Biostatistics

ANOVA - Analysis of Variance

Burkhardt Seifert & Alois Tschopp

Biostatistics Unit
University of Zurich

Analysis of variance

ANOVA = Analysis of variance

- simple example: Two-sample t-test = difference between means in two groups (not differences between variances!)
- analyses and interprets observations of several groups, treatments, conditions, etc.
- decomposes the total variance present in the data into contributions of the single sources of variation:
  systematic contributions = differences of means —
  and random rest = variability around group mean
- complicated example (Stoll, Brühlmann, Stucki, Seifert & Michel (1994). *J. Rheumatology*):
  Muscle strength of 7 patients was measured twice by 3 physicians (42 measurements — analysis of variance for repeated measures with 2 within-factors). Is the new measurement reliable?

Simple example

Example: (Amess et al. 1978) 22 bypass-patients are randomly divided into 3 treatment groups (different respiration). Differ the values of folic acid in red blood cells after 24 h?

Group	1	1	1	1	1	1	1	1
Red cell folate	243	251	275	291	347	354	380	392

Group	2	2	2	2	2	2	2	2	2
Red cell folate	206	210	226	249	255	273	285	295	309

Group	3	3	3	3	3
Red cell folate	241	258	270	293	328

Simple example

- Scientific hypothesis H_1:
  The values of folic acid in the red blood cells differ after 24 h, i.e. the 3 population means μ_1, μ_2, μ_3 are not all the same.
- Null hypothesis: H_0: $\mu_1 = \mu_2 = \mu_3$
- The central result of the analysis of variance is the ANOVA-table:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
(Intercept)	1	1764789.14	1764789.14	844.27	0.0000
group	2	15515.77	7757.88	3.71	0.0436
Residuals	19	39716.10	2090.32		

$R^2 = 0.281$, $R^2_{\text{adj}} = 0.205$

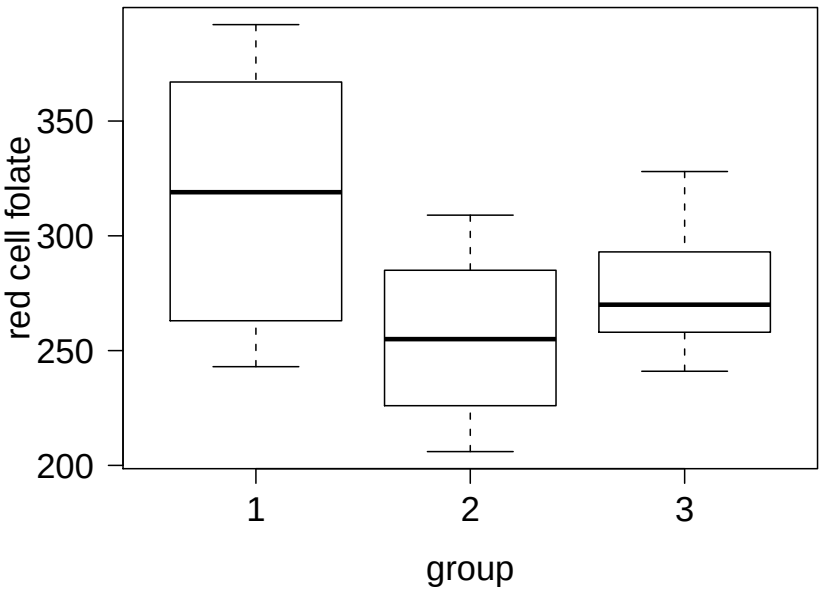
Simple example

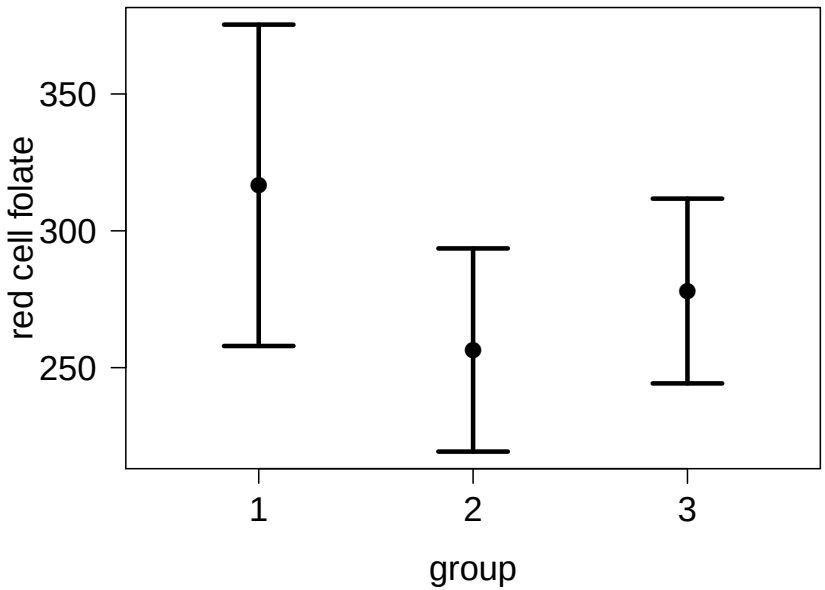
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
(Intercept)	1	1764789.14	1764789.14	844.27	0.0000
group	2	15515.77	7757.88	3.71	0.0436
Residuals	19	39716.10	2090.32		

