In an attempt to improves its clarity, this document someimes refers back to the parameters in the NAGRepeatedANOVA routines. In these cases notation of the form, [**r**], is used, which in this case, indicates the parameter 'r' in this case. In addition, differences between the NAG implementation at that supplied within SAS and SPSS are highlighted.

1 Mauchly's W Statistic, Greenhouse-Geisser, Huynh-Feldt and Lower-Bound

With two minor exception, the W statistc and all three of the degrees of freedom adjustments are calculated using the same formula as SAS and SPSS, that is:

Mauchly's W Statistic

$$\begin{aligned} W &= \frac{\det(A)}{(\mathrm{trace}(A)/d)^d} \\ \chi^2 &= \left(\frac{2d^2+d+2}{6d} - n - r_X\right)\log(W) \\ df &= \frac{d(d+1)}{2} - 1 \end{aligned}$$

Greenhouse-Geisser

$$\epsilon_{gg} = \frac{(\mathrm{trace}(A))^2}{d \times \mathrm{trace}(A^T A)}$$

Huynh-Feldt

$$\epsilon_{hf} = \min\left(\frac{nd\epsilon_{gg} - 2}{d(n - r_X) - d^2\epsilon_{gg}}, 1\right)$$

Lower-Bound

$$\epsilon_{lb} = \frac{1}{d}$$

where $A = M^T SM$, S is the error sums of squares and cross product matrix, M is an orthonormal matrix associated with the within-subject effects being tested, r_X is the rank of the design matrix, and n is the number of subjects. For main effects, d is $l-1$, where l is the number of levels for the variable being tested. For the two-way interactions, $d = (l_1 - 1)(l_2 - 1)$ where l_1 and l_2 are the respective number of levels for the two variables making up the interaction. The values of l, l_1 and l_2 can be those supplied in any of [**r**], [**nlx1**] or [**nlx2**], depending on the comparisons being made.

In some of its documentation SPSS quotes a formula for calculating the p-value associated with chi-squared statistic given above for Mauchly's W Statistic. The p-value returned by these routines is based on the standard chi-squared distribution using the statistic and p-value given above, rather then the formula given in the SPSS documentation.

When calculating the Huynh-Feldt adjustment (ϵ_{hf}), SAS allows values > 1, i.e. the minimum in the above formula is not used. SPSS does limit ϵ_{hf}) to be less than or equal to 1.

2 Pillai's Trace, Wilks Lambda, Hotelling's Trace and Roy's Largest Root

With one exception, all of the these statistics use the same formula as SAS and SPSS, that is:

- Pillai's Trace

$$\begin{aligned} T &= \sum_{i=1}^{c} \frac{\lambda_1}{1+\lambda_1} \\ df_1 &= s(2m^* + s + 1) \\ df_2 &= s(2n^* + s + 1) \\ F &= \frac{df_2}{df_1} \frac{T}{(s - T)} \end{aligned}$$

- Wilks Lambda

$$\begin{aligned}
T &= \prod_{i=1}^{c} \frac{1}{1+\lambda_i} \\
df_1 &= n_H d \\
df_2 &= \eta\tau - \nu \\
\eta &= n_E - \frac{d - n_H + 1}{2} \\
\tau &= \begin{cases} \sqrt{\frac{(n_H^2 d^2 - 4)}{n_H^2 + d^2 - 5}} & \text{if } (n_H^2 + d^2 - 5) > 0 \\ 1 & \text{otherwise} \end{cases} \\
\nu &= \frac{n_H d - 2}{2} \\
F &= \frac{df_2}{df_1} \frac{(1 - T^{1/\tau})}{T^{1/\tau}}
\end{aligned}$$

- Hotelling's Trace

$$T = \sum_{i=1}^{c} \lambda_i \tag{1}$$

if $(n^* > 1)$

$$\begin{aligned}
df_1 &= n_H d \\
df_2 &= 4 + \frac{n_H d + 2}{\eta} \\
\eta &= \frac{(d + 2n^*)(n_H + 2n^*)}{2(2n^* + 1)(n^* - 1)} - 1 \\
\tau &= \frac{2\eta + n_H d + 2)}{2\eta n^*} \\
F &= \frac{df_2}{df_1} \frac{T}{\tau}
\end{aligned}$$

else

$$\begin{aligned}
df_1 &= s(2m^* + s + 1) \\
df_2 &= 2(sn^* + 1) \\
F &= \frac{df_2}{df_1} \frac{T}{s}
\end{aligned}$$

