NAG Library Routine Document

G04DAF

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

G04DAF computes sum of squares for a user-defined contrast between means.

2 Specification

```
SUBROUTINE G04DAF (NT, TMEAN, IREP, RMS, RDF, NC, CT, LDCT, EST, TABL,     &
                   LDTABL, TOL, USETX, TX, IFAIL)

INTEGER            NT, IREP(NT), NC, LDCT, LDTABL, IFAIL
REAL (KIND=nag_wp) TMEAN(NT), RMS, RDF, CT(LDCT,NC), EST(NC),              &
                   TABL(LDTABL,*), TOL, TX(NT)
LOGICAL            USETX
```

3 Description

In the analysis of designed experiments the first stage is to compute the basic analysis of variance table, the estimate of the error variance (the residual or error mean square), $\hat{\sigma}^2$, and the (variance ratio) F-statistic for the t treatments. If this F-test is significant then the second stage of the analysis is to explore which treatments are significantly different.

If there is a structure to the treatments then this may lead to hypotheses that can be defined before the analysis and tested using linear contrasts. For example, if the treatments were three different fixed temperatures, say 18, 20 and 22, and an uncontrolled temperature (denoted by N) then the following contrasts might be of interest.

	18	20	22	N
(a)	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	-1
(b)	-1	0	1	0

The first represents the average difference between the controlled temperatures and the uncontrolled temperature. The second represents the linear effect of an increasing fixed temperature.

For a randomized complete block design or a completely randomized design, let the treatment means be $\hat{\tau}_i$, $i = 1, 2, \ldots, t$, and let the jth contrast be defined by λ_{ij}, $i = 1, 2, \ldots, t$, then the estimate of the contrast is simply:

$$\Lambda_j = \sum_{i=1}^{t} \hat{\tau}_i \lambda_{ij}$$

and the sum of squares for the contrast is:

$$\mathrm{SS}_j = \frac{\Lambda_j^2}{\sum_{i=1}^{t} \lambda_{ij}^2 / n_i} \tag{1}$$

where n_i is the number of observations for the ith treatment. Such a contrast has one degree of freedom so that the appropriate F-statistic is $\mathrm{SS}_j / \hat{\sigma}^2$.

The two contrasts λ_{ij} and $\lambda_{ij'}$ are orthogonal if $\sum_{i=1}^{t}\lambda_{ij}\lambda_{ij'} = 0$ and the contrast λ_{ij} is orthogonal to the overall mean if $\sum_{i=1}^{t}\lambda_{ij} = 0$. In practice these sums will be tested against a small quantity, ϵ. If each of a set of contrasts is orthogonal to the mean and they are all mutually orthogonal then the contrasts provide a partition of the treatment sum of squares into independent components. Hence the resulting F-tests are independent.

If the treatments come from a design in which treatments are not orthogonal to blocks then the sum of squares for a contrast is given by:

$$\mathrm{SS}_j = \frac{\Lambda_j \Lambda_j^*}{\sum_{i=1}^{t}\lambda_{ij}^2/n_i} \tag{2}$$

where

$$\Lambda_j^* = \sum_{i=1}^{t}\tau_i^*\lambda_{ij}$$

with τ_i^*, for $i = 1, 2, \ldots, t$, being adjusted treatment means computed by first eliminating blocks then computing the treatment means from the block adjusted observations without taking into account the non-orthogonality between treatments and blocks. For further details see John (1987).

4 References

Cochran W G and Cox G M (1957) *Experimental Designs* Wiley

John J A (1987) *Cyclic Designs* Chapman and Hall

Winer B J (1970) *Statistical Principles in Experimental Design* McGraw–Hill

5 Arguments

1: NT – INTEGER *Input*

On entry: t, the number of treatment means.

Constraint: NT ≥ 2.

2: TMEAN(NT) – REAL (KIND=nag_wp) array *Input*

On entry: the treatment means, $\hat{\tau}_i$, for $i = 1, 2, \ldots, t$.

3: IREP(NT) – INTEGER array *Input*

On entry: the replication for each treatment mean, n_i, for $i = 1, 2, \ldots, t$.

4: RMS – REAL (KIND=nag_wp) *Input*

On entry: the residual mean square, $\hat{\sigma}^2$.

Constraint: RMS > 0.0.

5: RDF – REAL (KIND=nag_wp) *Input*

On entry: the residual degrees of freedom.

Constraint: RDF ≥ 1.0.

6: NC – INTEGER *Input*

On entry: the number of contrasts.

Constraint: NC ≥ 1.

