NAG Toolbox

nag_contab_tabulate_margin (g11bc)

1 Purpose

nag_contab_tabulate_margin (g11bc) computes a marginal table from a table computed by nag_contab_tabulate_stat (g11ba) or nag_contab_tabulate_percentile (g11bb) using a selected statistic.

2 Syntax

```
[stable, mcells, mdim, mlevel, auxt, ifail] = nag_contab_tabulate_margin(stat, table, idim, isdim, maxst, 'ncells', ncells, 'ndim', ndim)
[stable, mcells, mdim, mlevel, auxt, ifail] = g11bc(stat, table, idim, isdim, maxst, 'ncells', ncells, 'ndim', ndim)
```

3 Description

For a dataset containing classification variables (known as factors) the functions nag_contab_tabulate_stat (g11ba) and nag_contab_tabulate_percentile (g11bb) compute a table using selected statistics, for example the mean or the median. The table is indexed by the levels of the selected factors, for example if there were three factors A, B and C with 3, 2 and 4 levels respectively and the mean was to be tabulated the resulting table would be $3 \times 2 \times 4$ with each cell being the mean of all observations with the appropriate combination of levels of the three factors. In further analysis the table of means averaged over C for A and B may be required; this can be computed from the full table by taking the mean over the third dimension of the table, C.

In general, given a table computed by nag_contab_tabulate_stat (g11ba) or nag_contab_tabulate_percentile (g11bb), nag_contab_tabulate_margin (g11bc) computes a sub-table defined by a subset of the factors used to define the table such that each cell of the sub-table is the selected statistic computed over the remaining factors. The statistics that can be used are the total, the mean, the median, the variance, the smallest and the largest value.

4 References

John J A and Quenouille M H (1977) *Experiments: Design and Analysis* Griffin

Kendall M G and Stuart A (1969) *The Advanced Theory of Statistics (Volume 1)* (3rd Edition) Griffin

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: **stat** – CHARACTER(1)

Indicates which statistic is to be used to compute the marginal table.

stat = 'T'
The total.

stat = 'A'
The average or mean.

stat = 'M'
The median.

stat = 'V'
The variance.

stat = 'L'
The largest value.

stat = 'S'
The smallest value.

Constraint: **stat** = 'T', 'A', 'M', 'V', 'L' or 'S'.

2: **table**(**ncells**) – REAL (KIND=nag_wp) array

The table as computed by nag_contab_tabulate_stat (g11ba) or nag_contab_tabulate_percentile (g11bb).

3: **idim**(**ndim**) – INTEGER array

The number of levels for each dimension of **table** as returned by nag_contab_tabulate_stat (g11ba) or nag_contab_tabulate_percentile (g11bb).

Constraint: **idim**$(i) \geq 2$, for $i = 1, 2, \ldots,$ **ndim**.

4: **isdim**(**ndim**) – INTEGER array

Indicates which dimensions of **table** are to be included in the sub-table. If **isdim**$(i) > 0$ the dimension or factor indicated by **idim**(i) is to be included in the sub-table, otherwise it is excluded.

5: **maxst** – INTEGER

The maximum size of sub-table to be computed.

Constraint: **maxst** $\geq$ the product of the levels of the dimensions of **table** included in the sub-table, **stable**.

5.2 Optional Input Parameters

1: **ncells** – INTEGER

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

The number of cells in **table** as returned by nag_contab_tabulate_stat (g11ba) or nag_contab_tabulate_percentile (g11bb).

2: **ndim** – INTEGER

Default: the dimension of the arrays **idim**, **isdim**. (An error is raised if these dimensions are not equal.)

The number of dimensions for **table** as returned by nag_contab_tabulate_stat (g11ba) or nag_contab_tabulate_percentile (g11bb).

Constraint: **ndim** ≥ 2.