- Important: p-value ($\Pr(> F)$) $= 0.044$
- Sum of squares (Sum Sq, SS)
- Mean square (Mean Sq, MS) = SS/"degress of freedom (Df)"
- Hypothesis H_0:
  "Groups have the same true mean" $\longrightarrow$ under H_0 have MS_{group} (later MS_T) and $\text{MS}_{\text{Residuals}}$ (later MS_{res}) the same mean.
- Test statistic: $F = \text{MS}_T/\text{MS}_{\text{res}} = 3.71$ times larger than expected under H_0.
- Assumption: Data are normally distributed.
- p-value $p = 0.044$ from $F \sim F_{2,19}$ (see Df)
- MS_{res} is estimated based on all groups, as in the t-test.

Simple example

Graphical presentation





Error Bars show mean ± 1.0 sd
Dots show mean

Simple example

Question: Is it possible to provide evidence of the group differences without an analysis of variance?

3 group comparisons!

	Mean diff.	df	t-value	p-value
1 vs. 2	60.181	15	2.558	0.0218
1 vs. 3	38.625	11	1.327	0.2115
2 vs. 3	-21.556	12	-1.072	0.3046

- significant difference between group 1 versus 2.
- testing of 3 hypotheses
  Bonferroni correction: $p < 0.05/3 = 0.017$ significant
  $\longrightarrow$ no significance
- ANOVA provides p-value for the question:
  "Is there a difference at all?"
- observations pooled for estimation of variance
  $\longrightarrow$ better discriminatory power

Two-sample problem is an ANOVA

- unpaired t-test

t	df	p-value	Mean diff.	lower	upper
2.558	15	0.022	60.18	10.039	110.322

- ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
(Intercept)	1	1378545.94	1378545.94	588.15	0.0000
group	1	15338.96	15338.96	6.54	0.0218
Residuals	15	35158.10	2343.87		

$R^2 = 0.304$, $R^2_{\text{adj}} = 0.257$

Note: $F = t^2$, p-values are identical.

Unpaired t-test as ANOVA

Given 2 samples

$$y_{11}, y_{12}, \ldots, y_{1n_1}$$
$$y_{21}, y_{22}, \ldots, y_{2n_2}$$

with:

- means μ_1 and μ_2
- same variance σ^2
- $n = n_1 + n_2$ observations

Model: $y_{ij} = \mu_i + \varepsilon_{ij} = \mu + \alpha_i + \varepsilon_{ij}$ $(i = 1, 2;\ j = 1, \ldots, n_i)$

$\alpha_i = \mu_i - \mu$ is called (treatment-) effect

Unpaired t-test as ANOVA

Decompose total sum of squares $\mathrm{SS}_{\mathrm{total}}$:

$$\begin{aligned}
\mathrm{SS}_{\mathrm{total}} &= \sum_{j=1}^{n_1} (y_{1j} - \bar{y})^2 + \sum_{j=1}^{n_2} (y_{2j} - \bar{y})^2 \\
&= \sum_{j=1}^{n_1} (y_{1j} - \bar{y}_1 + \bar{y}_1 - \bar{y})^2 + \sum_{j=1}^{n_2} (y_{2j} - \bar{y}_2 + \bar{y}_2 - \bar{y})^2 \\
&= \underbrace{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}_{} + \underbrace{n_1(\bar{y}_1 - \bar{y})^2 + n_2(\bar{y}_2 - \bar{y})^2}_{}
\end{aligned}$$

(mixed products disappear)

$$\begin{aligned}
&= \mathrm{SS}_{\mathrm{res}} + \mathrm{SS}_{\mathrm{T}} \\
(\quad &= \text{residual SS} + \text{Treatment SS} \quad) \\
&= \text{SS within groups} + \text{SS between groups}
\end{aligned}$$

Unpaired t-test as ANOVA

- SS_T corresponds to squared enumerator $(\bar{y}_1 - \bar{y}_2)^2$ of the t-statistic

$$SS_T = n_1(\bar{y}_1 - \bar{y})^2 + n_2(\bar{y}_2 - \bar{y})^2$$

$$= n_1 \left(\bar{y}_1 - \frac{n_1\bar{y}_1 + n_2\bar{y}_2}{n_1 + n_2} \right)^2 + n_2 \left(\bar{y}_2 - \frac{n_1\bar{y}_1 + n_2\bar{y}_2}{n_1 + n_2} \right)^2$$

$$= \frac{n_1 n_2}{n_1 + n_2}(\bar{y}_1 - \bar{y}_2)^2$$

- SS_{res} corresponds to denominator of the t-statistic

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Unpaired t-test as ANOVA

Definition degrees of freedom (df):

