Factorial design

		Factor B		
		B_1	B_2	B_3
Factor A	A_1	y_{11k}	y_{12k}	y_{13k}
	A_2	y_{21k}	y_{22k}	Y_{23k}
	A_3	y_{31k}	y_{32k}	y_{33k}

Example of a factorial design with two factors (A and B). Each factor has three levels. y_{ijk} represents the kth observation in the condition defined by the ith level of factor A and jth level of factor B.

- The most common design for a *n*-way ANOVA is the factorial design.

- In a factorial design, there are two or more experimental factors, each with a given number of levels.

- Observations are made for each combination of the levels of each factor (see example)

- In a completely randomized factorial design, each experimentally unit is randomly assigned to one of the possible combination of the existing level of the experimental factors.

Advantages of the factorial design

- A two-way design enables us to examine the *joint* (or interaction) effect of the independent variables on the dependent variable. An *interaction* means that the effect of one independent variable has on a dependent variable is not the same for all levels of the other independent variable. We cannot get this information by running separate one-way analyses.

- Factorial design can lead to more powerful test by reducing the error (within cell) variance. This point will appear clearly when will compare the result of one-way analyses with the results of a two-way analyses or t-tests.

Interaction plot

	B_1	B_2	B_3
A_1	$\mu_{11}=6$	$\mu_{12}=10$	$\mu_{13}=20$
A_2	$\mu_{21}=10$	$\mu_{22}=14$	$\mu_{23}=24$
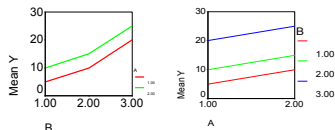
	B_1	B_2	B_3
A_1	$\mu_{11}=6$	$\mu_{12}=10$	$\mu_{13}=20$
A_2	$\mu_{21}=10$	$\mu_{22}=10$	$\mu_{23}=4$

Exercise. Make the interaction plots for the second table. Describe the interaction (if any).

- An interaction plot represents the mean value m_{ij} observed in each one of the condition of a factorial design.
- The Y axis corresponds to the dependent (or criterion) variable. The various level of one of the two experimental factor are aligned on the X axis. The lines relate the mean values that corresponds to the same level of the second experimental factor.
- There is an *interaction* between the factors if the lines are not parallel because the effect of one factor depends on the value of the other factor.
- If the lines are a parallel, the effect of the second factor is independent from the value of the first factor. In other words, there is no interaction.

Structural model (factorial ANOVA)

	B_1	B_2	B_3	Mean
A_1	μ_{11}	μ_{12}	μ_{13}	$\mu_{1\bullet}$
A_2	μ_{21}	μ_{22}	μ_{23}	$\mu_{2\bullet}$
Mean	$\mu_{\bullet 1}$	$\mu_{\bullet 2}$	$\mu_{\bullet 3}$	μ

- Let y_{ijk} be the kth observation of the ith level of factor A and jth level of factor B.
- Let μ_{ij} be the population mean for the ith level of factor A and jth level of factor B (condition A_iB_j), let $\mu_{i\bullet}$ be the population mean in condition A_i, let $\mu_{\bullet j}$ be the population mean in condition B_j and let μ be the grand mean.
- By definition $\alpha_i = \mu_{i\bullet} - \mu$ is the effect of factor A and $\beta_j = \mu_{\bullet j} - \mu$ is the effect of factor B.

- The structural model of a two-way factorial ANOVA *without interaction* is

$$y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$$

- In absence of interaction, the mean value μ_{ij} in condition (A_iB_j) depends in a *additive* manner on the effect of each condition

$$\mu_{ij} = \mu + \alpha_i + \beta_j$$

- The complete model of the two-way factorial ANOVA is

$$y_{ijk} = \mu + \alpha_i + \beta_i + \alpha\beta_{ij} + \varepsilon_{ijk}$$

where $\alpha\beta_{ij} = \mu_{ij} - (\alpha_i + \beta_j + \mu) = \mu_{ij} - \mu_{i\bullet} - \mu_{\bullet j} + \mu$ is the interaction effect. The interaction effect represents the fact that the contribution of one factors depends on the value of the other factor in a non-additive way.

Exercise.

