**R-functions for the analysis of variance**

The following R functions may be downloaded from the directory http://www.uni-koeln.de/~luepsen/R/

All file names of the functions have the extension ...R.

A complete set of all functions contains the file anova.lib, which can be made available after a download by means of `attach(path/anova.lib)`.

Usage advices:

- Variables used as factors have to declared as „factor“ by the user.
- The following command should be entered before usage:
  ```r
  options(contrasts=c("contr.sum","contr.poly"))
  ```
- In contrary to most other R anova functions all variables must be part of the dataframe.
- Data frames with missing values are not allowed, neither for repeated measures nor for between subject analyses.
  In such cases the function `na.omit` should be applied before performing the analysis.

There is warranty for correct functions.

*Haiko Lüpsen, (last revision 12.5.2021)*

---

**Index**

<table>
<thead>
<tr>
<th>Function</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. box.f:</td>
<td>3</td>
</tr>
<tr>
<td>Box-F-Test for heterogeneous variances</td>
<td></td>
</tr>
<tr>
<td>2. bf.f:</td>
<td>4</td>
</tr>
<tr>
<td>Brown &amp; Forsythe-F-test for heterogeneous variances</td>
<td></td>
</tr>
<tr>
<td>3. mbf.f:</td>
<td>5</td>
</tr>
<tr>
<td>modified Brown &amp; Forsythe-F-test for heterogeneous variances in mixed (split-plot) designs</td>
<td></td>
</tr>
<tr>
<td>4. wj.anova:</td>
<td>6</td>
</tr>
<tr>
<td>Welch-James-anova for heterogeneous variances in between subject designs</td>
<td></td>
</tr>
<tr>
<td>5. wj.spanova:</td>
<td>7</td>
</tr>
<tr>
<td>Welch-James-anova for heterogeneous covariance matrices in mixed (split plot) designs</td>
<td></td>
</tr>
<tr>
<td>6. box.andersen.f:</td>
<td>8</td>
</tr>
<tr>
<td>F-test for nonnormal distributed variables</td>
<td></td>
</tr>
<tr>
<td>7. check.covar:</td>
<td>9</td>
</tr>
<tr>
<td>several tests for homogeneity of covariance matrices</td>
<td></td>
</tr>
<tr>
<td>8. check.corr:</td>
<td>11</td>
</tr>
<tr>
<td>several tests for homogeneity of covariance matrices</td>
<td></td>
</tr>
<tr>
<td>9. check.sphere:</td>
<td>12</td>
</tr>
<tr>
<td>several tests for sphericity of a covariance matrix</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Method Name</td>
</tr>
<tr>
<td>---</td>
<td>----------------</td>
</tr>
<tr>
<td>10</td>
<td>ats.2</td>
</tr>
<tr>
<td></td>
<td>ats.3</td>
</tr>
<tr>
<td>11</td>
<td>np.anova</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>art1.anova</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>art2.anova</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>art3.anova</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>koch.anova</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>iga and iga.anova</td>
</tr>
<tr>
<td>17</td>
<td>ap.anova</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>simple.effects</td>
</tr>
<tr>
<td>19</td>
<td>gee.anova</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1. **box.f: Box-F-Test for heterogeneous variances**

1- or 2-factorial analysis of variance using the robust Box F-tests for heterogeneous variances.

Call: `box.f (model, dataframe)`

**Parameter:**

- **model**
  - anova model (as in function `aov`)
  - example: `x ~ A*B`

- **dataframe**
  - Data, object of type data.frame

**Result:**

- anova table
  - object of type data.frame and anova

**Method:**

Let there be \( k \) groups (main or interaction effect) with variances \( s_i^2 \). The F-test

\[
F = \frac{MS_{Effekt}}{MS_{Fehler}}
\]

can be corrected according to the heterogeneity of the variances by adjusting (exactly reducing) the numerator and denominator degrees of freedom. The numerator \( df \) \( df1 \) are multiplied with \( \varepsilon_1 \), the denominator \( df \) \( df2 \) with \( \varepsilon_2 \). The corrections are computed as followst:

\[
s^2 = \frac{\sum s_i^2}{k}
\]

\[
c^2 = \left( \frac{\sum (s_i^2 - \bar{s}^2)^2}{(k \cdot \bar{s}^4)} \right)
\]

\[
\varepsilon_1 = \left( 1 + \frac{k - 2}{k - 1} c^2 \right)^{-1} \quad \varepsilon_2 = (1 + c^2)^{-1}
\]

Here \( s^2 \) may be interpreted as an average variance and \( c^2 \) as a dispersion of the variances. It is easy to recognize that in the case of equal variances \( c^2=0 \) results and therefore \( \varepsilon_1 \) and \( \varepsilon_2 \) become 1.