7: CT(LDCT, NC) – REAL (KIND=nag_wp) array *Input*

On entry: the columns of CT must contain the NC contrasts, that is CT(i, j) must contain λ_{ij}, for $i = 1, 2, \ldots, t$ and $j = 1, 2, \ldots,$ NC.

8: LDCT – INTEGER *Input*

On entry: the first dimension of the array CT as declared in the (sub)program from which G04DAF is called.

Constraint: LDCT $\geq$ NT.

9: EST(NC) – REAL (KIND=nag_wp) array *Output*

On exit: the estimates of the contrast, Λ_j, for $j = 1, 2, \ldots,$ NC.

10: TABL(LDTABL, $*$) – REAL (KIND=nag_wp) array *Input/Output*

Note: the second dimension of the array TABL must be at least 5.

On entry: the elements of TABL that are not referenced as described below remain unchanged.

On exit: the rows of the analysis of variance table for the contrasts. For each row column 1 contains the degrees of freedom, column 2 contains the sum of squares, column 3 contains the mean square, column 4 the F-statistic and column 5 the significance level for the contrast. Note that the degrees of freedom are always one and so the mean square equals the sum of squares.

11: LDTABL – INTEGER *Input*

On entry: the first dimension of the array TABL as declared in the (sub)program from which G04DAF is called.

Constraint: LDTABL $\geq$ NC.

12: TOL – REAL (KIND=nag_wp) *Input*

On entry: the tolerance, ϵ used to check if the contrasts are orthogonal and if they are orthogonal to the mean. If TOL ≤ 0.0 the value ***machine precision*** is used.

13: USETX – LOGICAL *Input*

On entry: if USETX = .TRUE. the means τ_i^* are provided in TX and the formula (2) is used instead of formula (1).

If USETX = .FALSE. formula (1) is used and TX is not referenced.

14: TX(NT) – REAL (KIND=nag_wp) array *Input*

On entry: if USETX = .TRUE. TX must contain the means τ_i^*, for $i = 1, 2, \ldots, t$.

15: IFAIL – INTEGER *Input/Output*

On entry: IFAIL must be set to 0, -1 or 1. If you are unfamiliar with this argument you should refer to Section 3.4 in How to Use the NAG Library and its Documentation 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, because for this routine the values of the output arguments may be useful even if IFAIL $\neq 0$ on exit, the recommended value is -1. **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).

Note: G04DAF may return useful information for one or more of the following detected errors or warnings.

Errors or warnings detected by the routine:

IFAIL $= 1$

On entry, NC < 1,
or NT < 2,
or LDCT $<$ NT,
or LDTABL $<$ NC,
or RMS ≤ 0.0,
or RDF < 1.0.

IFAIL $= 2$

On entry, a contrast is not orthogonal to the mean,
or at least two contrasts are not orthogonal.

If IFAIL $= 2$ full results are returned but they should be interpreted with care.

IFAIL $= -99$

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

See Section 3.9 in How to Use the NAG Library and its Documentation for further information.

IFAIL $= -399$

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

See Section 3.8 in How to Use the NAG Library and its Documentation for further information.

IFAIL $= -999$

Dynamic memory allocation failed.

See Section 3.7 in How to Use the NAG Library and its Documentation for further information.

7 Accuracy

The computations are stable.

8 Parallelism and Performance

G04DAF is not threaded in any implementation.

9 Further Comments

If the treatments have a factorial structure G04CAF should be used and if the treatments have no structure the means can be compared using G04DBF.

10 Example

The data is from a completely randomized experiment on potato scab with seven treatments representing amounts of sulphur applied, whether the application was in spring or autumn and a control treatment. The one-way anova is computed using G02BBF. Two contrasts are analysed, one comparing the control with use of sulphur, the other comparing spring with autumn application.