Mean values

m_{ij}	B_1	B_2	B_3	$m_{i\bullet}$
A_1	6	10	20	12
A_2	10	14	24	16
$m_{\bullet j}$	8	12	22	14

Table of effects

$\alpha\beta_{ij}$	B_1	B_2	B_3	α_i
A_1	0	0	0	-2
A_2	0	0	0	+2
β_j	-6	-2	+8	

m_{ij}	B_1	B_2	B_3	$m_{i\bullet}$
A_1	6	10	20	12
A_2	10	10	4	8
$m_{\bullet j}$	8	10	12	10

$\alpha\beta_{ij}$	B_1	B_2	B_3	α_i
A_1	-4	-2	+6	+2
A_2	+4	+2	-6	-2
β_j	-2	0	+2	

- Compute the main and interaction effects from the mean values (see tables in the left column). Answer: see tables in the right column.

Data Analysis (draft) - Gabriel Baud-Bovy

Sum of squares

$$SST = \sum_{i,j,k} (y_{ijk} - m)^2$$

$$SStr = \sum_{i,j,k} (m_{ij} - m)^2$$

$$SSE = \sum_{i,j,k} (y_{ijk} - m_{ij})^2$$

$$SSA = \sum_{i,j,k} (m_{i\bullet} - m)^2$$

$$SSB = \sum_{i,j,k} (m_{\bullet j} - m)^2$$

$$SSAB = \sum_{i,j,k} \left(m_{ij} - m_{i\bullet} - m_{\bullet j} + m\right)^2$$

- Like in the one-way ANOVA, the total sum of squares (SST) can be decomposed into a between-groups sum of square (the treatment effect, SStr) and a within-group sum of square (SSE) which corresponds to the residual variance:

$$SST = SStr + SSE$$

  Note that the group (or experimental condition) in a factorial designed is determined by the value of two or more experimental factors.

- The between-group variations (SStr) can themselves be decomposed further into a variations that are explained by factor A (SSA), variations that are explained by factor B (SSB) and variation that are explained by the interaction between both factors (SSAB)

$$SStr = SSA + SSB + SSAB$$

Data Analysis (draft) - Gabriel Baud-Bovy

F tests

- A factorial design aims at answering three different questions:
  1. Is there an effect of the first experimental factor?

  $H_0: \alpha_i = 0 \quad (y_{ijk} = \mu + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk})$

  2. Is there an effect of the second experimental factor?

  $H_0: \beta_j = 0 \quad (y_{ijk} = \mu + \alpha_i + \alpha\beta_{ij} + \varepsilon_{ijk})$

  3. Is there an interaction?

  $H_0: \alpha\beta_{ij} = 0 \quad (y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk})$

- In all cases, the alternative hypothesis is the complete model

  $H_1: y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}$

- The residual variance (within-group variance) for this model is:

  $$MSE = \frac{SSE}{n_A n_B (n-1)} = \frac{SSE}{N - n_A n_B} \quad (N = n_A n_B n)$$

- In all cases, the F test is constructed by computing the percentage of variance that is explained by the parameters of interest divided by the residual variance of the more complex model.
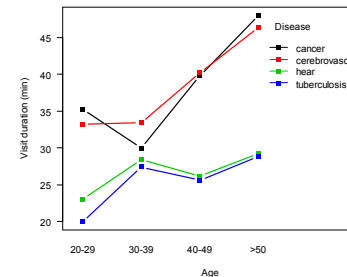
$$F_A = \frac{SSA/(n_A - 1)}{MSE}$$

$$F_B = \frac{SSB/(n_B - 1)}{MSE}$$

$$F_{AB} = \frac{SSAB/(n_A - 1)(n_B - 1)}{MSE}$$

If the null hypotheses are true, the F ratios follow a Fisher distribution with the corresponding degrees of freedom.

In that can be shown that the numerator is also an estimate the residual variance.

Data Analysis (draft) - Gabriel Baud-Bovy

[R] Interaction plot

- Lines are not parallel which is the tell-tale sign of an interaction. This plot suggests that the visit time increase with the older age groups for the cancer and cerebro-vascular diseases while it remained constant for the heart and tuberculosis diseases.

```
> visits<-read.table("visits.dat",header=TRUE)
> visits$age<-ordered(visits$age,c("20-29","30-39","40-49",">50"))
> interaction.plot(visits$age,visits$disease,visits$duration,
+       type="b",col=1:4,lty=1,lwd=2,pch=c(15,15,15,15),las=1,
+       xlab="Age",ylab="Visit duration (min)",trace.label="Disease")
```

Data Analysis (draft) - Gabriel Baud-Bovy

[R] 2-way ANOVA

- two-way ANOVA without interaction

```
> fit<-aov(duration~disease+age,visits)
> anova(fit)
Analysis of Variance Table
Response: duration
          Df  Sum Sq Mean Sq F value    Pr(>F)
disease    3 2992.45  997.48  47.037 < 2.2e-16
age        3 1201.05  400.35  18.879 3.649e-09
Residuals 73 1548.05   21.21
```

- two-way ANOVA with interaction

```
> fit<-aov(duration~disease+age+disease:age,visits)
> fit<-aov(duration~disease*age,visits)
> anova(fit)
Analysis of Variance Table
Response: duration
            Df Sum Sq Mean Sq F value    Pr(>F)
disease      3 2992.45  997.48 67.9427 < 2.2e-16
age          3 1201.05  400.35 27.2695 1.763e-11
disease:age  9  608.45   67.61  4.6049 0.0001047
Residuals   64  939.60   14.68
```

- The ANOVA table shows statistically significant main effects of the age and disease factors as well as a statistically significant interaction.