**References:**

2. **bf.f: Brown & Forsythe-F-test for heterogeneous variances**

1- or 2-factorial analysis of variance using the robust method of Brown & Forsythe for heterogeneous variances.

The function consists of bf.f, bf2.f, bf3.f, bf.main, bf.orthog, bf.fratio.

**Call:** `bf.f (model, dataframe, mod=T)`

**Parameter:**
- `model` anova model (as in function `aov`)
  - example: `x ~ A*B`
- `dataframe` Data, object of type data.frame
- `mod` T: modification of error degrees of freedom by Mehrotra is used
  - F: original version of error degrees of freedom by Brown & Forsythe

**Result:**
- anova table object of type data.frame and anova

**Method:**

**1-factorial analysis:**

Let there be \( k \) groups with variances \( s_i^2 \), cell frequencies \( n_i \). Brown & Forsythe use the following ratio which follows approximately an F distribution:

\[
F = \frac{SS_{Effekt}}{SS_{Fehler}}
\]

Here \( SS_{Error} \) is computed as (with \( n = \sum n_i \))

\[
SS_{Error} = \sum \left( 1 - \frac{n_i}{n} \right) n_i^2
\]

The denominator degrees of freedom of the F-test are computed as follows

\[
df = \left( \sum \frac{m_i^2}{n_i - 1} \right)^{-1}
\]

\[
m_i = \left( 1 - \frac{n_i}{n} \right) \frac{s_i^2}{SS_{Error}}
\]

The 2-factorial analysis requires much more computation. For details the reader is referred to the article cited below.

**References:**


3. **mbf.f: modified Brown & Forsythe-F-test for heterogeneous variances**

2-factorial split-plot analysis of variance using the robust method of Brown & Forsythe for heterogeneous variances in the version of Vallejo et al., which allows nonsphericity and heterogeneous covariance matrices. `mbf.f` allows data in wide as well as in long format.

The function requires the function `bf.f` (from this package).

Call (wide format): `mbf.f(df, group)`

Call (long format): `with(df, mbf.f(y, groups=g, trial=w, id=id))`

Parameter:
- `df`: Dataframe, object of type data.frame, in case of wide format: only the repeated measurement variables
- `group`: grouping factor
- `trial`: repeated measurement factor
- `id`: case id

Result:
- anova table: object of type data.frame and anova

Examples:

Call (wide format): `mbf.f(winer518[,3:5],winer518,2)`

Call (long format): `with(winer518t, mbf.f(y, groups=Geschlecht, trial=Zeit, id=Vpn))`

References:


4. **wj.anova: Welch-James-anova for heterogeneous variances in between subject designs**

1- or 2-factorial analysis of variance for between subject designs using the robust method of Welch & James for heterogeneous variances

Call: `wj.anova (dataframe, dependent variable, grouping factors, Ftest)`

Parameter:

- **dataframe**: Data, object of type data.frame
- **dependent variable**: variable name (in “...“)
- **grouping factor 1**: variable name (in “...“)
- **grouping factor 2**: variable name (in “...“), optional
- **Ftest**: F (default): chi²-test as proposed by Algina & Olejnik
  T: F-test as proposed by Lix, Algina & Keselman

Result:

- **anova table**: object of type data.frame

Example:

`wj.anova (mydata, "y", "g1", "g2")`

References:


5. **wj.spanova: Welch-James-anova for heterogeneous covariance matrices in split plot designs**

1- or 2-factorial analysis of variance for mixed designs (split plot designs) using the robust method of Welch & James for heterogeneous covariance matrices. The dataframe must have the same structure as it is requested by `aov` oder `ezANOVA`.

