NAG Library Routine Document

G11BCF

Note: before using this routine, please read the Users' Note for your implementation to check the interpretation of ***bold italicised*** terms and other implementation-dependent details.

1 Purpose

G11BCF computes a marginal table from a table computed by G11BAF or G11BBF using a selected statistic.

2 Specification

```
SUBROUTINE G11BCF (STAT, TABLE, NCELLS, NDIM, IDIM, ISDIM, STABLE, MAXST,       &
                   MCELLS, MDIM, MLEVEL, AUXT, IWK, WK, IFAIL)

INTEGER            NCELLS, NDIM, IDIM(NDIM), ISDIM(NDIM), MAXST, MCELLS,       &
                   MDIM, MLEVEL(NDIM), IWK(3*NDIM), IFAIL
REAL (KIND=nag_wp) TABLE(NCELLS), STABLE(MAXST), AUXT(*), WK(NCELLS)
CHARACTER(1)       STAT
```

3 Description

For a dataset containing classification variables (known as factors) the routines G11BAF and G11BBF 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 G11BAF or G11BBF, G11BCF 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

1: STAT – CHARACTER(1) *Input*

On entry: 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 *Input*

On entry: the table as computed by G11BAF or G11BBF.

3: NCELLS – INTEGER *Input*

On entry: the number of cells in TABLE as returned by G11BAF or G11BBF.

4: NDIM – INTEGER *Input*

On entry: the number of dimensions for TABLE as returned by G11BAF or G11BBF.

Constraint: NDIM ≥ 2.

5: IDIM(NDIM) – INTEGER array *Input*

On entry: the number of levels for each dimension of TABLE as returned by G11BAF or G11BBF.

Constraint: IDIM$(i) \geq 2$, for $i = 1, 2, \ldots,$ NDIM.

6: ISDIM(NDIM) – INTEGER array *Input*

On entry: 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.

7: STABLE(MAXST) – REAL (KIND=nag_wp) array *Output*

On exit: 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 8.

8: MAXST – INTEGER *Input*

On entry: 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.

9: MCELLS – INTEGER *Output*

On exit: the number of cells in the sub-table in STABLE.

10: MDIM – INTEGER *Output*

On exit: the number of dimensions to the sub-table in STABLE.

11: MLEVEL(NDIM) – INTEGER array *Output*

On exit: the first MDIM elements contain the number of levels for the dimensions of the sub-table in STABLE. The remaining elements are not referenced.

12: AUXT(∗) – REAL (KIND=nag_wp) array *Output*

Note: the dimension of the array AUXT must be at least MAXST if STAT = 'V', and at least 1 otherwise.

On exit: if STAT = 'V' AUXT contains the sub-table of means corresponding to the sub-table of variances in STABLE. Otherwise AUXT is not referenced.

13: IWK(3 × NDIM) – INTEGER array *Workspace*

14: WK(NCELLS) – REAL (KIND=nag_wp) array *Workspace*

15: IFAIL – INTEGER *Input/Output*

On entry: IFAIL must be set to 0, −1 or 1. If you are unfamiliar with this parameter you should refer to Section 3.3 in the Essential Introduction for details.

For environments where it might be inappropriate to halt program execution when an error is detected, the value −1 or 1 is recommended. If the output of error messages is undesirable, then the value 1 is recommended. Otherwise, if you are not familiar with this parameter, the recommended value is 0. **When the value −1 or 1 is used it is essential to test the value of IFAIL on exit.**

On exit: IFAIL = 0 unless the routine detects an error or a warning has been flagged (see Section 6).

6 Error Indicators and Warnings

If on entry IFAIL = 0 or −1, explanatory error messages are output on the current error message unit (as defined by X04AAF).

Errors or warnings detected by the routine:

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.

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 G11BCF 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

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

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

```
    Program g11bcfe

!     G11BCF Example Program Text

!     Mark 24 Release. NAG Copyright 2012.

!     .. Use Statements ..
      Use nag_library, Only: g11bcf, nag_wp
!     .. Implicit None Statement ..
      Implicit None
!     .. Parameters ..
      Integer, Parameter               :: nin = 5, nout = 6
!     .. Local Scalars ..
      Integer                          :: i, ifail, k, lauxt, maxst, mcells,   &
                                          mdim, ncells, ncol, ndim, nrow
      Character (1)                    :: stat
!     .. Local Arrays ..
      Real (Kind=nag_wp), Allocatable  :: auxt(:), stable(:), table(:), wk(:)
      Integer, Allocatable             :: idim(:), isdim(:), iwk(:), mlevel(:)
!     .. Executable Statements ..
      Write (nout,*) 'G11BCF Example Program Results'
      Write (nout,*)

!     Skip heading in data file
      Read (nin,*)

!     Read in the problem size
      Read (nin,*) stat, ncells, ndim

      Allocate (table(ncells),idim(ndim),isdim(ndim))

!     Read in data
      Read (nin,*) table(1:ncells)
      Read (nin,*) idim(1:ndim)
      Read (nin,*) isdim(1:ndim)

!     Calculate MAXST
      maxst = 1
      Do i = 1, ndim
        If (isdim(i)>0) Then
          maxst = maxst*idim(i)
        End If
      End Do

      If (stat=='V' .Or. stat=='v') Then
        lauxt = maxst
      Else
        lauxt = 0
      End If
      Allocate (stable(maxst),mlevel(ndim),auxt(lauxt),iwk(3*ndim),wk(ncells))

!     Compute marginal table
      ifail = 0
      Call g11bcf(stat,table,ncells,ndim,idim,isdim,stable,maxst,mcells,mdim, &
        mlevel,auxt,iwk,wk,ifail)
```

```
!     Display results
      Write (nout,*) ' Marginal Table'
      Write (nout,*)
      ncol = mlevel(mdim)
      nrow = mcells/ncol
      k = 1
      Do i = 1, nrow
        Write (nout,99999) stable(k:(k+ncol-1))
        k = k + ncol
      End Do

99999 Format (10F8.2)
    End Program g11bcfe
```

9.2 Program Data

```
G11BCF Example Program Data
'A' 54 3
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
3 6 3
0 1 1
```

9.3 Program Results

```
 G11BCF Example Program 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
```