NAG Toolbox

nag_tsa_uni_means (g13au)

1 Purpose

nag_tsa_uni_means (g13au) calculates the range (or standard deviation) and the mean for groups of successive time series values. It is intended for use in the construction of range-mean plots.

2 Syntax

```
[y, mean, ifail] = nag_tsa_uni_means(z, m, rs, 'n', n)
[y, mean, ifail] = g13au(z, m, rs, 'n', n)
```

3 Description

Let $Z_1, Z_2, \ldots, Z_n$ denote n successive observations in a time series. The series may be divided into groups of m successive values and for each group the range or standard deviation (depending on a user-supplied option) and the mean are calculated. If n is not a multiple of m then groups of equal size m are found starting from the end of the series of observations provided, and any remaining observations at the start of the series are ignored. The number of groups used, k, is the integer part of n/m. If you wish to ensure that no observations are ignored then the number of observations, n, should be chosen so that n is divisible by m.

The mean, M_i, the range, R_i, and the standard deviation, S_i, for the ith group are defined as

$$M_i = \frac{1}{m}\sum_{j=1}^{m} Z_{l+m(i-1)+j}$$

$$R_i = \max_{1\leq j\leq m}\left\{Z_{l+m(i-1)+j}\right\} - \min_{1\leq j\leq m}\left\{Z_{l+m(i-1)+j}\right\}$$

and

$$S_i = \sqrt{\left(\frac{1}{m-1}\right)\sum_{j=1}^{m}\left(Z_{l+m(i-1)+j} - M_i\right)^2}$$

where $l = n - km$, the number of observations ignored.

For seasonal data it is recommended that m should be equal to the seasonal period. For non-seasonal data the recommended group size is 8.

A plot of range against mean or of standard deviation against mean is useful for finding a transformation of the series which makes the variance constant. If the plot appears random or the range (or standard deviation) seems to be constant irrespective of the mean level then this suggests that no transformation of the time series is called for. On the other hand an approximate linear relationship between range (or standard deviation) and mean would indicate that a log transformation is appropriate. Further details may be found in either Jenkins (1979) or McLeod (1982).

You have the choice of whether to use the range or the standard deviation as a measure of variability. If the group size is small they are both equally good but if the group size is fairly large (e.g., $m = 12$ for monthly data) then the range may not be as good an estimate of variability as the standard deviation.

4 References

Jenkins G M (1979) *Practical Experiences with Modelling and Forecasting Time Series* GJP Publications, Lancaster

McLeod G (1982) *Box–Jenkins in Practice. 1: Univariate Stochastic and Single Output Transfer Function/Noise Analysis* GJP Publications, Lancaster

5 Parameters

5.1 Compulsory Input Parameters

1: **z**(**n**) – REAL (KIND=nag_wp) array

z(t) must contain the tth observation Z_t, for $t = 1, 2, \ldots, n$.

2: **m** – INTEGER

m, the group size.

Constraint: **m** ≥ 2.

3: **rs** – CHARACTER(1)

Indicates whether ranges or standard deviations are to be calculated.

rs = 'R'
Ranges are calculated.

rs = 'S'
Standard deviations are calculated.

Constraint: **rs** = 'R' or 'S'.

5.2 Optional Input Parameters

1: **n** – INTEGER

Default: the dimension of the array **z**.

n, the number of observations in the time series.

Constraint: **n** $\geq$ **m**.

5.3 Output Parameters

1: **y**($ngrps$) – REAL (KIND=nag_wp) array

$ngrps$ = int(**n**/**m**).

y(i) contains the range or standard deviation, as determined by **rs**, of the ith group of observations, for $i = 1, 2, \ldots, k$.

2: **mean_p**($ngrps$) – REAL (KIND=nag_wp) array

$ngrps$ = int(**n**/**m**).

mean(i) contains the mean of the ith group of observations, for $i = 1, 2, \ldots, k$.

3: **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 $= 1$

On entry, **n** $<$ **m**,
or **m** < 2,
or $ngrps \neq$ integer part of **n**/**m**.

ifail $= 2$

On entry, **rs** is not equal to `R' or `S'.

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

The computations are believed to be stable.

8 Further Comments

The time taken by nag_tsa_uni_means (g13au) is approximately proportional to n.

9 Example

The following program produces the statistics for a range-mean plot for a series of 100 observations divided into groups of 8.

9.1 Program Text

```
function g13au_example

fprintf('g13au example results\n\n');

% Data
z = [101;  82;  66;  35;  31;   6;  20;  90; 154; 125;
      85;  68;  38;  23;  10;  24;  83; 133; 131; 118;
      90;  67;  60;  47;  41;  21;  16;   6;   4;   7;
      14;  34;  45;  43;  49;  42;  28;  10;   5;   2;
       0;   1;   3;  12;  14;  35;  47;  41;  30;  24;
      16;   7;   4;   2;   8;  13;  36;  50;  62;  67;
      72;  48;  29;   8;  13;  57; 122; 139; 103;  86;
      63;  37;  26;  11;  15;  40;  62;  98; 124;  96;
      65;  64;  54;  39;  21;   7;   4;  23;  53;  94;
      96;  77;  59;  44;  47;  30;  16;   7;  37;  74];

% Number of groups
m = nag_int(8);

% Calculate summary statistic
rs = 'RANGE';
[y, mean_p, ifail] = g13au( ...
                           z, m, rs);
```

```
% Display results
fprintf('     Mean      Range\n');
fprintf('   -------------------\n');
fprintf('%11.3f%11.3f\n', [mean_p y]');

fig1 = figure;
plot(mean_p,y,'+','Color','Red');
xlabel('Mean');
ylabel('Range');
title('Plot of Range vs Mean (Y vs Mean)');
```

9.2 Program Results

```
g13au example results

    Mean      Range
  -------------------
    72.375    148.000
    70.000    123.000
    43.500     84.000
    29.750     45.000
     7.625     28.000
    26.750     40.000
    30.250     65.000
    61.000    131.000
    47.625     92.000
    75.250     85.000
    46.875     92.000
    39.250     67.000
```