**Call:** `wj.spanova (dataframe, dep var, group factor, repmes factor id variable)`

**Parameter:**
- **dataframe** Data, object of type `data.frame`
- **dependent variable** variable name (in “...“)
- **grouping factor** variable name (in “...“)
- **rep.meas. factor** variable name (in “...“)
- **id variable** variable for identifying cases (in “...“)

**Result:**
- **anova table** object of type `data.frame`

**Example:**
```
wj.spanova (mydata,"y", "group", "time", "Vpn")
```

**References:**

6. **box.andersen.f: F-test for nonnormal distributed variables**

1- or 2-factorial analysis of variance using the robust method of Box & Andersen for nonnormal distributed dependent variables.

**Call:** `box.andersen.f (model, dataframe)`

**Aufrufparameter:**
- `model`: anova model (as in function aov)
  - example: `x ~ A*B`
- `dataframe`: Data, object of type data.frame

**Result:**
- `anova`: anova table: object of type data.frame and anova
- `eps`: factor for the correction of the degrees of freedom

**Method:**

Box and Andersen use a modified F-test where the numerator and denominator degrees of freedom are multiplied with a parameter $d$. This is computed by means of the variance and kurtosis of the variable $x$. The following computation of the correction $d$ is valid for approximately equal $n_i$. Therefore let $n$ be the number of observations in each group. (It should be remarked that there exists also a more extensive formula for strongly varying $n_i$.

$$S_2 = \sum_{i}^{k} \sum_{j}^{n} (x_{ij} - \bar{x})^2 \quad S_4 = \sum_{i}^{k} \sum_{j}^{n} (x_{ij} - \bar{x})^4$$

These sums are used to compute the following intermediate parameters:

$$k_2 = S_2/(n - 1)$$
$$k_4 = [n(n + 1)S_4 - 3(n - 1)S_2^2] / [(n - 1)(n - 2)(n - 3)]$$

And finally $d$ is computed as

$$d = 1 + \frac{1}{n} \frac{k_4}{k_2^2}$$

**References:**

check.covar: test of homogeneity of covariance matrices

Several tests of homogeneity of covariance matrices:

- Likelihood Ratio test,
- Box M-test, based on multivariate normal distribution,
- Schott’s T1, an improvement of Box M-test, based on multivariate normal distribution,
- Schott’s T2, for elliptic distributions, incorporating their kurtoses for measuring the deviation from the multivariate normal distribution, assuming all distributions as equal,
- Schott’s T3, for elliptic distributions, incorporating their kurtoses, allowing different distributions,
- multivariate Levene test, assuming at least an ordinal scale (details see below),
- a pure dispersion test.

The T2 and T3 allow a wider area of distributions and are generalisations of T1. In his paper Schott recommends T2 for general usage, but limited to balanced designs. But own simulations showed that only the Levene test controls the type error rate in nearly all situations (e.g. different distributions) and has on the other side a good power performance in case of variance heterogeneity. For details see Luepsen (2020).

The function accepts the data in wide format as well as in long format.

Call (wide format): check.covar (dependent variables, grouping variable)
Call (long format): check.covar (dep varianle, groups=.., trial=.., id=..)

Parameters (wide format):

  dependent variables
  repeated measurement variables as a dataframe or matrix

  grouping variable
  vector containing the values of the grouping factor

Parameters (long format):

  dependent variable
  vector containing the values of the repeated measurement variable

  grouping variable
  vector containing the values of the grouping factor

  trial factor
  vector containing the values of the trial factor

  case id
  vector containing the case ids

Examples: check.covar (winer[,c("V3","V4","V5")], winer$V2) with(winer,check.covar(v,sex,time,id))

Output:

dataframe with $\chi^2$-values, df and p value for the 6 tests, NAs if a test is not computable, e.g.
if there are too less cases in relation to the number of dependent variables.