(df of SS) = (# squared elements) - (# linear restrictions)

$\text{df}(\text{SS}_{\text{res}}) = n_1 - 1 + n_2 - 1 = n - 2$

2 restrictions: $\sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i) = 0$

$\text{df}(\text{SS}_{\text{T}}) = 2 - 1 = 1$

1 restriction: $n_1(\bar{y}_1 - \bar{y}) + n_2(\bar{y}_2 - \bar{y}) = 0$

- Degrees of freedom sum up to $n - 1$

Definition mean squares (MS): $\text{MS} = \text{SS}/\text{df}$

Pooled variance:

Mean variability around μ_1 and μ_2

$$\hat{\sigma}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)} = \text{MS}_{\text{res}}$$

Unpaired t-test as ANOVA

Null hypothesis H_0: $\mu_1 = \mu_2$ or $\alpha_1 = \alpha_2 = 0$

F-test

$$(\bar{Y}_1 - \bar{Y}_2) \sim \mathcal{N}\left(\mu_1 - \mu_2, \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\sigma^2\right)$$

$$\longrightarrow \mathrm{E}\left[\bar{Y}_1 - \bar{Y}_2\right]^2 = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\sigma^2 + (\mu_1 - \mu_2)^2$$

$$\longrightarrow \mathrm{E}\left[\mathrm{MS_T}\right] = \mathrm{E}\left[\frac{n_1\ n_2}{n_1 + n_2}\left(\bar{Y}_1 - \bar{Y}_2\right)^2\right] = \sigma^2 + \underbrace{\frac{n_1\ n_2}{n_1 + n_2}(\mu_1 - \mu_2)^2}_{\geq 0}$$

$$\mathrm{E}\left[\mathrm{MS_{res}}\right] = \sigma^2$$

$$\boxed{F = \mathrm{MS_T}/\mathrm{MS_{res}}}$$

Here: $F = t^2$

One-way ANOVA

Generalisation of the two-sample t-test from 2 to m groups

Model: "completely randomized design"

$$y_{ij} = \mu_i + \varepsilon_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \qquad i = 1, \ldots, m\,, j = 1, \ldots, n_i$$

$$\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$$

- Decomposition of the observations:

$$\begin{aligned} y_{ij} &= \hat{\mu} + (\bar{y}_i - \hat{\mu}) + (y_{ij} - \bar{y}_i) \\ &= \hat{\mu} + \hat{\alpha}_i + e_{ij} \\ &= \text{"overall mean"} + \text{effekt} + \text{residual} \\ &\quad \text{(everything estimated)} \end{aligned}$$

  - well-defined by restrictions; What does "overall mean" stand for?
  - meaningful and usual: $\hat{\mu} = \frac{1}{m}\sum_{i=1}^{m} \bar{y}_i \longrightarrow \sum_{i=1}^{m} \alpha_i = 0$

Scientific hypothesis H_1: at least one $\alpha_i \neq 0$

Null hypothesis H_0: all $\alpha_i = 0$; "all group means are equal"

One-way ANOVA

Central: ANOVA-table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
(Intercept)	1	1764789.14	1764789.14	844.27	0.0000
group	2	15515.77	7757.88	3.71	0.0436
Residuals	19	39716.10	2090.32		

- ANOVA decomposes variance of the observations ("total") into contributions of the single sources (sources of variation):
  - group = between groups: variability of the group means (treatments $\longrightarrow SS_T$), systematic contribution
  - Residuals = within groups: variability of the observations within one group (residuals $\longrightarrow SS_{res}$), random contribution

One-way ANOVA

- Degrees of freedom (df)
  $=$ (number of squared elements)$-$(number of restrictions)

  (total $n-1$, like for the variance s^2)

