NAG Toolbox

nag_anova_confidence (g04db)

1 Purpose

nag_anova_confidence (g04db) computes simultaneous confidence intervals for the differences between means. It is intended for use after nag_anova_random (g04bb) or nag_anova_rowcol (g04bc).

2 Syntax

```
[cil, ciu, isig, ifail] = nag_anova_confidence(typ, tmean, rdf, c, clevel, 'nt',
nt)
[cil, ciu, isig, ifail] = g04db(typ, tmean, rdf, c, clevel, 'nt', nt)
```

3 Description

In the computation of analysis of a designed experiment the first stage is to compute the basic analysis of variance table, the estimate of the error variance (the residual or error mean square), $\hat{\sigma}^2$, the residual degrees of freedom, ν, and the (variance ratio) F-statistic for the t treatments. The second stage of the analysis is to compare the treatment means. If the treatments have no structure, for example the treatments are different varieties, rather than being structured, for example a set of different temperatures, then a multiple comparison procedure can be used.

A multiple comparison procedure looks at all possible pairs of means and either computes confidence intervals for the difference in means or performs a suitable test on the difference. If there are t treatments then there are $t(t-1)/2$ comparisons to be considered. In tests the type 1 error or significance level is the probability that the result is considered to be significant when there is no difference in the means. If the usual t-test is used with, say, a 6% significance level then the type 1 error for all $k = t(t-1)/2$ tests will be much higher. If the tests were independent then if each test is carried out at the 100α percent level then the overall type 1 error would be $\alpha^* = 1 - (1-\alpha)^k \simeq k\alpha$. In order to provide an overall protection the individual tests, or confidence intervals, would have to be carried out at a value of α such that α^* is the required significance level, e.g., five percent.

The $100(1-\alpha)$ percent confidence interval for the difference in two treatment means, $\hat{\tau}_i$ and $\hat{\tau}_j$ is given by

$$\left(\hat{\tau}_i - \hat{\tau}_j\right) \pm T^*_{(\alpha,\nu,t)} se\left(\hat{\tau}_i - \hat{\tau}_j\right),$$

where $se()$ denotes the standard error of the difference in means and $T^*_{(\alpha,\nu,t)}$ is an appropriate percentage point from a distribution. There are several possible choices for $T^*_{(\alpha,\nu,t)}$. These are:

(a) $\frac{1}{2}q_{(1-\alpha,\nu,t)}$, the studentized range statistic, see nag_stat_inv_cdf_studentized_range (g01fm). It is the appropriate statistic to compare the largest mean with the smallest mean. This is known as Tukey–Kramer method.

(b) $t_{(\alpha/k,\nu)}$, this is the Bonferroni method.

(c) $t_{(\alpha_0,\nu)}$, where $\alpha_0 = 1 - (1-\alpha)^{1/k}$, this is known as the Dunn–Sidak method.

(d) $t_{(\alpha,\nu)}$, this is known as Fisher's LSD (least significant difference) method. It should only be used if the overall F-test is significant, the number of treatment comparisons is small and were planned before the analysis.

(e) $\sqrt{(k-1)F_{1-\alpha,k-1,\nu}}$ where $F_{1-\alpha,k-1,\nu}$ is the deviate corresponding to a lower tail probability of $1-\alpha$ from an F-distribution with $k-1$ and ν degrees of freedom. This is Scheffe's method.

In cases (b), (c) and (d), $t_{(\alpha,\nu)}$ denotes the α two tail significance level for the Student's t-distribution with ν degrees of freedom, see nag_stat_inv_cdf_students_t (g01fb).

The Scheffe method is the most conservative, followed closely by the Dunn–Sidak and Tukey–Kramer methods.

To compute a test for the difference between two means the statistic,

$$\frac{\hat{\tau}_i - \hat{\tau}_j}{se(\hat{\tau}_i - \hat{\tau}_j)}$$

is compared with the appropriate value of $T^*_{(\alpha,\nu,t)}$.

4 References

Kotz S and Johnson N L (ed.) (1985a) Multiple range and associated test procedures *Encyclopedia of Statistical Sciences* **5** Wiley, New York

Kotz S and Johnson N L (ed.) (1985b) Multiple comparison *Encyclopedia of Statistical Sciences* **5** Wiley, New York

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

5 Parameters

5.1 Compulsory Input Parameters

1: **typ** – CHARACTER(1)

Indicates which method is to be used.

typ = 'T'
The Tukey–Kramer method is used.

typ = 'B'
The Bonferroni method is used.

typ = 'D'
The Dunn–Sidak method is used.

typ = 'L'
The Fisher LSD method is used.

typ = 'S'
The Scheffe's method is used.

Constraint: **typ** = 'T', 'B', 'D', 'L' or 'S'.

2: **tmean**(**nt**) – REAL (KIND=nag_wp) array

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

3: **rdf** – REAL (KIND=nag_wp)

ν, the residual degrees of freedom.

Constraint: **rdf** ≥ 1.0.

4: **c**(ldc, **nt**) – REAL (KIND=nag_wp) array

ldc, the first dimension of the array, must satisfy the constraint $ldc \geq$ **nt**.

The strictly lower triangular part of **c** must contain the standard errors of the differences between the means as returned by nag_anova_random (g04bb) and nag_anova_rowcol (g04bc). That is **c**(i,j), $i > j$, contains the standard error of the difference between the ith and jth mean in **tmean**.

Constraint: **c**$(i,j) > 0.0$, for $i = 2, 3, \ldots, t$ and $j = 1, 2, \ldots, i - 1$.

5: **clevel** – REAL (KIND=nag_wp)

The required confidence level for the computed intervals, $(1 - \alpha)$.

Constraint: $0.0 <$ **clevel** < 1.0.

