



Introduction and basic concepts

Analysis of repeated measurements 2015

Julie Lyng Forman & Lene Theil Skovgaard

Department of Biostatistics, University of Copenhagen



Outline

About the course

What are repeated measurements?

Warm up: Paired data

Basics of longitudinal data

The multivariate normal distribution

Analysis of response profiles

SAS proc mixed

Baseline adjustment



Aim of the course

To make the participants able to:

- ▶ understand and interpret advanced statistical analyses
- ▶ judge the assumptions behind the use of various methods of analyses
- ▶ perform own analyses using SAS
- ▶ understand output from a statistical program package - in general, i.e. other than SAS
- ▶ present results from a statistical analysis - numerically and graphically

To create a better platform for communication between 'users' of statistics and statisticians, to benefit subsequent collaboration

We expect students to ...

Be **motivated** (from your own research project)

Have an **open mind** towards mathematical model descriptions.

Have **basic knowledge** of statistical concepts such as:

- ▶ the normal distribution
- ▶ mean, average, variance, standard deviation
- ▶ estimate, standard error, confidence interval
- ▶ correlation, regression, ANOVA, linear models.
- ▶ t-test, χ^2 -test, F-test
- ▶ generalized linear models (logistic/possession regression)

Have **patience** to learn statistical programming.



Topics for the course

Models for **dependent data**.

Quatitative outcomes (normal distribution):

- ▶ Linear mixed models
- ▶ Variance component models

Binary and count outcomes:

- ▶ Generalized linear mixed models
- ▶ Population average models (Generalized estimating equations)

Not covered:

- ▶ Censored data (survival analysis)
- ▶ Multivariate data (several different outcomes at once)



Recommended reading

The lecture notes, exercises etc at:

- ▶ <http://staff.pubhealth.ku.dk/~jufo/RepeatedMeasuresE2015>

The book:

- ▶ G.M. Fitzmaurice, N.M. Laird & J.H. Ware :
Applied Longitudinal Analysis (2nd edition),
John Wiley & sons, 2011

We teach **SAS** programming.

- ▶ Additional R and Stata examples can be found at:
www.biostat.harvard.edu/~fitzmaur/ala2e

Teaching activities

Lectures:

- ▶ Thursday and Friday mornings (10.15–13.00)
- ▶ **Copies of overheads must be downloaded in advance**
- ▶ 1-2 coffee breaks (coffee+cake)

Computers labs:

- ▶ In the afternoon (13.30-16.30) following each lecture
- ▶ Coffee, tea, and fruit will be served
- ▶ **Problems and datasets must be downloaded in advance**
- ▶ Solutions can be downloaded after classes



Course diploma

To pass the course 80% attendance is required.

- ▶ It is **your responsibility** to sign the lists
- ▶ ... **each morning**
- ▶ ... **each afternoon**.
- ▶ Note: $6 \times 2 = 12$ lists, at least 10 half days.

There is **no compulsory home work** ...

- ▶ but to benefit from the course you need to **work with the material at home!**
- ▶ or perhaps with your own data :)



Topics for the course

Today: Basic concepts for correlated and clustered data.

- ▶ Describing quantitative data (normal distribution?).
- ▶ Analysis of response profiles for balanced longitudinal data.
- ▶ **Case:** Baseline-follow up study.

Rest of the course:

- ▶ Unbalanced data (models for mean and covariance).
- ▶ Clustered data (variance component and multi-level models)
- ▶ Various designs (linear mixed models): e.g. cross over and repeatability and reproducibility of measurement methods.
- ▶ Non-normal outcomes: Binary and count data.
- ▶ Model diagnostics and Missing data.



Disclaimer: Some of the datasets we use in our case studies are in fact **too small to yield interesting conclusions**. But due to their small size they are very useful for **illustrative purposes**.

Outline

About the course

What are repeated measurements?