References:


Levene test:

For $i=1,...,I$ groups, $k=1,...,n_i$, $N (N = \sum n_i)$ subjects and $j=1,...,J$ repeated measurements $y_{k(i)j}$ shall denote the $j$ th measurement of variable $y$ for subject $k(i)$ and $m_{ij}$ the median in group $i$. Then for each subject $k(i)$ the following covariances $s_{j_1,j_2}$ are computed:

$$s_{k(i)j_1,j_2} = (y_{k(i)j_1} - m_{ij_1})(y_{k(i)j_2} - m_{ij_2}) \quad (j_1, j_2 = 1,...,J)$$

which are transformed into

$$\hat{s}_{k(i)j_1,j_2} = \text{sgn}(s_{k(i)j_1,j_2}) \cdot \sqrt{|s_{k(i)j_1,j_2}|}$$

where $\text{sgn}$ denotes the sign function. In the next step for each subject $k(i)$ the lower triangular of $\hat{s}_{k(i)j_1,j_2}$ is transformed into a vector which results in a data matrix $Y$ with $N$ rows and $(J+1)J/2$ columns. Finally a multivariate analysis of variance, e.g. Wilks Lambda test, is applied on $Y$ which checks the homogeneity of the covariance matrices of groups $i=1,...,I$. 
8. **check.corr: test of homogeneity of correlation matrices**

Several tests of homogeneity of correlation matrices:
- Jennrich Test,
- Larntz & Perlman Test,
- a modified, restricted to correlations, Levene like test by O’Brien (see above),
- a modified, restricted to correlations, Box-M-Test

Details will be found in Luepsen (2020), see check.covar.

The function accepts the data in wide format as well as in long format.

**Call (wide format):** `check.corr (dependent variables, grouping variable)`

**Call (long format):** `check.corr (dep varianle, groups=.., trial=.., id=..)`

**Parameters (wide format):**
- `dependent variables`: repeated measurement variables as a dataframe or matrix
- `grouping variable`: vector containing the values of the grouping factor

**Parameters (long format):**
- `dependent variable`: vector containing the values of the repeated measurement variable
- `grouping variable`: vector containing the values of the grouping factor
- `trial factor`: vector containing the values of the trial factor
- `case id`: vector containing the case ids

**Examples:**
- `check.corr (winer[,c("V3","V4","V5")], winer$V2)`
- `with(winer,check.corr(v,sex,time,id))`

**References:**
9. **check.sphere: test for sphericity of a covariance matrix**

Several tests for sphericity of a covariance matrix:

- Likelihood ratio test,
- a test by Muirhead & Waternaud (1980) for elliptic distributions, incorporating their kurtoses for measuring the deviation from the multivariate normal distribution, allowing a wider area of distributions (a generalization of the LR test),
- John’s V test (see John, 1972), approximated by a beta distribution,
- John’s V test, considering the deviation of the kurtosis from the value for the normal distribution (Hallin & Paindaveine, 2006), approximated by a $\chi^2$-distribution
- John’s V test, using a more accurate computation of the p value (see Nagao, 1973),
- Mauchly’s test (e.g. Winer, 1991, p. 255),
- multisample Mauchly test (Mendoza, 1980),
- multisample Mauchly test (Harris, 1984),
- a test for compound symmetry, as described by Winer (1991, p. 517).

The most favorable tests are said to be John’s and Nagao’s tests, though the two for elliptic distributions offer a wider range of applications. One remark to John’s test: many different versions of this test can be found in the publications. Some of them differ only in the presentation of the formulae, but some also lead to different results. Therefore the formula used is presented below. For more information see Luepsen (2020).

The function accepts the data in wide format as well as in long format.

Call (wide format): check.sphere (dependent variables [, grouping variable])
Call (long format): check.sphere (dep variable, trial=.., Id=.. [, groups=..])

Parameters (wide format):

- **dependent variables**
  repeated measurement variables as a dataframe or matrix
- **grouping variable**
  vector containing the values of the grouping factor (optional)

Parameters (long format):

- **dependent variables**
  vector containing the values of the repeated measurement variable
- **groups**
  vector containing the values of the grouping factor (optional)
- **trial**
  vector containing the values of the trial factor
- **Id**
  vector containing the case ids

Example: check.sphere (winer[,c("V3","V4","V5")])
        with(winer,check.sphere(v,time,id))
Output:

$results: dataframe with $\chi^2$-values, df and p value for the 8 tests.
For John’s test with the kurtosis correction the column chisquare contains the kurtosis.