- Note the change significativity of the main effect when the interaction is included. This is due to the smaller denominator (residual error) in the F ratio.

- Manual compuation:

$$F_{AGE} = \frac{1201.050/3}{939.6/64} = \frac{400.350}{14.681} = 27.269$$

$$F_{DISEASE} = \frac{997.483}{14.681} = 67.943$$

$$F_{AGE:DISEASE} = \frac{67.606}{14.681} = 4.605$$

$$\sigma^2 = 14.681$$

Data Analysis (draft) - Gabriel Baud-Bovy

Exercise

```
> anova(aov(duration~disease,visits))
Analysis of Variance Table
Response: duration
          Df Sum Sq Mean Sq F value   Pr(>F)
disease    3 2992.45  997.48  27.576 3.600e-12
Residuals 76 2749.10   36.17

> anova(aov(duration~age,visits))
Analysis of Variance Table

Response: duration
          Df Sum Sq Mean Sq F value    Pr(>F)
age        3 1201.0  400.3   6.7012 0.0004502
Residuals 76 4540.5   59.7

> anova(aov(duration~disease*age,visits))
Analysis of Variance Table
Response: duration
            Df Sum Sq Mean Sq F value    Pr(>F)
disease      3 2992.45  997.48 67.9427 < 2.2e-16
age          3 1201.05  400.35 27.2695 1.763e-11
disease:age  9  608.45   67.61  4.6049 0.0001047
Residuals   64  939.60   14.68
```

- **Exercise.** Analyze the effect of the age and disease variable on the visit time by doing two separate one-way ANOVAs. Compare the results with the main effects of a two-way factorial ANOVA.

- As expected, the sum of square that corresponds to the tested hypothesis (red circles) are the same for the one-way ANOVA and the two-way ANOVA. However, the F statistics (green circles) for the one-way ANOVA are quite different from the main effect of the two-way ANOVA because the estimates of the residual variance are different (see blue circles).

Data Analysis (draft) - Gabriel Baud-Bovy

Exercise

- What is total sum of square?

```
> m<-mean(visits$duration)
> sum((visits$duration-m)^2)
[1] 5741.55
> fit<-aov(duration~disease*age,visits)
> sum(anova(fit)[,"Sum Sq"])
[1] 5741.55
```

Can you reconstruct the results of a one-way ANOVA from the results of a two-way factorial ANOVA? What would be the $SS_{Between}$, SS_{Within} and SS_{Total} of a one-way ANOVA performed with the disease experimental factor?

Answer: SS_{Total} = 5741.550 (no difference)

$SS_{Between}$ = $SS_{Disease}$ = 2992.45 (no difference)

SS_{Within} = SS_{Age} + $SS_{Age:Disease}$ + SSE = 1201.050 + 608.450 + 939.6 = 2749.1

(In the one-way ANOVA, the age and the interaction are ignored and considered as unexplained part of the variation of dependent variable).

- What is the value of the F statistics in the one-way ANOVA?

Answer: F = (2992.45/3)/(2747.1/76)=27.576

Data Analysis (draft) - Gabriel Baud-Bovy

Unbalanced data

```
> v0<-visits[3:nrow(visits),] # remove two first cases
> table(v0$age,v0$disease)
       cancer cerebrovascular hear tuberculosis
20-29    5                    5   3            5
30-39    5                    5   5            5
40-49    5                    5   5            5
>50      5                    5   5            5

> fit<-aov(duration~disease+age,v0)
> anova(fit)
Analysis of Variance Table
Response: duration
          Df Sum Sq Mean Sq F value    Pr(>F)
disease    3 2839.64  946.55 43.767 3.964e-16 ***
age        3 1161.89  387.30 17.908 9.339e-09 ***
Residuals 71 1535.51   21.63

> fit<-aov(duration~age+disease,v0)
> anova(fit)
Analysis of Variance Table

Response: duration
          Df Sum Sq Mean Sq F value    Pr(>F)
age        3 1072.64  357.55 16.532 3.017e-08 ***
disease    3 2928.88  976.29 45.142 < 2.2e-16 ***
Residuals 71 1535.51   21.63
```

- When the dataset is balanced (same number of observation per group), the order in which factor are specified is not important.