Warm up: Paired data

Basics of longitudinal data

The multivariate normal distribution

Analysis of response profiles

SAS proc mixed

Baseline adjustment

What are repeated measurements?



Repeated measurements refer to data where the same outcome has been measured in different situations (or at different spots) **on the same subjects**.

- ▶ Special case: **longitudinal** means **repeatedly over time**.



Examples of repeated measurements

Subjects should be understood in a **wide sense**:

- ▶ Repeated measurements on a patient or person.
- ▶ or on a mouse, dog, blood sample, or cell line.

Replicates can be made:

- ▶ Over time.
- ▶ Under different circumstances/treatments.
- ▶ With different measurement device.
- ▶ On different limbs or locations of the body.



Repeated measurements are termed **clustered data** when the same outcome is measured **on groups of subjects** that are somehow related.

Examples of clustered data

Clusters could be:

- ▶ Siblings, families, or school classes.
- ▶ Clinics, hospitals, or GPs.

But also:

- ▶ Litters or cages.
- ▶ Plates (in a laboratory experiment).

Or any kind of measurements in an experiment clustered within:

- ▶ Operators, days, or sessions.



Statistical analysis

The **usual assumption** is that observations are **independent**.

If you have clustered or repeated measurements the assumption of independence is **violated**.

- ▶ Your analyses must account for the repetitions/clustering.

Warning: Ignoring the repetitions/clustering and doing a standard analysis most often leads to:

- ▶ p-values that are too small or too large.
- ▶ confidence intervals that are too wide or too narrow.



Outline

About the course

What are repeated measurements?

Warm up: Paired data

Basics of longitudinal data

The multivariate normal distribution

Analysis of response profiles

SAS proc mixed

Baseline adjustment

17 / 80

Paired data

The most simple example of clustered or repeated measurements.

- ▶ Two replicates or two subjects per cluster

Examples of paired data:

- ▶ Same person with treatment and placebo.
- ▶ Baseline-follow up study.
- ▶ Twin study.
- ▶ Comparison of two measurement methods or reliability of a measurement method

Quantitative outcome analysed with the **paired t-test**.



18 / 80



Example: Energy intake

Daily dietary intake for 10 women recorded over 10 pre-menstrual and 10 post-menstrual days.

| Subject | Pre-menstrual | Post-menstrual | Difference |
|---------|---------------|----------------|------------|
| 1 | 5260 | 3910 | 1350 |
| 2 | 5470 | 4220 | 1250 |
| 3 | 5640 | 3885 | 1755 |
| 4 | 6180 | 5160 | 1020 |
| 5 | 6390 | 5645 | 745 |
| 6 | 6515 | 4680 | 1835 |
| 7 | 6805 | 5265 | 1540 |
| 8 | 7515 | 5975 | 1540 |
| 9 | 7515 | 6790 | 725 |
| 10 | 8230 | 6900 | 1330 |
| 11 | 8770 | 7335 | 1435 |
| Mean | 6753.6 | 5433.2 | 1320.5 |
| SD | 1142.1 | 1216.8 | 366.7 |

19 / 80



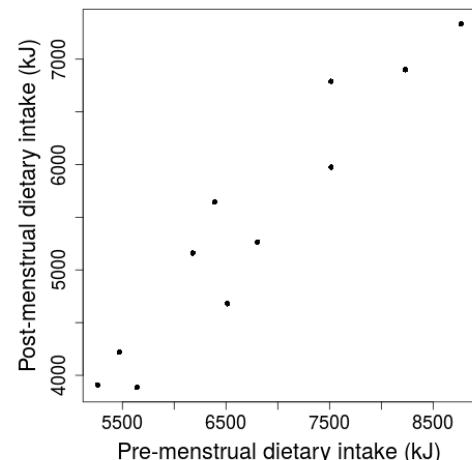
20 / 80



Example: Energy intake

Strong association between pre- and post-menstrual dietary intake:

- ▶ **Correlation:** 0.95 (95% CI: 0.83;0.99).