$Box\text{. }\epsilon: the value of Box \epsilon$

$k\text{. deviation}: deviation of the kurtosis from the value of the normal distribution

$error: error code:
2: data matrix and vector of grouping variable have different lengths.

References:


The Annals of Statistics, Vol. 34, No. 6, pp 2707–2756


Luepsen, Haiko (2020): *Anmerkungen zum Testen der Spherizität*
URL: http://www.uni-koeln.de/~luepsen/statistik/texte/sphericity.pdf


John’s test:

Let $S(s_{ij})$ be the covariance matrix of the K repeated measurements. The statistic is computed in several steps:

$$U = \left( \sum_{i}^{K} \lambda_i^2 \right) / \left( \sum_{i}^{K} \lambda_i \right)^2$$

where $\lambda_i$ are the eigenvalues of the matrix $CSC^\prime$ with a normalized orthogonal transformation matrix $C$. Next the value

$$T = (K \cdot U - 1) / (K - 1)$$
is computed, for which several approximate distributions exist. Normally the $\chi^2$-distribution is used, noting that this is only satisfying for larger samples $n$:

$$X = \frac{n}{2} \cdot K \cdot (K - 1) \cdot T$$

$X$ is $\chi^2$-distributed with $(K-1)/K - 1$ degrees of freedom. For smaller samples also the statistic $T$ can be used which is approximately beta-distributed or which can be transformed in an approximately $F$-distributed statistic. For details see John (1972).

As $U = 1/(K \cdot \varepsilon)$, where $\varepsilon$ is the Box correction factor measuring the deviation from sphericity, above $X$ can be written as

$$X = \frac{n}{2} \cdot K \cdot \left(\frac{1}{\varepsilon} - 1\right)$$

which directly shows that $X$ is a statistic for the deviation from sphericity and e.g. $X=0$ for $\varepsilon=1$. 


10. **ats.2 : 2-factorial analysis of variance**  
**ats.3 : 3-factorial analysis of variance**

2-factorial or 3-factorial analysis of variance using the ATS (anova type statistic) method of Akritas, Arnold and Brunner. Empty cells are not allowed.

Call: `ats.2 (model, dataframe)`  
`ats.3 (model, dataframe)`

Parameter:
- **model**: anova model (as in function `aov`)  
  example: `x ~ A*B`
- **dataframe**: Data, object of type `data.frame`

Result:
- **anova**: anova table: object of type `data.frame`

References:


Note:

For repeated measures designs there exists the package `nparLD` on Cran.
11. **np.anova: nonparametric analysis of variance using the KWF-, Puri & Sen- and van der Waerden-method**

factorial analysis of variance, with and without repeated measurements, either using the KWF-method, which is a generalisation of the well known Kruskal-Wallis- and Friedman-analyses, the Puri & Sen (L-statistic) method, or the van der Waerden method based on the inverse normal transformation. In the case of repeated measurements the dataframe must have the same structure as it is requested by `aov` or `ezANOVA`.

**Call:**
- `np.anova (model, dataframe)` generalized Kruskal-Wallis-Friedman
- `np.anova (model, dataframe, method=1)` van der Waerden method

**Parameter:**
- **model** anova model (as in function `aov`)
  examples: `x ~ A*B` or `score ~ group*time+Error(Vpn/time)`
- **dataframe** Data, object of type `data.frame` (long format)
- **method**
  - 0: KWF generalized Kruskal-Wallis-Friedman tests
  - 1: generalized van der Waerden method
  - 2: Puri & Sen method
  - 3: Puri & Sen method based on normal scores
- **compact** for repeated measurements only:
  - T: all tests in one table (dataframe) (default)
  - F: one table (dataframe) for each error term (similar to `summary(aov)`)
- **pseudo** ranking method:
  - F: standard ranking (default)
  - T: pseudo ranks

**Result:**
- anova table object of type `data.frame` and `anova`

**References:**


12. **art1.anova: nonparametric analysis of variance using the ART (Aligned Rank Transform) for between subjects designs**

factorial nonparametric analysis of variance.