- When the data set is unbalanced, the percentage of variance by a factor explained depend on its position.

Data Analysis (draft) - Gabriel Baud-Bovy

Type I, Type II and Type III sum of squares

- **Type I (sequential):**
  - Terms are entered sequentially in the model.
  - Type I SS depend on the order in which terms are entered in the model
  - Type I SS can be added to yield to the total SS.
- **Type II (hierarchical)**:
  - see textbook
- **Type III (marginal)**
  - Type III SS correspond to the SS explained by a term *after* all other terms have already been included in the model.
  - Type III SS do not add.

- The analysis of unbalanced data sets (different number of observation in each group) present speficial difficulties because there are different ways of computing the sum of squares. These different ways corresponding to different hypotheses and, correspondly, the F tests are different.
- The R function anova yields Type I sum of square.
- Most textbooks suggest using the Type III sum of squares and many statistical sofftwares use Type III sum of square as a default but many stasticians think it does not make sense when there are statistically significant interactions.

Overall & Spiegel (1969) Psychol. Bull., 72:311- 322, for a detailed discussion of factorial designs.

[R] Type III sum of square

- The function Anova in the library car compute type II and type III sum of squares

```
> library(car)
> Anova(aov(duration~disease+age,v0),type="III")
Anova Table (Type III tests)
Response: duration
              Sum Sq Df  F value     Pr(>F)
(Intercept) 29261.2  1 1353.001 < 2.2e-16 ***
age          1161.9  3   17.908  9.34e-09 ***
disease      2928.9  3   45.142 < 2.2e-16 ***

Residuals    1535.5 71
```

- Compare with Type I sum of squares

```
> anova(aov(duration~disease+age,v0))        > anova(aov(duration~age+disease,v0))
Analysis of Variance Table                   Analysis of Variance Table
Response: duration                           Response: duration
          Df  Sum Sq Mean Sq F value   Pr(>F)           Df  Sum Sq Mean Sq F value   Pr(>F)
disease    3 2839.64  946.55  43.767 3.964e-16 age       3 1072.64  357.55  16.532 3.017e-08
age        3 1161.89  387.30  17.908 9.339e-09 disease   3 2928.88  976.29  45.142 < 2.2e-16
Residuals 71 1535.51   21.63                  Residuals 71 1535.51   21.63
```

Repeated-measure designs

	Treatment			
Subject	1	2	3	Mean
1	2	4	7	4.33
2	10	12	13	11.67
3	22	29	30	27.00
4	30	31	34	31.67
Mean	16	19	21	18.67

This example of one-way repeated measure ANOVA shows only small differences between treatments and large difference between subjects. In the repeated-measure ANOVA, we neglect the variations between subjects and consider only the variation for each treatment within each subject.

- In repeated-measure designs, several observations are made on the same experimental units. For a example, one of the most common research paradigm is that where subjects are observed at several different point in time (e.g., before and after treatment, longitudinal studies).

- In repeated measure design, it is important to distinguish between-subject and within-subject factors.
  - **Within-subject factors** are variables (like time or treatment or repetition) that identify the differences between conditions or treaments that have been assigned to each subject.
  - **Between-subject factors** are varables (like age or sex or group) that identify differences between the subjects.

Statistical approaches

- There are three approaches to repeated-measure designs:

1. **The univariate approach:** This approach uses the classic univariate F test of the ANOVA. However, the data must satisfy the so-called sphericity condition in addition of the usual assumptions for the test to be valid. It is possible to adjust degrees of freedom to account for possible violation of the sphericity assumption.

2. **The multivariate approach:** This sphericity condition does not need to be satisfied. However, this approach requires a larger number of observation (number of subjects must be larger than number of experimental conditions) and, in general, is less powerful than the univariate approach.

3. **The linear mixed model approach:** This approach is probably the best approach from a theoretical point of view but it is quite complex.

- References: Keselman, H. J., Algina, J., & Kowalchuk, R. K. (2001). The analysis of repeated measures designs: a review. British Journal of Mathematical and Statistical Psychology, 54, 1-20.

The univariate approach

- In the previous examples of ANOVAs, we have assumed that the observations between experimental conditions are uncorrelated (or independent). This assumption is valid if different subjects are used in different experimental conditions. However, this assumption is no more valid if the same subjects are used in several (or all) experimental condition because better subjects in one condition are also likely to perform better in the other conditions.