Covariance and correlation

Used to describe (linear) association between dependent variables
assumed to have a joint normal distribution.

The **covariance** between two measurements is:

$$\text{Cov}(Y_1, Y_2) = E\{(Y_1 - \mu_1)(Y_2 - \mu_2)\}$$

It has the squared units of the measurements.

The **correlation** between two measurements

$$\text{Cor}(Y_1, Y_2) = \frac{\text{Cov}(Y_1, Y_2)}{\text{SD}(Y_1)\text{SD}(Y_2)}$$

It has no units - interpretation is free of scale.

- ▶ If Y_1 and Y_2 are **independent**, the correlation is = 0.
- ▶ If a perfect linear association is found the correlation is ± 1 .

Example: Paired vs unpaired comparison

To compare pre-menstrual and post-menstrual dietary intake.

- ▶ Test $H_0 : \mu_1 = \mu_2$.
- ▶ Find a confidence interval for $\mu_1 - \mu_2$.

Very different results from the **paired t-test (correct analysis)** and the **two-sample t-test (wrong analysis)**:

| Analysis | Estimate (95% CI) | P-value |
|-------------------|-------------------|-----------|
| paired t-test | 1320 (1074;1567) | 0.0000003 |
| two-sample t-test | 1320 (271; 2370) | P=0.01625 |

Explanation: Standard error in paired testing

Estimated difference between the two means:

$$\bar{Y}_2 - \bar{Y}_1$$

Statistical uncertainty is assessed by the **standard error** of the differences between the sample means:

$$\text{s.e.}(\bar{Y}_2 - \bar{Y}_1) = \sqrt{\frac{\text{Var}(Y_1)}{n} + \frac{\text{Var}(Y_2)}{n} - 2\frac{\text{Cov}(Y_1, Y_2)}{n}}$$

Note the covariance!

Outline

About the course

What are repeated measurements?

Warm up: Paired data

Basics of longitudinal data

The multivariate normal distribution

Analysis of response profiles

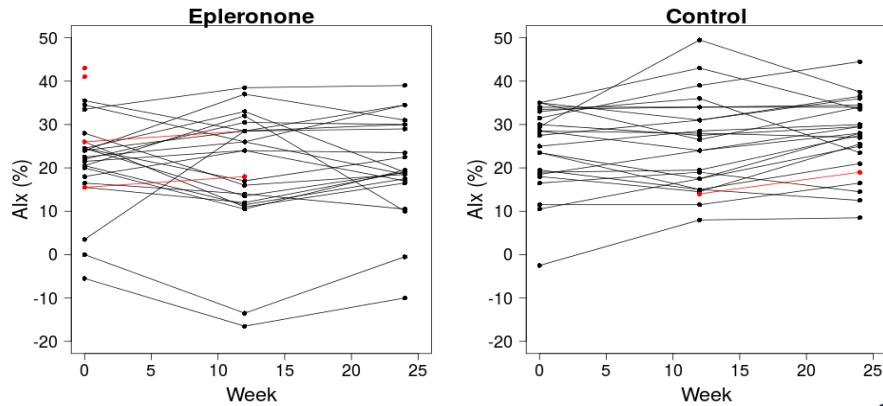
SAS proc mixed

Baseline adjustment

Case study: Baseline follow-up

Response: Augmentation index (aix) in patients with CKD.

- ▶ Comparison of *Eplerenone* to standard treatment.
- ▶ Follow-up after 12 and 24 weeks.



Boesby et al: *Eplerenone Attenuates Pulse Wave Reflection in Chronic Kidney Disease Stage 3-4 - A Randomized Controlled Study*, PLOS ONE 2013.

Making spaghettiplots with SAS

The spaghettiplots were made with:

```
PROC SGPANEL DATA=ckd;
PANELBY treat;
SERIES x = week y = aix / GROUP=id;
RUN;
```

Data must be in the *long format* for drawing the spaghettiplots.