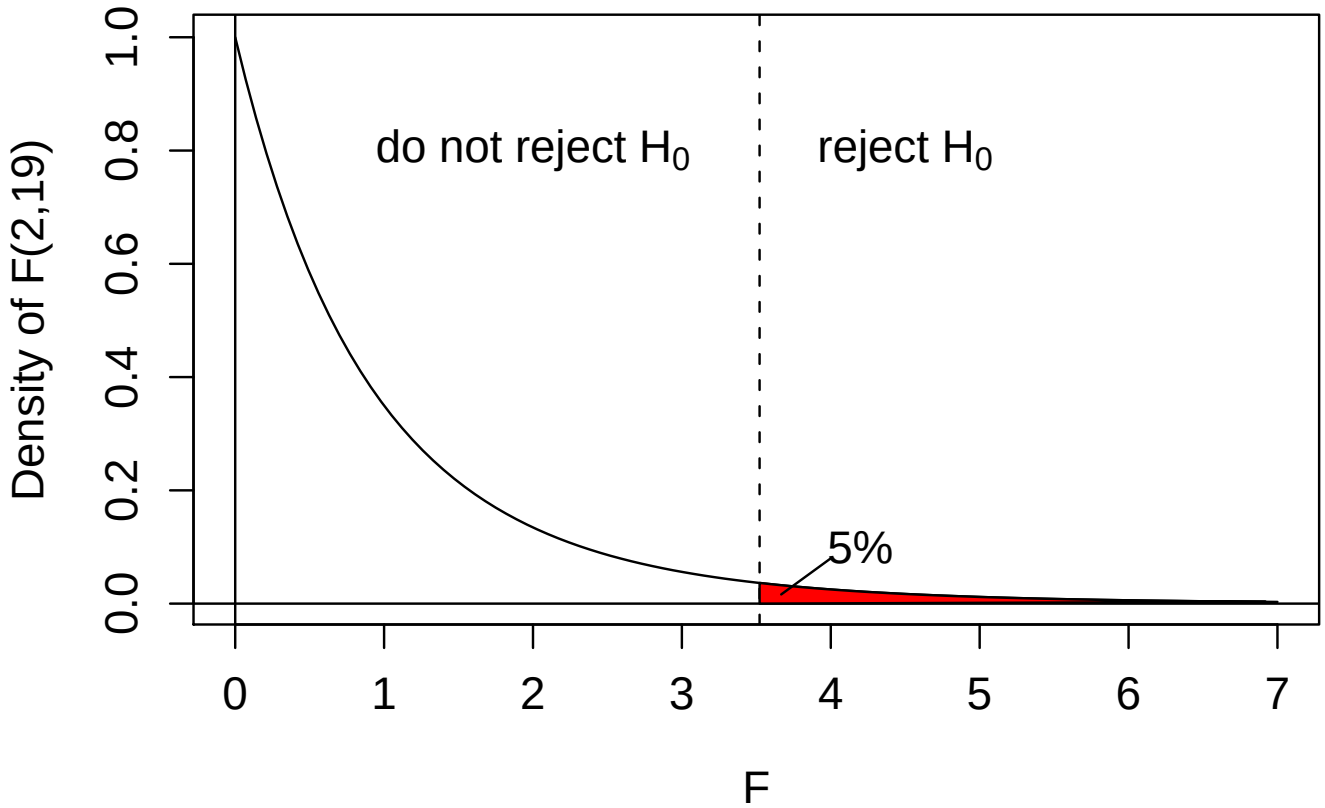
  are also decomposed:
  - between groups:
    m group means $-$ 1 restriction $= m-1 = 2$
  - within groups: n observations $-$ m groups $= n-m = 19$
- mean squares: SS/df
- sum of squares SS_T and $\mathrm{SS}_{\mathrm{res}}$ are independent,
- under H_0 have MS_{T} and $\mathrm{MS}_{\mathrm{res}}$ the same mean σ^2.
- under H_1 is MS_{T} large, $\mathrm{MS}_{\mathrm{res}}$ not influenced.

$$\longrightarrow \boxed{F = \mathrm{MS}_{\mathrm{T}} \,/\, \mathrm{MS}_{\mathrm{res}} \sim F_{m-1,n-m}}$$

One-way ANOVA

In the example: (m=3; n=22)

$F = 3.7 \longrightarrow p - \text{value } p = 0.044$



Test always two-sided.

♣ Confidence intervals

In the case of two groups ("t-test") we received:

$$\bar{y}_1 - \bar{y}_2 - t_{n-2,1-\alpha/2} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{y}_1 - \bar{y}_2 + t_{n-2,1-\alpha/2} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Generalisation

- $\hat{\sigma}^2 = s^2 = \text{MS}_{\text{res}} = 2090$ is the pooled residual variance estimation for all groups

$\longrightarrow SE(\bar{y}_i) = s_{\text{res}} / \sqrt{n_i}$

$\longrightarrow$ Confidence interval for μ_i:

$$\bar{y}_i - s_{\text{res}}\, t_{n-m,1-\alpha/2} \,/\, \sqrt{n_i} \leq \mu_i \leq \bar{y}_i + s_{\text{res}}\, t_{n-m,1-\alpha/2} \,/\, \sqrt{n_i}$$

♣ Confidence intervals

$$\mathrm{SE}(\bar{y}_{i_1} - \bar{y}_{i_2}) = s_{\mathrm{res}} \times \sqrt{\frac{1}{n_{i_1}} + \frac{1}{n_{i_2}}}$$

$\longrightarrow$

$(1-\alpha)$ Confidence interval for difference of the means:

$$\bar{y}_{i_1} - \bar{y}_{i_2} - t_{n-m,1-\alpha/2} s_{\mathrm{res}} \sqrt{\frac{1}{n_{i_1}} + \frac{1}{n_{i_2}}}$$

$$\leq \mu_{i_1} - \mu_{i_2} \leq \bar{y}_{i_1} - \bar{y}_{i_2} + t_{n-m,1-\alpha/2} s_{\mathrm{res}} \sqrt{\frac{1}{n_{i_1}} + \frac{1}{n_{i_2}}}$$

$\longrightarrow$ multiple decision problem

Post-hoc tests

- Analysis of variance answers global questions:
  Are there any differences between the means?
- More specific questions:
  Differ certain pairs or groups of mean values?