5.2 Optional Input Parameters

1: **nt** – INTEGER

Default: the dimension of the array **tmean** and the first dimension of the array **c** and the second dimension of the array **c**. (An error is raised if these dimensions are not equal.)

t, the number of treatment means.

Constraint: **nt** ≥ 2.

5.3 Output Parameters

1: **cil**(**nt** $\times$ (**nt** $-$ **1**)/**2**) – REAL (KIND=nag_wp) array

The $((i-1)(i-2)/2+j)$th element contains the lower limit to the confidence interval for the difference between ith and jth means in **tmean**, for $i = 2, 3, \ldots, t$ and $j = 1, 2, \ldots, i-1$.

2: **ciu**(**nt** $\times$ (**nt** $-$ **1**)/**2**) – REAL (KIND=nag_wp) array

The $((i-1)(i-2)/2+j)$th element contains the upper limit to the confidence interval for the difference between ith and jth means in **tmean**, for $i = 2, 3, \ldots, t$ and $j = 1, 2, \ldots, i-1$.

3: **isig**(**nt** $\times$ (**nt** $-$ **1**)/**2**) – INTEGER array

The $((i-1)(i-2)/2+j)$th element indicates if the difference between ith and jth means in **tmean** is significant, for $i = 2, 3, \ldots, t$ and $j = 1, 2, \ldots, i-1$. If the difference is significant then the returned value is 1; otherwise the returned value is 0.

4: **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, **nt** < 2,
or $ldc <$ **nt**,
or **rdf** < 1.0,
or **clevel** ≤ 0.0,
or **clevel** ≥ 1.0,
or **typ** $\neq$ 'T', 'B', 'D', 'L' or 'S'.

ifail $= 2$

On entry, **c**$(i,j) \leq 0.0$ for some i, j, $i = 2, 3, \ldots, t$ and $j = 1, 2, \ldots, i-1$.

ifail $= 3$

There has been a failure in the computation of the studentized range statistic. This is an unlikely error. Try using a small value of **clevel**.

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

For the accuracy of the percentage point statistics see nag_stat_inv_cdf_students_t (g01fb) and nag_stat_inv_cdf_studentized_range (g01fm).

8 Further Comments

If the treatments have a structure then the use of linear contrasts as computed by nag_anova_contrasts (g04da) may be more appropriate.

An alternative approach to one used in nag_anova_confidence (g04db) is the sequential testing of the Student–Newman–Keuls procedure. This, in effect, uses the Tukey–Kramer method but first ordering the treatment means and examining only subsets of the treatment means in which the largest and smallest are significantly different. At each stage the third argument of the Studentized range statistic is the number of means in the subset rather than the total number of means.

9 Example

In the example taken from Winer (1970) a completely randomized design with unequal treatment replication is analysed using nag_anova_random (g04bb) and then confidence intervals are computed by nag_anova_confidence (g04db) using the Tukey–Kramer method.

9.1 Program Text

```
function g04db_example

fprintf('g04db example results\n\n');

n1 = nag_int(1);
iblock = n1;
nt = 4*n1;
y  = [ 3  2  4  3  1  5  7  8  4 10 ...
       6  3  2  1  2  4  2  3  1 10 ...
      12  8  5 12 10  9];
it = [n1  1  1  1  1  1  2  2  2  2 ...
       2  3  3  3  3  3  3  3  3  4 ...
       4  4  4  4  4  4];

% Calculate ANOVA table
tol  = 0;
irdf = nag_int(0);
[gmean, bmean, tmean, table, c, irep, r, ef, ifail] = ...
  g04bb( ...
         y, iblock, nt, it, tol, irdf);

% Display ANOVA results
fprintf('ANOVA table\n\n');
fprintf(' Source        df         SS          MS          F        Prob\n\n');
fmt5 = '%s%5.0f%12.1f%12.1f%12.3f%11.4f\n';
fmt3 = '%s%5.0f%12.1f%12.1f\n';
fmt2 = '%s%5.0f%12.1f\n';
if iblock > 1
  fprintf(fmt5, 'Blocks      ', table(1,1:5));
end
fprintf(fmt5, 'Treatments  ', table(2,1:5));
fprintf(fmt3, 'Residual    ', table(3,1:3));
fprintf(fmt2, 'Total       ', table(4,1:2));
```

```
fprintf('\nTreatment Means\n\n');
for j = 1:8:nt
  fprintf('%8.3f', tmean(j:min(j+7,nt)));
  fprintf('\n');
end
fprintf('\n');

% Extract the residual degrees of freedom
rdf = table(3,1);

% Calculate simultaneous CIs
typ = 'T';
clevel = 0.95;
[cil, ciu, isig, ifail] = g04db( ...
                                  typ, tmean, rdf, c, clevel);

% Display CI results
fprintf('\nSimultaneous Confidence Intervals\n\n');
star(2) = '*';
star(1) = ' ';
ij = 0;
for i = 1:nt
  for j = 1:i-1
    ij = ij + 1;
    fprintf(' %2d%2d%15.3f%15.3f%5s\n', i, j, cil(ij), ciu(ij), ...
            star(isig(ij)+1));
  end
end
```

9.2 Program Results

```
     g04db example results

ANOVA table

 Source        df          SS          MS           F        Prob

Treatments      3       239.9        80.0      24.029     0.0000
Residual       22        73.2         3.3
Total          25       313.1

Treatment Means

   3.000   7.000   2.250   9.429

Simultaneous Confidence Intervals

  2 1          0.933          7.067    *
  3 1         -3.486          1.986
  3 2         -7.638         -1.862    *
  4 1          3.610          9.247    *
  4 2         -0.538          5.395
  4 3          4.557          9.800    *
```