26 / 80

Typical set-up for longitudinal measurements

Two or more **groups** of subjects

- ▶ Often receiving different treatments
- ▶ Possibly **randomised at baseline**.

Longitudinal measurements of the same quantity over time for each subject, typically as a function of

- ▶ time (i.e. duration of treatment or disease)
- ▶ age

Do the time courses differ between the groups?

Merits of longitudinal studies

In longitudinal studies measurements are taken repeatedly on the same subjects over time.

- ▶ This allows us to **study changes over time within subjects** and factors that influence these changes, e.g. treatment.
- ▶ By comparing each subjects responses at two or more occasions we eliminate extraneous but unavoidable sources of variability among subjects. Thus we obtain **more accurate estimates** and **more certain conclusions** about changes over time than in cross-sectional studies.



27 / 80

Balanced and complete data

In a planned study the times of measurements will usually be the same for all subjects. We have a **balanced design**

In practice data is most often somewhat unbalanced due to drop-out, missed visits, failed measurements.

- ▶ In this case we say that data is **incomplete**.
- ▶ But still **the design is balanced**.

Data from (retrospective) observational studies are most often **unbalanced** both by design and in practice.

Unbalanced designs are treated in lecture 2.

Missing data is treated in lecture 4.

Outline

About the course

What are repeated measurements?

Warm up: Paired data

Basics of longitudinal data

The multivariate normal distribution

Analysis of response profiles

SAS proc mixed

Baseline adjustment



The distribution of repeated outcomes

Repeated measurements are characterized by being

- ▶ mutually dependent or **correlated**.

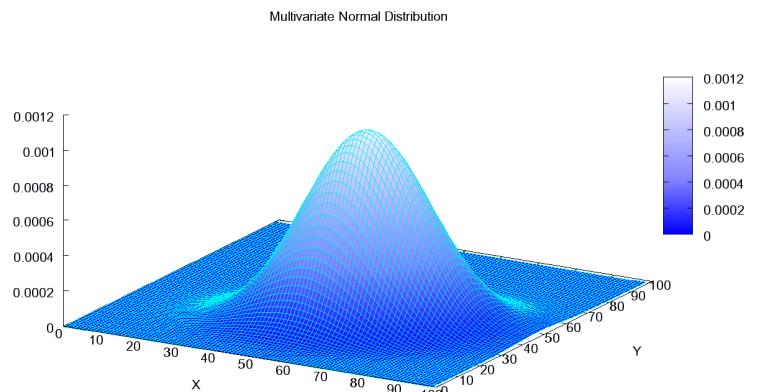
We need to characterize their **joint distribution**.

Standard model for quantitative data: The **multivariate normal**

- ▶ Location: mean-**vector**
- ▶ Variability: covariance-**matrix**



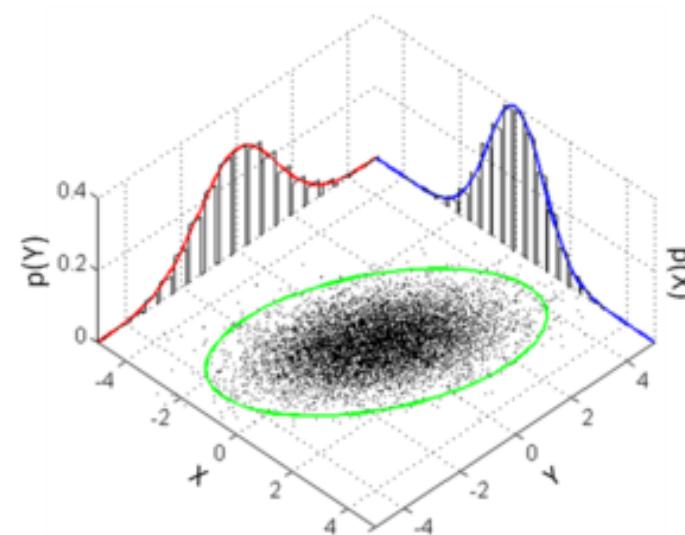
The multivariate normal distribution



Source: Wikipedia.



