  2.40;  3.90];
x2 = [-0.66; -2.39; -6.25;  1.23;  2.27; -2.27; 10.12;
       8.29; -2.99;  8.71; -0.74;  0.02;  1.22;  1.70;
       4.30;  2.99; -0.83; -1.00;  6.57;  2.32; -3.47;
      -1.41; -5.26;  0.53;  1.80;  4.79; -3.04;  1.20;
      -3.21; -3.75;  0.86;  1.27; -5.95; -5.27;  1.63;
       3.59; -0.01; -1.38; -4.71; -4.82;  3.55;  0.46;
       2.57;  1.76; -4.05;  1.23; -1.99;  3.20; -0.65;
       8.42; -6.01];
x3 = [ 1.13; -8.86;  5.92; -1.71; -3.99;  6.57; -2.01;
      -2.29; -1.11;  7.14;  4.84; -4.44; -3.32; 10.25;
      -2.11;  8.02; -7.31;  2.80; -1.20;  1.01;  1.37;
      -2.28;  1.28; -3.95;  3.43; -0.61; 4.85; -0.11];
data = {x1; x2; x3};

mrcomm = zeros(20,3);
% Initialise the number of valid observations processed so far
for i =1:3
  % Summarise this block of data
  if (i == 1)
    [pn, xmean, xsd, xskew, xkurt, xmin, xmax, mrcomm(:, 1), ifail] = ...
        g01at(x1, 'wt', wt1);
```

```
  else
    [pn, xmean, xsd, xskew, xkurt, xmin, xmax, mrcomm(:, i), ifail] = ...
        g01at(data{i});
  end

  % Display the results for this block
  fprintf('\nSummary for block %d\n', i);
  fprintf('%d valid observations\n', pn);
  fprintf('Mean          %13.2f\n', xmean);
  fprintf('Std devn      %13.2f\n', xsd);
  fprintf('Skewness      %13.2f\n', xskew);
  fprintf('Kurtosis      %13.2f\n', xkurt);
  fprintf('Minimum       %13.2f\n', xmin);
  fprintf('Maximum       %13.2f\n', xmax);
end

% Combine the summaries across all the blocks
[pn, xmean, xsd, xskew, xkurt, xmin, xmax, rcomm, ifail] = ...
    g01au(mrcomm);

% Display the combined results
fprintf('\nSummary for the combined data\n');
fprintf('%d valid observations\n', pn);
fprintf('Mean          %13.2f\n', xmean);
fprintf('Std devn      %13.2f\n', xsd);
fprintf('Skewness      %13.2f\n', xskew);
fprintf('Kurtosis      %13.2f\n', xkurt);
fprintf('Minimum       %13.2f\n', xmin);
fprintf('Maximum       %13.2f\n', xmax);
```

9.2 Program Results

```
     g01au example results

Summary for block 1
21 valid observations
Mean                   0.73
Std devn               4.40
Skewness              -0.05
Kurtosis              -1.00
Minimum               -8.74
Maximum                7.65

Summary for block 2
51 valid observations
Mean                   0.28
Std devn               3.96
Skewness               0.46
Kurtosis              -0.16
Minimum               -6.25
Maximum               10.12

Summary for block 3
28 valid observations
Mean                   0.48
Std devn               4.65
Skewness               0.19
Kurtosis              -0.58
Minimum               -8.86
Maximum               10.25

Summary for the combined data
100 valid observations
Mean                   0.51
```

```
Std devn                4.24
Skewness                0.18
Kurtosis               -0.59
Minimum                -8.86
Maximum                10.25
```