10.1 Program Text

```
    Program g04dafe

!     G04DAF Example Program Text

!     Mark 26 Release. NAG Copyright 2016.

!     .. Use Statements ..
      Use nag_library, Only: g04bbf, g04daf, nag_wp
!     .. Implicit None Statement ..
      Implicit None
!     .. Parameters ..
      Integer, Parameter               :: nin = 5, nout = 6
!     .. Local Scalars ..
      Real (Kind=nag_wp)               :: gmean, rdf, rms, tol
      Integer                          :: i, iblock, ifail, irdf, ldc, ldct,   &
                                          ldtabl, lit, n, nc, nt
      Logical                          :: usetx
!     .. Local Arrays ..
      Real (Kind=nag_wp), Allocatable  :: bmean(:), c(:,:), ct(:,:), ef(:),    &
                                          est(:), r(:), tabl(:,:), tmean(:),   &
                                          tx(:), wk(:), y(:)
      Integer, Allocatable             :: irep(:), it(:)
      Character (11), Allocatable      :: names(:)
!     .. Intrinsic Procedures ..
      Intrinsic                        :: abs
!     .. Executable Statements ..
      Write (nout,*) 'G04DAF Example Program Results'
      Write (nout,*)

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

!     Read in problem size for G04BBF
      Read (nin,*) n, nt, iblock

      ldc = nt
      If (nt>1) Then
        lit = n
      Else
        lit = 1
      End If
      Allocate (y(n),bmean(abs(iblock)),tmean(nt),irep(nt),c(ldc,nt),r(n),     &
        ef(nt),wk(3*nt),it(lit))

!     Read in the data and plot information for G04BBF
      Read (nin,*) y(1:n)
      If (nt>1) Then
        Read (nin,*) it(1:n)
      End If

!     Don't use TX when calling G04DAF
      usetx = .False.

!     Read in the number of contrasts
      Read (nin,*) nc

!     Using first 4 rows of TABL in G04BBF next NC rows in  G04DAF
      ldtabl = nc + 4

      ldct = nt
```

```
      Allocate (ct(ldct,nc),est(nc),tabl(ldtabl,5),tx(nt),names(nc))

!     Read in the contrasts and their names
      Do i = 1, nc
        Read (nin,*) ct(1:nt,i)
        Read (nin,99999) names(i)
      End Do

!     Use default tolerance
      tol = 0.0E0_nag_wp

!     Use standard degrees of freedom
      irdf = 0

!     Calculate the ANOVA table
      ifail = 0
      Call g04bbf(n,y,iblock,nt,it,gmean,bmean,tmean,tabl,ldtabl,c,ldc,irep,r, &
        ef,tol,irdf,wk,ifail)

!     Display results from G04BBF
      Write (nout,*) ' ANOVA table'
      Write (nout,*)
      Write (nout,*) '  Source        df         SS          MS          F',   &
        '        Prob'
      Write (nout,*)
      If (iblock>1) Then
        Write (nout,99998) ' Blocks    ', tabl(1,1:5)
      End If
      Write (nout,99998) ' Treatments', tabl(2,1:5)
      Write (nout,99998) ' Residual  ', tabl(3,1:3)
      Write (nout,99998) ' Total     ', tabl(4,1:2)
      Write (nout,*)

!     Extract the residual mean square and degrees of freedom from ANOVA
!     table
      rms = tabl(3,3)
      rdf = tabl(3,1)

!     Compute sums of squares for contrast
      ifail = -1
      Call g04daf(nt,tmean,irep,rms,rdf,nc,ct,ldct,est,tabl(5,1),ldtabl,tol,   &
        usetx,tx,ifail)
      If (ifail/=0) Then
        If (ifail/=2) Then
          Go To 100
        End If
      End If

!     Display results from G04DAF
      Write (nout,*) ' Orthogonal Contrasts'
      Write (nout,*)
      Write (nout,99998)(names(i),tabl(i+4,1:5),i=1,nc)

100   Continue

99999 Format (A)
99998 Format (A,3X,F3.0,2X,F10.1,2X,F10.1,2X,F10.3,2X,F9.4)
    End Program g04dafe
```

10.2 Program Data

```
G04DAF Example Program Data
32 7 1                              :: N, NT, IBLOCK (see G04BBF)
12 10 24 29 30 18 32 26
9 9 16 4 30 7 21 9 16 10 18 18
18 24 12 19 10 4 4 5 17 7 16 17     :: End of Y (see G04BBF)
1 1 1 1 1 1 1 1
2 2 2 2 3 3 3 3 4 4 4 4
5 5 5 5 6 6 6 6 7 7 7 7             :: End of IT (see G04BBF)
```

```
2                                    :: NC
6 -1 -1 -1 -1 -1 -1                  :: First contrast
 Cntl v S                            :: Name of first contrast
0 1 -1 1 -1 1 -1                     :: Second contrast
 Spring v A                          :: Name of second contrast
```

10.3 Program Results

```
 G04DAF Example Program Results

  ANOVA table

   Source        df         SS          MS          F        Prob

 Treatments     6.        972.3       162.1       3.608     0.0103
 Residual      25.       1122.9        44.9
 Total         31.       2095.2

  Orthogonal Contrasts

 Cntl v S       1.        518.0       518.0      11.533     0.0023
 Spring v A     1.        228.2       228.2       5.080     0.0332
```