The multivariate normal distribution



Source: Wikipedia.

Eplerenone: Correlations

Pairs of correlation for the three time points:

| | | Pearson Correlation Coefficients | | |
|------|---------|----------------------------------|---------|--|
| | | Number of Observations | | |
| | aix0 | aix1 | aix2 | |
| aix0 | 1.00000 | 0.78707 | 0.76148 | |
| | 24 | 23 | 23 | |
| aix1 | 0.78707 | 1.00000 | 0.79525 | |
| | 23 | 24 | 24 | |
| aix2 | 0.76148 | 0.79525 | 1.00000 | |
| | 23 | 24 | 24 | |

| | | Pearson Correlation Coefficients | | |
|------|---------|----------------------------------|---------|--|
| | | Number of Observations | | |
| | aix0 | aix1 | aix2 | |
| aix0 | 1.00000 | 0.67942 | 0.72694 | |
| | 26 | 24 | 22 | |
| aix1 | 0.67942 | 1.00000 | 0.81741 | |
| | 24 | 24 | 22 | |
| aix2 | 0.72694 | 0.81741 | 1.00000 | |
| | 22 | 22 | 22 | |

But correlations can be misleading if data are not normal.

Eplerenone: Summary statistics

Means and std.devs for the three time points:

trt=0

Simple Statistics

| Variable | N | Mean | Std Dev | Sum | Minimum | Maximum |
|----------|----|----------|----------|-----------|----------|----------|
| aix0 | 24 | 24.64583 | 9.37559 | 591.50000 | -2.50000 | 35.00000 |
| aix1 | 24 | 25.31250 | 10.60333 | 607.50000 | 8.00000 | 49.50000 |
| aix2 | 24 | 27.33333 | 8.70490 | 656.00000 | 8.50000 | 44.50000 |

trt=1

Simple Statistics

| Variable | N | Mean | Std Dev | Sum | Minimum | Maximum |
|----------|----|----------|----------|-----------|-----------|----------|
| aix0 | 26 | 22.28846 | 11.11321 | 579.50000 | -5.50000 | 43.00000 |
| aix1 | 24 | 19.93750 | 13.69966 | 478.50000 | -16.50000 | 38.50000 |
| aix2 | 22 | 20.38636 | 11.43192 | 448.50000 | -10.00000 | 39.00000 |

But this does not tell us anything about the joint distribution.



34 / 80



Matrix notation

Covariances and correlations of the 3D (normal) distribution:

$$\text{Cov} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{pmatrix}, \quad \text{Cor} = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{pmatrix}$$

NOTE:

- ▶ Variances $\sigma_1^2, \sigma_2^2, \sigma_3^2$ along the diagonal in Cov.
- ▶ Covariances are symmetric $\sigma_{ij} = \sigma_{ji}$.
- ▶ 1's along the diagonal in Cor.
- ▶ Correlations are $\rho_{ij} = \sigma_{ij}/\sqrt{\sigma_i^2 \cdot \sigma_j^2}$.