Suggestion:

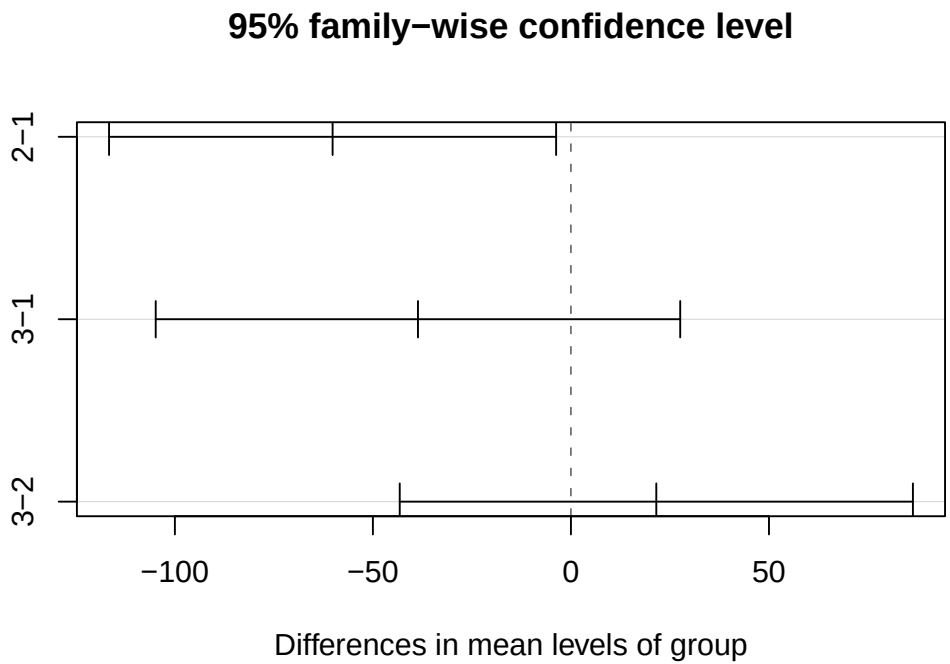
- continue with post-hoc tests only, if the p-value of the analysis of variance < 0.05
- choose a priori plausible and interesting differences
  (the less, the better)
- modified t-tests with joint s_{res} calculated from ANOVA and p-values corrected using the Bonferroni-method (Bonferroni-Dunn-test).

Example

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = lm_red)

$group
         diff        lwr       upr     p adj
2-1 -60.18056 -116.61904 -3.742070 0.0354792
3-1 -38.62500 -104.84037 27.590371 0.3214767
3-2  21.55556  -43.22951 86.340620 0.6802018
```



Non-parametric analysis of variance: Kruskal-Wallis test

Generalisation of the Mann-Whitney test without assuming a normal distribution and based on ranks

- Have observations in all groups the same distribution? (the same variance is indirectly postulated)
- Analysis of variance of the ranks

Example:

Group	N	Mean rank
1	8	15.00
2	9	8.56
3	5	11.20