**Call:** `art1.anova (model, dataframe, method=.., main=.., adjust=.., INT=..)`

**Parameter:**
- `model`: anova model (as in function `aov`)
  example: `x ~ A*B`
- `dataframe`: Data, object of type `data.frame`
- `method`: 0: Alignment using a regression (default)
  1: Alignment by computing the deviations from the cell means
- `main`: F: for tests of main effects use the simple RT-technique (default)
  T: for tests of main effects use also the ART-technique
- `INT`: F: no transformation into normal scores after ranking (default)
  T: transformation into normal scores after ranking

**Result:**
- `anova table`: object of type `data.frame` and `anova`

**References:**


**Note:**

Meanwhile there exists also the package `ARTool` on Cran for between subjects designs.
13. **art2.anova: nonparametric analysis of variance using the ART (Aligned Rank Transform) for pure within subjects designs**

1- and 2-factorial nonparametric analysis of variance for pure within subjects designs. For the tests of main effects the simple RT-technique is applied.

Call: `art2.anova(model, dataframe, main=..., INT=...)`

Parameter:

- **model**: anova model (as in function `aov`)  
  example: `x ~ Medi*Aufgabe+Error(Vpn/( Medi*Aufgabe))`
- **dataframe**: Data, object of type `data.frame` (long format)
- **main**: F: for tests of main effects use the simple RT-technique (default)  
  T: for tests of main effects use also the ART-technique
- **INT**: F: no transformation into normal scores after ranking (default)  
  T: transformation into normal scores after ranking

Result:

- anova table: object of type `data.frame` and `anova`

References:

14. **art3.anova: nonparametric analysis of variance using the ART (Aligned Rank Transform) for mixed designs (split plot designs)**

Factorial nonparametric analysis of variance for mixed designs with at least one grouping factor and one or two repeated measurement factors. For the tests of main effects the simple RT-technique is applied.

**Call:** art3.anova (model, dataframe, method=..., main=..., INT=...)

**Parameter:**
- `model`: anova model (as in function aov)
  
  example: `score ~ gruppe*Zeit+Error(Vpn/Zeit)`

- `dataframe`: Data, object of type data.frame (long format)

- `method`: 0: Alignment using a regression (default)
  1: Alignment by computing the deviations from the cell means

- `main`: F: for tests of main effects use the simple RT-technique (default)
  T: for tests of main effects use also the ART-technique

- `INT`: F: no transformation into normal scores after ranking (default)
  T: transformation into normal scores after ranking

**Result:**
- `anova table`: object of type data.frame and anova

**References:**
15. koch.anova: nonparametric anova for split plot designs using the procedure by G. Koch

2-factorial nonparametric analysis of variance based on ranks for mixed designs (split plot designs) using the method of Gary Koch. This method does not assume sphericity of the covariance matrix of the dependent variables. But there are several alternatives for this method:

The grouping main effect can be tested with a multivariate Kruskal-Wallis test. It has to be pointed out to the fact that this test is not independent of the interaction effect, with the consequence that any interaction will lead to larger type I error rates for the test of the grouping effect. But in practice that is not severe because in the case of a significant interaction effect the tests of the main effects are no longer of interest. Alternatively the grouping main effect can be tested by means of a univariate Kruskal-Wallis test on the subject means (default).

For the test of the repeated measures main effect there exists a test $W_{ni}$* assuming distributions which are equally shaped for all groups. Additionally, the sample sizes $n_i$ should be larger than the number of repeated measurements. Here also the test is not independent of the interaction effect, with the consequence that any interaction will lead to larger type I error rates for the test of the repeated measurements effect. Alternatively there is a version of $W_{ni}$* which is independent of the interaction at the cost of the power. On the other hand there exists a test $W$ assuming arbitrary distribution shapes which is also unaffected by the interaction (default).

The dataframe must have the „wide“ format, i.e. all repeated measures variables must be in one row. The structure as it is requested by aov oder ezANOVA is not allowed. Missing values (NAs) have to be eliminated befor usage.

Call: koch.anova (dataframe, grouping factor, A=.., B=..)