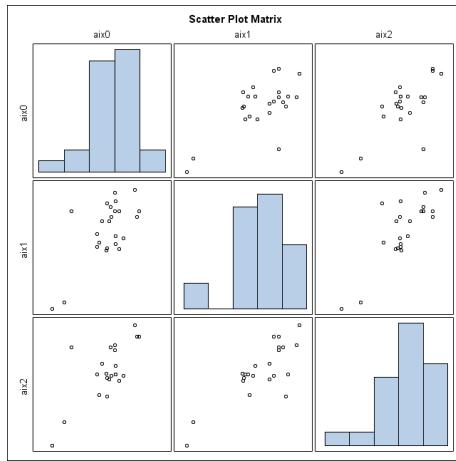


36 / 80

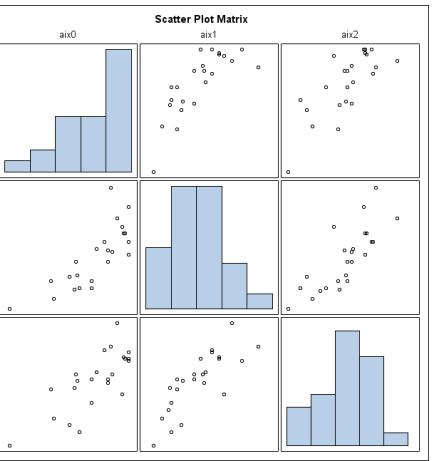


Eplerenone: Scatter plots

Left: Eplerenone.



Right: Controls.



Better check of normal distribution: use residual diagnostics (lecture 4).

37 / 80



Summary statistics and plot made in SAS

Program-file: eplerenone.sas

```
PROC SORT DATA=ckd;
BY treat;
RUN;
```

```
ODS GRAPHICS ON;
PROC CORR DATA=ckd PLOT=MATRIX(HISTOGRAM) NOPROB;
BY treat;
VAR aix0-aix2;
RUN;
ODS GRAPHICS OFF;
```

Note: Applies to data in the *wide format*.

38 / 80



Is normality really needed?

The **standard assumption** is that outcomes from the same subject follow a **multivariate normal distribution**.

But: the **linear mixed models for repeated outcomes are robust**.

- ▶ If sample size is not too small.
- ▶ If the distribution of the data is not too skewed.

Highly skewed data should always be **transformed**.

Outline

About the course

What are repeated measurements?

Warm up: Paired data

Basics of longitudinal data

The multivariate normal distribution

Analysis of response profiles

SAS proc mixed

Baseline adjustment



Analysis of response profiles

Comparison of change over n time points within g groups of subjects (e.g. different treatments).

- ▶ Similar to **two-way ANOVA** only with correlated data.
- ▶ **Covariates:** group and time
- ▶ **Balanced design**, but possibly incomplete data.

Do the groups evolve differently with time?

We further need to model the covariance-matrix:

- ▶ With a balanced design and few different time points we don't have to make any restrictions.
- ▶ An **unstructured covariance** is assumed.



Unstructured covariance

(aka unrestricted covariance)

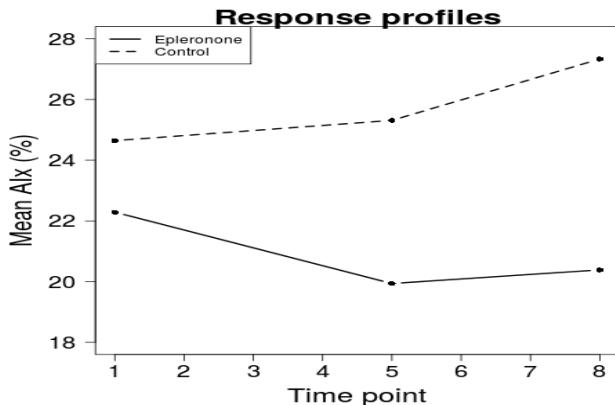
Fully flexible because no assumptions are made about the covariance as a function of time.

- ▶ One variance parameter for each time point
- ▶ One correlation parameter for each pair of time points
- ▶ $n + \frac{n(n-1)}{2}$ parameters in total with n time points.

Becomes **infeasible with many time points**.

Case study: Eplerenone

Individual curves are roughly parallel and few data are missing, so we look at averages over time.



Seeming improvement over time with Eplerenone.