- In the repeated-measure ANOVA, the data must also satisfy the so-called *sphericity* **(or *circularity*)** condition or the *compound symmetry* condition in addition of the usual assumptions (independence, homogeneity of the variances, and normality) .

- The compound symmetry condition is a stronger assumption than the sphericity condition.

- The sphericity condition needs to apply only to within-subject factors. It is automatically satisfied if the within-subject factor has only two levels.

# Adjusting of the degrees of freedom

- While tests for the sphericity or compound symmetry exist (e.g. Mauchly's test), they are not very reliable because they are quite sensitive to deviations of the normality assumption.

- A better approach is to *adjust the degrees of freedom* in order to make the tests of the repeated measure ANOVA more conservative. Several correction factors exist: Greenhouse-Geisser (1959), Huynh-Feldt (1990) and a lower-bound value which is most conservative (see relevant literature for more details). SPSS will automatically compute the value of these factors.

- To adjust the F test, it is necessary to multiply the two degrees of freedom of the F distribution by the correction factor. Since the value of the correction factor is smaller than 1, this will decrease the degrees of freedom of the F distribution and make, in general, the test more conservative.

Example. RQ data set

RQ data set (see Wayne, Table 8.4.): Analysis of the respiratory quotient (RQ) of 8 patients who followed a special diet. The RQ was measured at the beginning of the diet (day=0), after three days, and after seven days.

```
> rq.l<-read.table("RQ.dat",header=TRUE,sep="\t")
> interaction.plot(rq.l$day,rq.l$su,rq.l$rq,type="b",las=1,col=1,lty=1,fixed=T,pch=1:8,
  trace.label="Subject",xlab="Days",ylab="Respiratory quotient",)
```
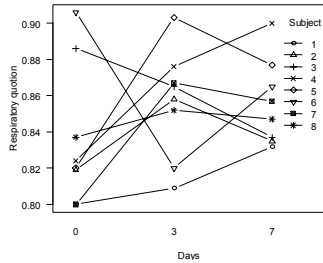

Data Analysis (draft) - Gabriel Baud-Bovy

This repeated-measure experimental design has only **one within-subject factor** (time, with 3 levels). It is an example of longitudinal study.

The underlying model for this experimental design is

$$y_{ij} = \mu + b_i + \beta_j + \varepsilon_{ij}$$

where y_{ij} is the measure done on the ith subject at point j in time ($i =1,..,8$, $j =1,..,3$), b_i is the subject effect and β_j is the diet effect measured at several points in time. Note that this model assumes that there is no interaction between the subject and the time (the hypothetical effect of the diet after three and seven days is the same for all subjects).

Example. RQ data set

• The Error() term in the formula is used to indicate how to compute the denominator (eror) term of the F test

```
> rq.l$su<-factor(rq.l$su)    # codes factors
> rq.l$day<-factor(rq.l$day)  #
> fit<-aov(rq~day+Error(su/day),rq.l)  # ANOVA
> summary(fit)
Error: su
          Df   Sum Sq   Mean Sq F value Pr(>F)
Residuals  7 0.0074380 0.0010626
Error: su:day
          Df   Sum Sq   Mean Sq F value Pr(>F)
day        2 0.0020803 0.0010402  1.0791 0.3666
Residuals 14 0.0134950 0.0009639
```

• Note that the error term of the F test of a within-subject factor corresponds to the interaction between the factor and the subject:

```
> anova(aov(rq~day*su,rq.l))
Analysis of Variance Table

Response: rq
          Df   Sum Sq   Mean Sq F value Pr(>F)
day        2 0.0020803 0.0010402
su         7 0.0074380 0.0010626
day:su    14 0.0134950 0.0009639
Residuals  0 0.0000000
```

• the null hypothesis being tested is H0: βj= 0 (the diet has no effect).

• The dofs for the F test are k-1=2 for the hypothesis being tested where k is the number of the within-subject factor, and $(n$-1)$(k$-1)=14 for the error term where n is the number of subjects

Data Analysis (draft) - Gabriel Baud-Bovy

Example. RQ data set

• **Long (univariate) format**: one colum contain all the observations, additional specify levels corresponding to within as well as between subject factors

```
> head(rq.l)
  su day    rq
1  1   0 0.800
2  1   3 0.809
3  1   7 0.832
4  2   0 0.819
5  2   3 0.858
```

