NAG Toolbox for MATLAB

g04da

1 Purpose

g04da computes sum of squares for a user-defined contrast between means.

2 Syntax

```
[est, tabl, ifail] = g04da(tmean, irep, rms, rdf, ct, tabl, tol, usetx, tx, 'nt', nt, 'nc', nc)
```

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 and Morgan 1993.

4 References

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

John J A 1987 *Cyclic Designs* Chapman and Hall

Morgan G W 1993 Analysis of variance using the NAG Fortran Library: Examples from Cochran and Cox *NAG Technical Report TR 3/93* NAG Ltd, Oxford

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

5 Parameters

5.1 Compulsory Input Parameters

1: **tmean(nt) – double array**

The treatment means, $\hat{\tau}_i$, for $i = 1, 2, \ldots, t$.

2: **irep(nt) – int32 array**

The replication for each treatment mean, n_i, for $i = 1, 2, \ldots, t$.

3: **rms – double scalar**

The residual mean square, $\hat{\sigma}^2$.

Constraint: **rms** > 0.0.

4: **rdf – double scalar**

The residual degrees of freedom.

Constraint: **rdf** ≥ 1.0.

5: **ct(ldct,nc) – double array**

ldct, the first dimension of the array, must be at least **nt**.

The columns of **ct** must contain the **nc** contrasts, that is **ct**(i,j) must contain λ_{ij}, for $i = 1, 2, \ldots, t$; $j = 1, 2, \ldots,$ **nc**.

6: **tabl(ldtabl,$*$) – double array**

The first dimension of the array **tabl** must be at least **nc**

The second dimension of the array must be at least 5

The elements of **tabl** that are not referenced as described below remain unchanged.

7: **tol – double scalar**

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.

8: **usetx – logical scalar**

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.

9: **tx**(**nt**) **– double array**

If **usetx** = **true** **tx** must contain the means τ_i^*, for $i = 1, 2, \ldots, t$.

5.2 Optional Input Parameters

1: **nt – int32 scalar**

Default: The dimension of the arrays **irep**, **tmean**, **tx**. (An error is raised if these dimensions are not equal.)

t, the number of treatment means.

Constraint: **nt** ≥ 2.

2: **nc – int32 scalar**

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

the number of contrasts.

Constraint: **nc** ≥ 1.

5.3 Input Parameters Omitted from the MATLAB Interface

ldct, ldtabl

5.4 Output Parameters

1: **est**(**nc**) **– double array**

The estimates of the contrast, Λ_j, for $j = 1, 2, \ldots,$ **nc**.

2: **tabl**(**ldtabl**,$*$) **– double array**

The first dimension of the array **tabl** must be at least **nc**

The second dimension of the array must be at least 5

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.

3: **ifail – int32 scalar**

0 unless the function detects an error (see Section 6).

6 Error Indicators and Warnings

Note: g04da may return useful information for one or more of the following detected errors or warnings.

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.

7 Accuracy

The computations are stable.

8 Further Comments

If the treatments have a factorial structure g04ca should be used and if the treatments have no structure the means can be compared using g04db.

9 Example

```
tmean = [22.625;
     9.5;
     16.75;
     15.5;
     18.25;
     5.75;
     14.25];
irep = [int32(8);
     int32(4);
     int32(4);
     int32(4);
     int32(4);
     int32(4);
     int32(4)];
rms = 44.915;
rdf = 25;
ct = [6, 0;
     -1, 1;
     -1, -1;
     -1, 1;
     -1, -1;
     -1, 1;
     -1, -1];
table = zeros(2, 5);
tol = 5e-06;
usetx = false;
tx = zeros(7, 1);
[est, tableOut, ifail] = g04da(tmean, irep, rms, rdf, ct, table, tol,
usetx, tx)
```

```
est =
   55.7500
  -18.5000
tableOut =
    1.0000  518.0104  518.0104   11.5331    0.0023
    1.0000  228.1667  228.1667    5.0800    0.0332
```

```
ifail =
           0
```