```
        Kruskal-Wallis rank sum test

data:  redcellfolate by group
Kruskal-Wallis chi-squared = 4.1852, df = 2, p-value = 0.1234
```

Random effects

- Example: Muscle strength was measured for each of 7 patients three times (21 measurements).
  Is the new measurement technique reliable?

  The one-way analysis of variance answers:
  Are there differences in the patients?
- However, individual patients are not of interest.
- Assumption: Patients are randomly chosen. Muscle strength is normally distributed.
- Fixed effects make statements about the levels of the factor (not generalizable).
- Random effects make statements about the population (generalizable).

$\longrightarrow$ Decision fixed/random depends on the goal of the analysis.

Random effects

Model

$y_{ij} = \mu + a_i + \varepsilon_{ij}$

$i = 1, \ldots, m\,,\ j = 1, \ldots, n_i$

$a_i \sim \mathcal{N}(0, \sigma_A^2)$ – Patient–effekt

$\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ – measurement error

a_i and ε_{ij} are supposed to be independent

$\mathrm{Var}(y_{ij}) = \sigma_A^2 + \sigma^2$

Variance components

- σ_A^2 and σ^2 are called variance components
- In the balanced model ($n_1 = \ldots = n_m = J$)
  variance components estimated from

  $$\mathrm{E}(\mathrm{MS_T}) = J\sigma_A^2 + \sigma^2$$

  $$\mathrm{E}(\mathrm{MS_{res}}) = \sigma^2$$

  $$\longrightarrow \hat{\sigma}_A = (\mathrm{MS_T} - \mathrm{MS_{res}}) \,/\, J \quad \text{(ANOVA method)}$$

- Always plan to use balanced designs!

Intraclass correlation

- $\mathrm{Cov}(y_{ij_1}, y_{ij_2}) = \sigma_A^2$

Intraclass correlation coefficient:

$$\rho_I = \frac{\sigma_A^2}{\sigma_A^2 + \sigma^2}$$

- Measure for the reliability of the measuring method:

  $0 \leq \rho_I \leq 1$

  $\rho_I = 0 \longrightarrow \sigma^2 = \infty \longrightarrow$ completely useless

  $\rho_I = 1 \longrightarrow \sigma^2 = 0 \longrightarrow$ no measurement error

Two-way ANOVA

Goal: Comparisons of means with respect to two factors

Example: Expiratory flow with cystic fibrosis

PEmax	BMP	sex	status
95	68	0	light
85	65	1	light
100	64	0	light
85	67	1	light
95	93	0	normal
...	...	...	...

1. Factor A: underweight:
   BMP (BMI as % of the age-specific median for healthy people) grouped into light ($< 80\%$) and normal ($\geq 80\%$)
2. Factor B: gender

Model: two-way cross classification

"completely randomised block design"

$$
\begin{aligned}
y_{ijk} &= \mu_{ij} + \varepsilon_{ijk} \\
i &= 1, \ldots, m_1 && \text{— levels of A} \\
j &= 1, \ldots, m_2 && \text{— levels of B} \\
k &= 1, \ldots, n_{ij} \geq 0 && \text{— replications} \\
\varepsilon_{ijk} &\sim \mathcal{N}(0, \sigma^2)
\end{aligned}
$$

All levels of A are "crossed" with all levels of B

Two-way cross classification

- Decomposition of means:

$$\begin{aligned}\mu_{ij} &= \mu + (\mu_i - \mu) + (\mu_j - \mu) + (\mu_{ij} - \mu_i - \mu_j + \mu) \\ &= \mu + \alpha_i + \beta_j + \gamma_{ij} \\ &= \text{“overall mean”} + \text{main effect of A} \\ &\quad + \text{main effect of B} + \text{interaction of A and B}\end{aligned}$$

- unique through restrictions
  What does “overall mean”, “main effect” mean?

$\longrightarrow$ Type I, II, III, IV sums of squares

New: interactions γ_{ij} (“specific effects”)

$\longrightarrow$ 2 models:

- additive model: $\mu_{ij} = \mu + \alpha_i + \beta_j$