- Roy's Largest Root

$$
\begin{aligned}
T &= \lambda_1 \\
df_1 &= \eta \\
df_2 &= (n_E - \eta + n_H) \\
\eta &= \max(n_h, d) \\
F &= \frac{df_2}{df_1} T
\end{aligned}
$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_c > 0$ are the c non-zero eigenvalues of the matrix $S_H S_E^{-1}$, where S_H is hypothesis sums of squares and cross product matrix, S_E is error sums of squares and cross product matrix, n_E is the degrees of freedom for S_E, n_H is the degrees of freedom for S_H, $s = \min(n_H, d)$, $2m^* = \text{abs}(d - n_H) - 1$ and $2n^* = n_E - d - 1$. For main effects, d is $l - 1$, where l is the number of levels for the variable being tested. For the two-way interactions, $d = (l_1 - 1)(l_2 - 1)$ where l_1 and l_2 are the respective number of levels for the two variables making up the interaction. The values of l, l_1 and l_2 can be those supplied in any of [**r**], [**nlx1**] or [**nlx2**], depending on the comparisons being made.

SPSS and SAS use different formula for the F test and degrees of freedom for Hotellings Trace when $n^* > 1$. SPSS uses the same formula irrespective of the value of n^*.

3 Multiple Treatment Comparisons

3.1 Tukey-Kramer, Bonnferonni, Dunn-Sidak and Fisher-LSD

As per the documentation for NAG library G04DBF, with the number of treatments (t) and the number of comparisons (k) given by:

One-way ANOVA

- Mean of repeated variable [**dm**]

$$
\begin{aligned}
t &= l_1 \\
k &= \frac{l_1(l_1 - 1)}{2}
\end{aligned}
$$

where l_1 is the number of repeated measures on each subject [**r**].

Two-way ANOVA, one repeated variable

- Mean of repeated variable [**dm1**]

$$\begin{aligned} t &= l_1 \\ k &= \frac{l_1(l_1-1)}{2} \end{aligned}$$

- Mean of between subject variable [**dm2**]

$$\begin{aligned} t &= l_2 \\ k &= \frac{l_2(l_2-1)}{2} \end{aligned}$$

- Interaction (comparing levels of repeated variable, within levels of between subject variable) [**di1**]

$$\begin{aligned} t &= l_1 \\ k &= \frac{l_1 l_2(l_1-1)}{2} \end{aligned}$$

- Interaction (comparing levels of repeated variable, within levels of between subject variable) [**di2**]

$$\begin{aligned} t &= l_2 \\ k &= \frac{l_1 l_2(l_2-1)}{2} \end{aligned}$$

where l_1 is the number of repeated measures on each subject [**r**] and l_2 is the number of levels for the between subject variable [**nlx**].

Two-way ANOVA, two repeated variables

- Mean of repeated variable [**dm1**]

$$\begin{aligned} t &= l_1 \\ k &= \frac{l_1(l_1-1)}{2} \end{aligned}$$

- Mean of between subject variable [**dm2**]

$$\begin{aligned} t &= l_2 \\ k &= \frac{l_2(l_2-1)}{2} \end{aligned}$$

- Interaction (comparing levels of first repeated variable, within levels of second repeated variable) [**di1**]

$$\begin{aligned} t &= l_1 \\ k &= \frac{l_1 l_2(l_1-1)}{2} \end{aligned}$$

- Interaction (comparing levels of second repeated variable, within levels of first repeated variable) [**di2**]

$$\begin{aligned} t &= l_2 \\ k &= \frac{l_1 l_2(l_2-1)}{2} \end{aligned}$$

where l_1 is the number of levels for the first repeated variable [**nlx1**] and l_2 is the number of levels for the second repeated variable [**nlx2**].

4 Multiple Treatment Comparisons

4.1 Scheffe, Holm-Bonferonni and Holm-Sidak

The Scheffe comparison is carried out as described in the NAG routine G04DBF.

Both the Holm-Bonferonni and Holm-Sidak comparisons are performed in the following manner:

1. Calculate unadjusted p-values for each of the k comparisons.
2. Sort the p-values in ascending order
3. Adjust the i'th smallest p-value, as if there had been $k-i+1$ comparisons made. If the adjusted p-value is no longer significant, flag all other comparisons as having not been made and end. Otherwise increment i and repeat.