• **Wide (multivariate) format**: each raw correspond to a different subject, columns contains repeated measures that correspond to within-subject factors). Addtional columns specify the levels of between-subject factors.

```
> rq.w<-reshape(rq.l,
      v.names="rq",
      idvar="su",
      timevar="day",
      direction="wide")
> rq.l<-rq.l[order(rq.l$su,rq.l$day),c("su","day","rq")]
> head(rq.w)
  su  rq.0  rq.3  rq.7
1  1 0.800 0.809 0.832
4  2 0.819 0.858 0.835
7  3 0.886 0.865 0.837
...
```

Data Analysis (draft) - Gabriel Baud-Bovy

Example. RQ data set

• To obtain **DoF adjustements**, it is necesary to use the data in the wide format and the specify the columns with the repeated measures in the *rhs* of the formula with `cbind(...)` in `aov` or `lm`.

```
> fit<-aov(cbind(rq.0,rq.3,rq.7)~1,rq.w)
```

• the `idata` argument of anova.mlm is necessary to establish a correspondence between columns of the dataset and the levels of the within-subject factor(s)

```
> (idata<-data.frame(day=factor(c(0,3,7))))
  day
1   0
2   3
3   7
```

• Univariate tests with adjusted DoFs are obtained with `anova.mlm` (anova multivariate) with test="Spherical"

```
> anova(fit,idata=idata,X=~1,test="Spherical")
Analysis of Variance Table
Contrasts orthogonal to
~1
Greenhouse-Geisser epsilon: 0.6945
Huynh-Feldt epsilon:        0.8119
            Df      F num Df den Df  Pr(>F)  G-G Pr  H-F Pr
(Intercept)  1 1.0791      2     14 0.36656 0.35084 0.35805
Residuals    7
```

• The adjusted degrees of freedom are obtained by multiplying the original degree of freedom by the correction factor. For Greenhouse and Geisser correction factor (ε=0.694), the adjusted dofs are 1.289=2x0.694 for the first dof and 9.723=14x0.694 for the second dof.

• Using the adjusted dofs with the F distribution yields typically a more conservative p value (.351 instead of .367). .

Data Analysis (draft) - Gabriel Baud-Bovy

Example. RQ data set

- **Mauchly's test** is sued to to check if sphericity is statisfied

```
> mauchly.test(fit,idata=idata,X=~1)
        Mauchly's test of sphericity
        Contrasts orthogonal to ~1
   data:  SSD matrix from aov(formula = cbind(rq.0, rq.3, rq.7) ~ 1, data = rq.w)
   W = 0.5601, p-value = 0.1757
```

- **Multivariate tests** do not assume sphericity

```
> anova(fit,idata=idata,X=~1,test="Pillai")
Analysis of Variance Table
Contrasts orthogonal to ~1
             Df  Pillai approx F num Df den Df Pr(>F)
(Intercept)  1 0.17622  0.64175      2      6  0.559
Residuals    7
```

Argument `test="Spherical"` gives access to alternative multivariate tests (`"Wilks"`, `"Hotelling-Lawley"`, `"Roy"`, `"Spherical"`),

---

Example. Threshold dataset

trial	1	2	3	4	5	6	7	8	9	10
start	0	10	0	10	0	10	0	10	0	10

SU	Sex									
1	M	5.4	3.9	5.8	4.9	5.2	3.9	3.9	6.3	5.9	5.3
2	F	4.2	5.5	4.8	5.5	5.0	6.1	5.4	6.0	4.0	7.7
3	M	3.6	3.1	4.5	4.9	5.1	5.1	6.2	4.9	5.4	4.5
4	F	4.6	6.1	5.6	2.8	5.9	4.5	5.5	3.0	5.7	8.1
5	M	4.0	4.9	4.9	4.5	5.5	3.7	6.2	5.2	3.6	6.3
6	F	4.3	6.1	4.3	5.3	4.3	4.9	5.9	4.9	6.0	7.6

- The dataset contains the absolute thresholds of 6 subjects (3 males and 3 females). Each subject performed 10 trials with one of two possible different starting values (start=0 and 10).