- model with interactions: $\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$

Two-way cross classification

Scientific hypothesis H_1:

(1) underweight has an impact, i.e. certain $\alpha_i \neq 0$

(2) the expiratory flow differs among men and women
i.e. certain $\beta_j \neq 0$

(3) the difference between light and patients with normal weight is gender-specific, i.e. certain $\gamma_{ij} \neq 0$

Null hypothesis H_0:

(1)' All $\alpha_i = 0$

(2)' All $\beta_j = 0$

(3)' All $\gamma_{ij} = 0$

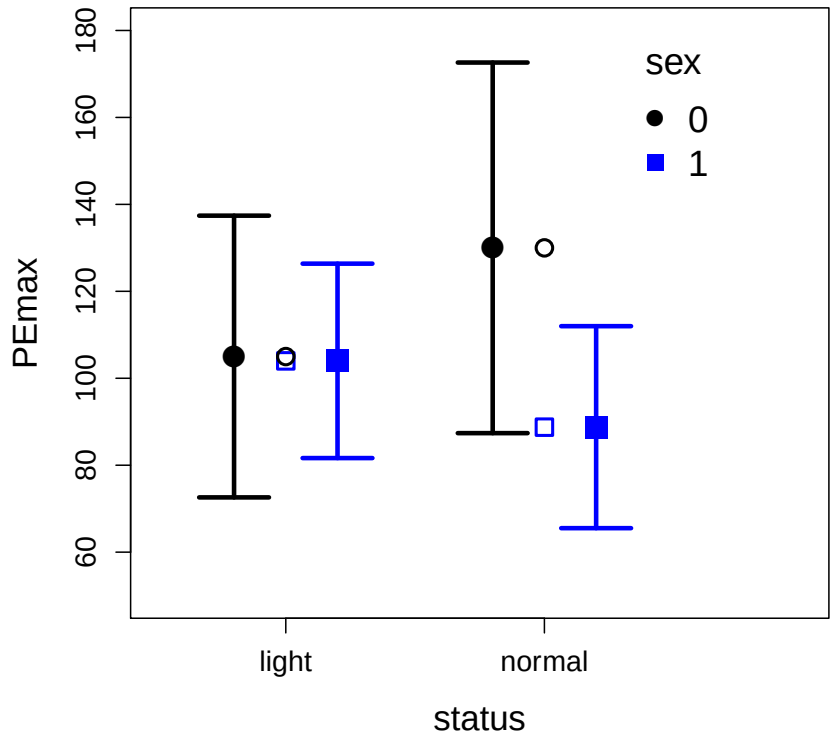
Example: expiratory flow with cystic fibrosis

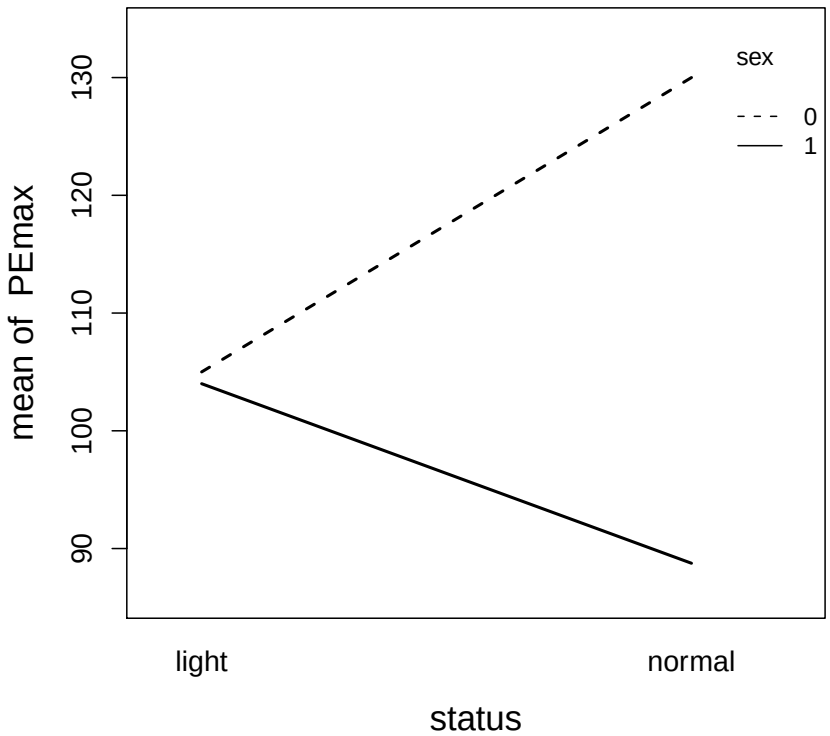
ANOVA Table (Type III tests)

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	269640.09	1	259.52	0.0000
sex	2630.62	1	2.53	0.1265
status	140.09	1	0.13	0.7171
sex:status	2387.46	1	2.30	0.1445
Residuals	21818.75	21		

(Function `Anova` in R-package `car`)

Example: expiratory flow with cystic fibrosis





Interaction if differences are not parallel.

Hierarchical ANOVA

Also: nested ANOVA

Example: X-rays of patients were rated by 3 general practitioners (GP) and 3 specialists (all different patients)

Questions:

- Do specialists rate better than GPs?
- How do specialists differ?
- How do GPs differ?

The person-related effect B (6 raters) is nested within (hierarchically subordinate to) the effect of qualification A.

Model: "hierarchical two-way classification"

$$y_{ijk} = \mu + \alpha_i + \beta_{j:i} + \varepsilon_{ijk}$$

(read "$j : i$" as "j within i")

$i = 1, \ldots, m_1$ – levels of A ($m_1 = 2$: 1–specialist, 2–GP)
$j = 1, \ldots, m_2(i)$ – levels of B:A ($m_2(1) = m_2(2) = 3$)
$k = 1, \ldots, n_{ij}$ – replications

$$\varepsilon_{ijk} \sim \mathcal{N}(0, \sigma^2)$$

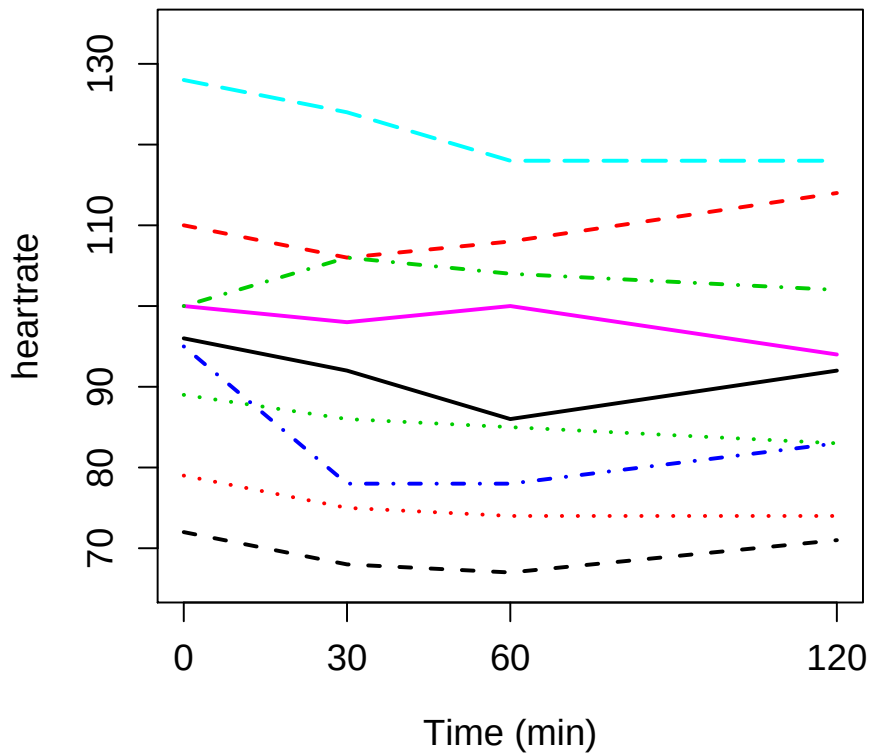
- b_j would not make sense!

Repeated measures ANOVA

- Earlier: Two investigations of the same sample
  $\longrightarrow$ paired t–test, not two-sample t–test
- Generalisation from 2 to more measuring times

Example: Short-term effect of a drug on the heart-rate of 9 patients with heart disease

	Time (min)			
Subject	0	30	60	120
1	96	92	86	92
2	110	106	108	114
3	89	86	85	83
4	95	78	78	83
5	128	124	118	118
6	100	98	100	94
7	72	68	67	71
8	79	75	74	74
9	100	106	104	102



Model: repeated measures ANOVA

$$y_{ij} = \mu + \alpha_i + b_j(t_i) + \varepsilon_{ij}$$

t_i – time points, measuring times, $i = 1, \ldots, m$

$\mu + \alpha_i$ – mean trend

$j = 1, \ldots, J$ – individuals

$b_j(t_i)$ – individual (random) effect of person j at time t_i

$\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$

1. multivariate one-way model (MANOVA)
   $\mathrm{Cov}(b_j(t_{i_1}), b_j(t_{i_2})) = \sigma_{i_1 i_2}$ (un-structured)

Repeated measures ANOVA

2. univariate ANOVA for repeated measures:
   $\mathrm{Cov}(b_j(t_{i_1}), b_j(t_{i_2})) = \sigma_s^2$ (compound symmetry)

$$
\begin{aligned}
y_{ij} &= \mu + \alpha_i + b_j + \varepsilon_{ij} \\
b_j &\sim \mathcal{N}(0, \sigma_s^2) \text{ – person (subject) effect}
\end{aligned}
$$

- Assumption of "compound symmetry" rarely valid for more than 2 measuring times
- Solution: Greenhouse–Geisser correction for deviations from "compound symmetry""

Idea: Estimation (reduction) of degrees of freedom

Example: short-term effect of drug on heartrate

(Function `Anova` in R-package `car`)