The adjustments made in step (3) for the Holm-Bonferonni and Holm-Sidak adjustments are done in a similiar manner to that described for the Bonferonni and Dunn-Sidak adjustments in the documentation of NAG routine G04DBF.

The values of t (number of treatments) and k (number of comparisons) for the comparison of the means is the same as given above in section 3.1, the other values of t and k are given by:

Two-way ANOVA, one repeated variable

- Interaction (comparing levels of repeated variable, within levels of between subject variable) [**di1**]

$$t = l_1$$
$$k = \frac{l_1(l_1 - 1)}{2}$$

- Interaction (comparing levels of repeated variable, within levels of between subject variable) [**di2**]

$$t = l_2$$
$$k = \frac{l_2(l_2 - 1)}{2}$$

where l_1 is the number of repeated measures on each subject [**r**] and l_2 is the number of levels for the between subject variable [**nlx**].

Two-way ANOVA, two repeated variables

- Interaction (comparing levels of first repeated variable, within levels of second repeated variable) [**di1**]

$$t = l_1$$
$$k = \frac{l_1(l_1 - 1)}{2}$$

- Interaction (comparing levels of second repeated variable, within levels of first repeated variable) [**di2**]

$$t = l_2$$
$$k = \frac{l_2(l_2 - 1)}{2}$$

where l_1 is the number of levels for the first repeated variable [**nlx1**] and l_2 is the number of levels for the second repeated variable [**nlx2**].

5 Multiple Treatment Comparisons

5.1 Dunnett

The values of t (number of treatments) and k (number of comparisons) for the comparison of the means is the same as given above in section 3.1

Two-way ANOVA, one repeated variable

- Interaction (comparing levels of repeated variable, within levels of between subject variable) [**di1**]

$$\begin{aligned} t &= l_1 \\ k &= n \end{aligned}$$

- Interaction (comparing levels of repeated variable, within levels of between subject variable) [**di2**]

$$\begin{aligned} t &= l_2 \\ k &= n \end{aligned}$$

where l_1 is the number of repeated measures on each subject [**r**], l_2 is the number of levels for the between subject variable [**nlx**] and n is the number of subjects [**n**].

Two-way ANOVA, two repeated variables

- Interaction (comparing levels of first repeated variable, within levels of second repeated variable) [**di1**]

$$\begin{aligned} t &= l_1 \\ k &= \frac{n}{nlx1} \end{aligned}$$

- Interaction (comparing levels of second repeated variable, within levels of first repeated variable) [**di2**]

$$t = l_2$$
$$k = \frac{n}{nlx2}$$

where l_1 is the number of levels for the first repeated variable [**nlx1**], l_2 is the number of levels for the second repeated variable [**nlx2**] and n is the number of subjects [**n**].

6 Degrees of Freedom Used in Comparisons and Confidence Intervals

6.1 One-way ANOVA

Parameter estimates [**bcil** etc]	between subject error DF
Differences in Means of:	
Repeated variable [**dmdf**]	within subject error DF

6.2 Two-way ANOVA, one repeated variables

Parameter estimates [**b1cil,b2cil,b3cil** etc]	between subject error DF
Differences in Means of:	
Repeated variable [**dm1df**]	within subject error DF
Between subject variable [**dm2df**]	between subject error DF
Interactions, comparing levels of:	
Repeated variable, within levels of between subect variable [**di1df**]	within subject error DF
Between subect variable, within levels of repeated variable [**di2df**]	within subject error DF*

* by default, SPSS uses Sattherwaites degrees of freedom in these cases

6.3 Two-way ANOVA, two repeated variables

Parameter estimates [**b1cil,b2cil,b3cil** etc]	between subject error DF
Differences in Means of:	
Repeated variable [**dm1df**]	within subject variable 1 error DF
Between subject variable [**dm2df**]	within subject variable 2 error DF
Interactions, comparing levels of:	
First repeated variable, within levels of second repeated variable [**di1df**]	overall within subject error DF*
Second variable, within levels of first repeated variable [**di2df**]	overall within subject error DF*

* by default, SPSS uses Sattherwaites degrees of freedom in these cases