5.3 Output Parameters

1: **stable**(**maxst**) – REAL (KIND=nag_wp) array

The first **mcells** elements contain the sub-table computed using the statistic indicated by **stat**. The table is stored in a similar way to **table** with the **mcells** cells stored so that for any two dimensions the index relating to the dimension given later in **idim** changes faster. For further details see Section 9.

2: **mcells** – INTEGER

The number of cells in the sub-table in **stable**.

3: **mdim** – INTEGER

The number of dimensions to the sub-table in **stable**.

4: **mlevel**(**ndim**) – INTEGER array

The first **mdim** elements contain the number of levels for the dimensions of the sub-table in **stable**. The remaining elements are not referenced.

5: **auxt**(:) – REAL (KIND=nag_wp) array

The dimension of the array **auxt** will be **maxst** if **stat** = 'V' and 1 otherwise

If **stat** = 'V' **auxt** contains the sub-table of means corresponding to the sub-table of variances in **stable**. Otherwise **auxt** is not referenced.

6: **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, **ndim** < 2,
or **stat** ≠ 'T', 'A', 'M', 'V', 'L' or 'S'.

ifail = 2

On entry, **idim**(i) ≤ 1, for some $i = 1, 2, \ldots,$ **ndim**,
or **ncells** is incompatible with **idim**,
or the requested sub-table is of dimension 0,
or the requested sub-table is the full table,
or **maxst** is too small, the minimum value is returned in **mdim**.

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

Only applicable when **stat** = 'V'. In this case a one pass algorithm is used as describe in West (1979).

8 Further Comments

The sub-tables created by nag_contab_tabulate_margin (g11bc) and stored in **stable** and, depending on **stat**, also in **auxt** are stored in the following way. Let there be m dimensions defining the table with dimension k having l_k levels, then the cell defined by the levels $i_1, i_2, \ldots, i_m$ of the factors is stored in sth cell given by

$$s = 1 + \sum_{k=1}^{m} [(i_k - 1)c_k],$$

where

$$c_j = \prod_{k=j+1}^{m} l_k \quad \text{for } j = 1, 2, \ldots, n-1 \quad \text{and} \quad c_m = 1.$$

9 Example

The data, given by John and Quenouille (1977), is for 3 blocks of a 3×6 factorial experiment. The data can be considered as a $3 \times 6 \times 3$ table (i.e., blocks $\times$ treatment with 6 levels $\times$ treatment with 3 levels). This table is input and the 6×3 table of treatment means for over blocks is computed and printed.

9.1 Program Text

```
function g11bc_example

fprintf('g11bc example results\n\n');

y   = [ 274;   361;   253;   325;   317;   339;
        326;   402;   336;   379;   345;   361;
        352;   334;   318;   339;   393;   358;
        350;   340;   203;   397;   356;   298;
        382;   376;   355;   418;   387;   379;
        432;   339;   293;   322;   417;   342;
         82;   297;   133;   306;   352;   361;
        220;   333;   270;   388;   379;   274;
        336;   307;   266;   389;   333;   353];

idim   = [nag_int(3); 6; 3];
isdim  = [nag_int(0); 1; 1];
maxt   = prod(idim(isdim~=0));
maxt   = nag_int(maxt);

% Compute marginal table
stat = 'A';
[table, mcells, mdim, mlevel, auxt, ifail] = ...
  g11bc( ...
         stat, y, idim, isdim, maxt);

% Display results
fprintf(' Marginal Table\n\n');
ncol = mlevel(mdim);
nrow = mcells/ncol;
table  = transpose(reshape(table,[ncol,nrow]));
for i = 1:nrow
  fprintf('%8.2f', table(i,:));
  fprintf('\n');
end
```

9.2 Program Results

```
     g11bc example results

 Marginal Table

  235.33  332.67  196.33
  342.67  341.67  332.67
  309.33  370.33  320.33
  395.00  370.33  338.00
  373.33  326.67  292.33
  350.00  381.00  351.00
```