Plotting time-averages with SAS

The plot of time-averages were made with:

```
PROC MEANS DATA=ckd NWAY NOPRINT;
CLASS treat week;
VAR aix;
OUTPUT OUT=ckdmeans MEAN=average;
RUN;
```

```
PROC SGLOT DATA=ckdmeans;
SERIES x = week y = average / GROUP = treat markers;
RUN;
```

Data must be in the *long format* for computing the time-averages.



Two-way ANOVA model for the means

Means for the six time-treatment combinations described as:

| | group = Control | group = Eplerenone |
|---------|---------------------|---|
| time=0 | β_1 | $\beta_1 + \beta_4$ |
| time=12 | $\beta_1 + \beta_2$ | $\beta_1 + \beta_2 + \beta_4 + \beta_5$ |
| time=24 | $\beta_1 + \beta_3$ | $\beta_1 + \beta_3 + \beta_4 + \beta_6$ |

- ▶ Mean of standard treatment at baseline is reference (intercept)
- ▶ Change over time with standard treatment (time estimates)
- ▶ Difference between groups at baseline (group estimates)
- ▶ Interaction: differences in time/group effect

Parameter estimates (program: slide 56)

| Effect | week | treat | Estimate | StdError | DF | t Value | Pr > t |
|------------|------|-------|----------|----------|------|---------|---------|
| Intercept | | | 24.3431 | 2.0793 | 49.4 | 11.71 | <.0001 |
| week | 12 | | 1.0887 | 1.7694 | 46.2 | 0.62 | 0.5414 |
| week | 24 | | 3.0895 | 1.4995 | 44.5 | 2.06 | 0.0452 |
| week | 0 | | 0 | . | . | . | . |
| treat | | 1 | -2.0547 | 2.8999 | 48.9 | -0.71 | 0.4820 |
| treat | | 0 | 0 | . | . | . | . |
| week*treat | 12 | 1 | -1.9493 | 2.4871 | 45.8 | -0.78 | 0.4372 |
| week*treat | 12 | 0 | 0 | . | . | . | . |
| week*treat | 24 | 1 | -3.6078 | 2.1298 | 45.3 | -1.69 | 0.0971 |
| week*treat | 24 | 0 | 0 | . | . | . | . |
| week*treat | 0 | 1 | 0 | . | . | . | . |
| week*treat | 0 | 0 | 0 | . | . | . | . |

Estimated R Matrix

| Row | Col1 | Col2 | Col3 |
|-----|---------|---------|---------|
| 1 | 106.23 | 96.3802 | 80.1893 |
| 2 | 96.3802 | 159.64 | 106.48 |
| 3 | 80.1893 | 106.48 | 106.38 |

Estimated R Correlation Matrix

| Row | Col1 | Col2 | Col3 |
|-----|--------|--------|--------|
| 1 | 1.0000 | 0.7401 | 0.7544 |
| 2 | 0.7401 | 1.0000 | 0.8171 |
| 3 | 0.7544 | 0.8171 | 1.0000 |

Two-way ANOVA: hypotheses

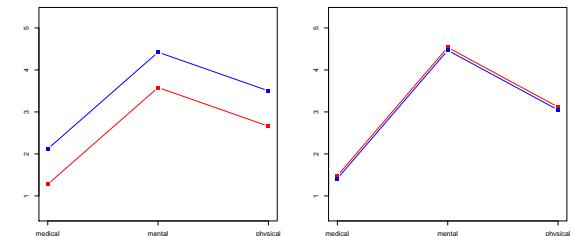
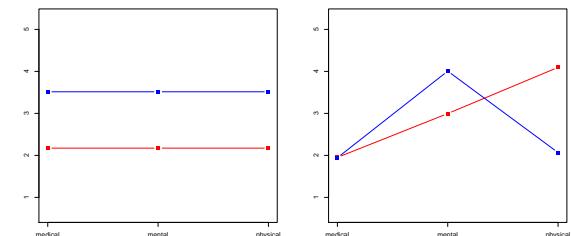
The starting point is a model with an interaction term is assumed

- ▶ No restrictions on the means of the six combination groups.