Parameter:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>dataframe</td>
<td>data, object of type data.frame</td>
</tr>
<tr>
<td>grouping factor</td>
<td>variable within the data.frame</td>
</tr>
<tr>
<td>A=</td>
<td>0: univariate Kruskal-Wallis test on the subject means 1: multivariate Kruskal-Wallis test</td>
</tr>
<tr>
<td>B=</td>
<td>0: $W$ test for the entire sample, assuming arbitrary distribution shapes 1: $W_{ni}$* test for the entire sample, assuming equally shaped distributions 2: $W_{ni}$* test pooled over all groups, assuming equally shaped distributions</td>
</tr>
</tbody>
</table>

Result:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>anova table</td>
<td>object of type data.frame</td>
</tr>
</tbody>
</table>

Example:

```r
koch.anova (mydata[ ,c("y1", "y2", "y3", "y4")], mydata$age)
```

References:

16. iga and iga.anova: the general approximation test (GA) and the improved general approximation test (IGA) by H.Huynh

Adjustments for the parametric F test in 2-factorial split-plot designs as proposed by H. Huynh, to allow for nonspherical and heterogeneous covariance matrices. For the repeated measures main effect as well as for the interaction effect a correction factor for the F value and adjusted degrees of freedom are computed. The GA can be considered as an improvement of the well-known degrees of freedom adjustment by Huynh & Feldt, whereas the IGA as an improvement of the GA for the case of heterogeneous covariance matrices.

The function iga computes the correction factor and the adjusted degrees of freedom for the two effects, whereas the function iga.anova performs a complete split-plot anova combining the results with those from iga resulting in an adjusted anova table. iga.anova allows data either in long or in wide format.

Call: iga (dataframe, grouping factor)

Parameter:

- dataframe: data of the dependent variables in wide format, object of class matrix or data.frame
- grouping factor: vector of class factor

Result: list of 4 vectors, each having 3 elements

- GA.B: GA-adjustments for the repeated measures main effect
- GA.AB: GA-adjustments for the interaction effect
- IGA.B: IGA-adjustments for the repeated measures main effect
- IGA.AB: IGA-adjustments for the interaction effect

Each of the above vectors has 3 elements:

1. correction factor c for the F value
2. adjusted degrees of freedom for the numerator df1
3. adjusted degrees of freedom for the denominator df2

If F is the unadjusted F value from the parametric anova, then cF is to be tested with (df1, df2) degrees of freedom

Examples:

iga(winer518[,3:5],winer518,2)

Call (wide format): iga.anova (df, group, ga=F)

Call (long format): with(df,iga.anova (y, groups=g, trial=w, id=id, ga=F))
iga and iga.anova: the general approximation test (GA) and the improved general approximation test (IGA) by

Parameters:

\( df \)  
- data frame either in wide format or in long format  
- which contains in case of wide format only the dependent variables

\( y \)  
- dependent variable

\( g \)  
- grouping factor of class factor

\( w \)  
- repeated measures factor of class factor

\( id \)  
- case id variable of class factor

\( iga \)  
- F: IGA adjustment, T: GA adjustment

Examples:

\[
\text{iga.anova(winer518[,3:5],winer518[,2])} \quad \# \text{ wide format}
\]

\[
\text{with(winer518t, iga.anova(score,Geschlecht,Zeit,Vpn))} \quad \# \text{ long format}
\]

References:

17. **ap.anova: nonparametric anova for repeated measures designs based on a multivariate test by Agresti & Pendergast**

1- or 2-factorial nonparametric analysis of variance based on ranks for repeated measures designs (e.g. split plot designs) using a nonparametric multivariate test by A. Agresti & J. Pendergast. This method does not assume sphericity of the covariance matrix of the dependent variables. Only the tests of the repeated measures main effect and of the interaction are performed.

The dataframe must have the „long“ format, the standard for repeated measures designs. Missing values (NAs) have to be eliminated before usage.

**Call:** ap.anova (dataframe, dependent var, case id, trial factor [,grouping factor])

**Parameter:**
- dataframe: data, object of class data.frame
- dependent variable: dependent variable within the data.frame
- case id: case identifier within the data.frame of class factor
- trial factor: repeated measures factor within the data.frame of class factor
- grouping factor: grouping factor within the data.frame of class factor (optional)

Variable names must be stated in „...“.