- We want to test whether there is a difference between the thresholds of the two sexes and ilf the threshold depend on the starting values

- **One between-subject factor** (sex) and **one within-subject factor** (start)

Example. Threshold dataset

```
# define threshold dataset (wide format)
th.w<-data.frame(
        sex=factor(c("M","F","M","F","M","F")),
        y=matrix( c(5.4,3.9,5.8,4.9,5.2,3.9,3.9,6.3,5.9,5.3,
                    4.2,5.5,4.8,5.5,5.0,6.1,5.4,6.0,4.0,7.7,
                    3.6,3.1,4.5,4.9,5.1,5.1,6.2,4.9,5.4,4.5,
                    4.6,6.1,5.6,2.8,5.9,4.5,5.5,3.0,5.7,8.1,
                    4.0,4.9,4.9,4.5,5.5,3.7,6.2,5.2,3.6,6.3,
                    4.3,6.1,4.3,5.3,4.3,4.9,5.9,4.9,6.0,7.6),nrow=6,byrow=TRUE)
)
# define within-subject factors
idata<-data.frame(
        half=factor(rep(1:2,each=5),
        start=factor(rep(c(0,10),5)))
# reshape into long format
th.l<-reshape(th.w,
        varying=paste("y",1:10,sep="."),
        v.names="y",
        idvar="su",
        timevar="trial",
        direction="long")
# add start value
th.l$start<-ifelse(th.l$trial%%2,0,10)
# reorder data
th.l<-th.l[order(th.l$su,th.l$trial),c("su","sex","trial","start","y")]
# define factors
th.l$su<-factor(th.l$su)
th.l$sex<-factor(th.l$sex)
th.l$start<-factor(th.l$start)
```

Example. Threshold data set

```
> summary(aov(y~sex*start+Error(su/start),th.l))
Error: su
          Df  Sum Sq Mean Sq F value  Pr(>F)
sex        1 2.81667 2.81667  13.662 0.02090 *
Residuals  4 0.82467 0.20617
Error: su:start
          Df Sum Sq Mean Sq F value Pr(>F)
start      1 0.3840  0.3840  0.2639 0.6345
sex:start  1 2.5627  2.5627  1.7615 0.2551
Residuals  4 5.8193  1.4548
Error: Within
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals 48 55.252   1.151
```

- Adjustement of dofs is necessary only if the number of dofs > 1.

```
># compute the means across repetition for each starting values
> th.l0<-aggregate(th.l[,"y",drop=FALSE],th.l[,c("su","sex","start")],mean)
> th.l0<-th.l0[order(th.l0$su),]
> # repeated measure ANOVA
> summary(aov(y~sex*start+Error(su/start),th.l0))
Error: su
          Df  Sum Sq Mean Sq F value  Pr(>F)
sex        1 0.56333 0.56333  13.662 0.02090 *
Residuals  4 0.16493 0.04123
Error: su:start
          Df  Sum Sq Mean Sq F value Pr(>F)
start      1 0.07680 0.07680  0.2639 0.6345
sex:start  1 0.51253 0.51253  1.7615 0.2551
Residuals  4 1.16387 0.29097
```

- Within group variability is not taken into acount in the test of withing subject factors
- => ANOVA can be done with mean values

Example. The Elashoff dataset

- The Elashoff dataset (Stevens, Table 13.10): Two groups of eight subjects were given three different doses of two drugs. This experimental design has **two within-subject factors** (dose and drug) and **one between-subject factor** (group).

```
> read.table("elashof.dat",header=TRUE)
> head(ela.l)
  group su drug dose dv
1     1  1    1    1 19
2     1  1    1    2 22
> ela.l$group<-factor(ela.l$group)
> ela.l$su<-factor(ela.l$su)
> ela.l$drug<-factor(ela.l$drug)
> ela.l$dose<-factor(ela.l$dose)
> ela.w<-data.frame(
    group=rep(1:2,each=8),
    matrix(ela.l$dv,nrow=16,byrow=T))
> names(ela.w)<-c("group",outer(c("v1","v2"),1:3,paste,sep=""))
> idata<-data.frame(
    drug=factor(rep(1:2,each=3)),
    dose=factor(rep(1:3,2)))
> ela.w$group<-factor(ela.w$group)
> head(ela.w)
  group su v11 v21 v31 v12 v22 v32
1     1  1  19  22  28  16  26  22
```

- The dataset is in long-format (ela.l)
- The datset is balanced.

- Define factors

- reshape in wide format (ela.w)

- define within-subject factors (idata)

- The questions of interest are: Will the drug be differentially effective for different groups? Is the effectiveness of the drugs dependent on the dose level? Is the effectiveness of the drug dependent on the does level and the group?