```
Type III Repeated Measures MANOVA Tests:
----------------------------------------
Sum of squares and products for the hypothesis:
           time1      time2      time3
time1 160.444444  8.4444444 -46.444444
time2   8.444444  0.4444444  -2.444444
time3 -46.444444 -2.4444444  13.444444

Sum of squares and products for error:
          time1     time2     time3
time1 217.55556  59.55556  26.44444
time2  59.55556 175.55556 125.44444
time3  26.44444 125.44444 143.55556

Multivariate Tests: time
                 Df test stat  approx F num Df den Df  Pr(>F)
Pillai            1 0.4930757 1.9453626      3      6 0.22367
Wilks             1 0.5069243 1.9453626      3      6 0.22367
Hotelling-Lawley  1 0.9726813 1.9453626      3      6 0.22367
Roy               1 0.9726813 1.9453626      3      6 0.22367
```

Example: short-term effect of drug on heartrate

```
Univariate Type III Repeated-Measures ANOVA Assuming Sphericity

               SS num Df Error SS den Df        F    Pr(>F)
(Intercept) 312295      1     8967      8 278.6307 1.678e-07 ***
time           151      3      297     24   4.0696   0.01802 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Mauchly Tests for Sphericity

     Test statistic p-value
time        0.47063 0.41220

Greenhouse-Geisser and Huynh-Feldt Corrections for Departure from Sphericity

      GG eps Pr(>F[GG])
time 0.70654    0.03412 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

     HF eps Pr(>F[HF])
time  0.968    0.01931 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Example: short-term effect of drug on heartrate

Usually, the course is not only observed in one, but two or more groups are considered.

Question: Do the courses differ across groups? If yes, how?

Principles of the analysis of variance

1. Assumption: Samples stem from normally distributed population with equal variances. Do not assume, but verify.
2. As variances are equal within all groups, all observations are used to estimate the variance (pooling). $\longrightarrow$ more degrees of freedom, better power
3. Estimated pooled variance s^2_{res} is also used for the computation of confidence intervals.
4. Following the analysis of variance, investigate residuals, i.e. the deviations of the individual observations from the respective group mean. If normal distribution or equal variances cannot be confirmed, transform data or use Kruskal-Wallis test.