**Result:**
- anova table: object of class data.frame

**Example:**

```
  ap.anova (winer518t,"score","Vpn","Geschlecht","Zeit")
```

**References:**


18. **simple.effects: parametric analysis of simple effects for between subject and mixed designs**

Parametric analysis of simple effects for designs with at least one grouping factor and at most one repeated measurement factors.

**Call:** `simple.effects (anova, interaction, dataframe, adjust=...)`

**Parameter:**

- **anova**
  anova result (as from function `aov`)

- **interaction**
  one or more interaction terms enclosed in “...“, e.g. `c("A*time", "A*time")`

- **dataframe**
  dataframe which was used for the analysis by `aov`

- **adjust**
  optional: $\alpha$-adjustment method (see R function `p.adjust`), default “none”, no adjustment

**References:**

B.J.Winer et al.: *Statistical Principles in Experiments*al Design*,
19. **gee.anova: Anova-like tests for GEE and GLMM models**

There are 2 Anova-like Wald tests for 2-factorial designs: `gee.anova` for the classical Wald-test and `gee.robanova` for a robust Wald-test according to Fan & Zhang. The classical Wald test is rather liberal, especially in the case of GEE and GLMM models where the covariance matrix of the parameter estimates is generally underestimated, resulting in too large $\chi^2$-values.

Call: `gee.anova(coefficients, covariance matrix, degrees of freedom, n)`

`gee.robanova(coefficients, covariance matrix, degrees of freedom)`

Parameter:

- `coefficients`: regression coefficients (details see below)
- `covariance matrix`
- `degrees of freedom`: Array with 3 df for 2 factors and the interaction
- `n`: sample size (required for the F test)

Result:

The result is a dataframe with 3 rows, one for each of the 3 effects with columns:

- `gee.anova`: degrees of freedom, $\chi^2$-value, p value
- `gee.robanova`: degrees of freedom, $\chi^2$-value, corresponding p value, F-value, corresponding p value

References:


Specification of the coefficients and covariance matrix:

In every model estimation the coefficients and covariance matrix are part of the resulting object. For the most popular of the functions to be used for the analyses the following table shows where these are to be found:

<table>
<thead>
<tr>
<th>function</th>
<th>package</th>
<th>coefficients</th>
<th>covariance matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>gls</td>
<td>nlme</td>
<td>...$coefficients</td>
<td>vcov(...)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$varBeta</td>
</tr>
<tr>
<td>glm</td>
<td>stats</td>
<td>...$coefficients</td>
<td>vcov(...)</td>
</tr>
<tr>
<td>lmer</td>
<td>lme4</td>
<td>...@beta summary(...)coeficients[,1]</td>
<td>vcov(...)</td>
</tr>
<tr>
<td>glmer</td>
<td>lme4</td>
<td>...@beta summary(...)coeficients[,1]</td>
<td>vcov(...)</td>
</tr>
<tr>
<td>glmmML</td>
<td>glmmML</td>
<td>...$coefficients</td>
<td>...$variance</td>
</tr>
<tr>
<td>glmmPQL</td>
<td>MASS</td>
<td>&quot;$coefficients$fixed</td>
<td>...$varFix</td>
</tr>
<tr>
<td>geeglm</td>
<td>geepack</td>
<td>...$coefficients</td>
<td>...$geese$vbetta</td>
</tr>
<tr>
<td>gee</td>
<td>drgee</td>
<td>...$coefficients</td>
<td>...$vcov</td>
</tr>
<tr>
<td>gee</td>
<td>gee</td>
<td>...$coefficients</td>
<td>...&quot;robust.variance&quot;</td>
</tr>
<tr>
<td>wgee</td>
<td>wgeesel</td>
<td>...$beta</td>
<td>...$var</td>
</tr>
<tr>
<td>JGee1</td>
<td>JGEE</td>
<td>...$coefficients</td>
<td>...&quot;robust.variance&quot;</td>
</tr>
<tr>
<td>MGEE</td>
<td>PGEE</td>
<td>...$coefficients</td>
<td>...&quot;robust.variance&quot;</td>
</tr>
</tbody>
</table>