Data Analysis (draft) - Gabriel Baud-Bovy

Example. The Elashoff dataset

```
> fit<-aov(dv~group*drug*dose+Error(su/(drug*dose)),ela.l)
> summary(fit)
Error: su
          Df Sum Sq Mean Sq F value Pr(>F)
group      1 270.01  270.01  7.0925 0.01855 *
Residuals 14 532.98   38.07
Error: su:drug
           Df Sum Sq Mean Sq F value  Pr(>F)
drug        1 348.84  348.84  13.001 0.002866 **
group:drug  1 326.34  326.34  12.163 0.003624 **
Residuals  14 375.65   26.83
Error: su:dose
           Df Sum Sq Mean Sq F value   Pr(>F)
dose        2 758.77  379.39 36.5097 1.580e-08 ***
group:dose  2  42.27   21.14  2.0339   0.1497
Residuals  28 290.96   10.39
Error: su:drug:dose
                Df Sum Sq Mean Sq F value Pr(>F)
drug:dose        2 12.062   6.031  0.6815 0.5140
group:drug:dose  2 14.812   7.406  0.8369 0.4436
Residuals       28 247.792  8.850
```

- The is a significant difference in the responses of the two groups ($F(1,14)=7.092,P=0.019$)
- Let us initially assume the sphericity condition.
- This analysis indicates a statistically significant effect of the drug ($F(1,14)=13.001, P=0.003$).
- The significant interactions DRUG*GP ($F(1,14)=12.163,P=0.04$) indicates that the effect of the drug is different for each group.
- The only other significant effect is the dose main effect ($F(2,28)=36.510, P<0.001$).

- The main effect of group is difficult to interpret because there is also statistically significant GROUP*DRUG interaction.)

Data Analysis (draft) - Gabriel Baud-Bovy

Example. The Elashoff dataset

dose main effect

group:drug interaction
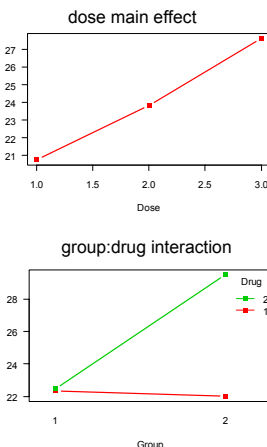
- The first plot shows that the average value of the response increases with the dose. The absence of interaction between the DOSE and the DRUG or the GROUP factor is independent from these factors.

- The interaction plot shows that the response to the second drug is much larger for the second than for the first group.

- Note that the main effect of group is misleading in this case. It is a side-effect of the observation that the response is much higher in the second group for the second drug.

Data Analysis (draft) - Gabriel Baud-Bovy

Example. The Elashoff dataset

- Tests corresponding to sequential (Type 1) SS.

```
> anova(fit,idata=idata,M=~drug,X=~1,test="Spherical")
Contrasts orthogonal to ~1
Contrasts spanned by ~drug
Greenhouse-Geisser epsilon: 1
Huynh-Feldt epsilon:        1
            Df      F num Df den Df    Pr(>F)     G-G Pr     H-F Pr
(Intercept)  1 13.001      1     14 0.0028659 0.0028659 0.0028659
group        1 12.163      1     14 0.0036242 0.0036242 0.0036242
Residuals   14
> anova(fit,idata=idata,M=~drug+dose,X=~drug,test="Spherical")
Contrasts orthogonal to ~drug
Contrasts spanned by ~drug + dose
Greenhouse-Geisser epsilon: 0.8787
Huynh-Feldt epsilon:        0.9949
            Df       F num Df den Df    Pr(>F)     G-G Pr     H-F Pr
(Intercept)  1 36.5097      2     28 1.580e-08 9.785e-08 1.705e-08
group        1  2.0339      2     28   0.14970   0.15648   0.14999
Residuals   14
> anova(fit,idata=idata,M=~drug:dose,X=~drug+dose,test="Spherical")
Contrasts orthogonal to ~drug + dose
Contrasts spanned to ~drug:dose
Greenhouse-Geisser epsilon: 0.7297
Huynh-Feldt epsilon:        0.7931
            Df      F num Df den Df  Pr(>F)  G-G Pr  H-F Pr
(Intercept)  1 0.6815      2     28 0.51404 0.47245 0.48341
group        1 0.8369      2     28 0.44360 0.41338 0.42147
Residuals   14
```

With this dataset, there is no difference between Type I and Type III SS.

Data Analysis (draft) - Gabriel Baud-Bovy

Repeated-measure summary

- Univariate tests with the sphericity condition assumed:
  - aov and the Error in the formula with data in the long format and summary

- Univariate tests with adjusted degrees of freedom or multivariate tests:
  - aov or lm with the data in the wide format and anova.mlm.

- Test of the sphericity condition
  - mauchly.test

- Notes:
  - For repeated measure, within-subject design needs to be balanced. If one cell has no data (missing data), the whole subject needs to be removed.
  - If the design includes replications, making the analysis with the avergage values gives the same result.
  - Testing repeated-measures the "old" way is a bit complicated in R but possible. The preferred R approach is to use mixed models.