Test hypotheses to simplify the description of the time and group effects:

- H_0 : The effects of time and group are additive in other words there is no interaction.
- H_1 : There is no effect of time.
- H_2 : There is no effect of group.
- H_3 : There is no effect of either treatment or group

Visualizing model and hypothesis



Hypothesis testing I: Overall test of interaction

H_0 : No group*time-interaction.

- I.e. mean changes over time are identical in all groups (we test whether the response profiles are parallel).

Type 3 Tests of Fixed Effects

| | Num | Den | | | |
|------------|-----|------|---------|--------|--|
| Effect | DF | DF | F Value | Pr > F | |
| week | 2 | 44.5 | 0.99 | 0.3794 | |
| treat | 1 | 47 | 1.84 | 0.1817 | |
| week*treat | 2 | 44.5 | 1.43 | 0.2490 | |

- For further testing the model must be updated.
Or use sequential TYPE1-tests instead.

Hypothesis testing II: effect at final follow-up

Usual interest in a randomized baseline follow-up study:

H_0 : Change at last follow-up is the same in both groups

- In term of the ANOVA-model $H_0 : \beta_6 = 0$.
- Hence, check the parameter estimates in your output.

| Effect | week | treat | Estimate | StdError | DF | t Value | Pr > t |
|------------|------|-------|----------|----------|------|---------|---------|
| week*treat | 24 | 1 | -3.6078 | 2.1298 | 45.3 | -1.69 | 0.0971 |

- Seemingly improvement, but non-significant difference.
- But:** we haven't made any **baseline adjustment**.

LSMeans (program on slide 64)

Test the differences between groups at each specific time point.

| Differences of Least Squares Means | | | | | | | | | | |
|------------------------------------|------|-------|-------|--------|----------|--------|------|---------|---------|--|
| Effect | week | treat | _week | _treat | Estimate | Error | DF | t Value | Pr > t | |
| week*treat | 12 | 1 | 12 | 0 | -4.0040 | 3.5909 | 45.5 | -1.12 | 0.2707 | |
| week*treat | 12 | 1 | 24 | 1 | -0.3423 | 1.5282 | 46.8 | -0.22 | 0.8237 | |
| week*treat | 12 | 1 | 24 | 0 | -6.0048 | 3.2739 | 54.8 | -1.83 | 0.0721 | |
| week*treat | 12 | 1 | 0 | 1 | -0.8606 | 1.7478 | 45.4 | -0.49 | 0.6248 | |
| week*treat | 12 | 1 | 0 | 0 | -2.9153 | 3.2720 | 62 | -0.89 | 0.3764 | |
| week*treat | 12 | 0 | 24 | 1 | 3.6617 | 3.2988 | 55.4 | 1.11 | 0.2718 | |
| week*treat | 12 | 0 | 24 | 0 | -2.0008 | 1.4868 | 44.9 | -1.35 | 0.1852 | |
| week*treat | 12 | 0 | 0 | 1 | 3.1434 | 3.2554 | 60.1 | 0.97 | 0.3381 | |
| week*treat | 12 | 0 | 0 | 0 | 1.0887 | 1.7694 | 46.2 | 0.62 | 0.5414 | |
| week*treat | 24 | 1 | 24 | 0 | -5.6625 | 2.9505 | 46 | -1.92 | 0.0612 | |
| week*treat | 24 | 1 | 0 | 1 | -0.5183 | 1.5125 | 46 | -0.34 | 0.7334 | |
| week*treat | 24 | 1 | 0 | 0 | -2.5730 | 2.9485 | 63.7 | -0.87 | 0.3861 | |
| week*treat | 24 | 0 | 0 | 1 | 5.1442 | 2.9019 | 60.7 | 1.77 | 0.0813 | |
| week*treat | 24 | 0 | 0 | 0 | 3.0895 | 1.4995 | 44.5 | 2.06 | 0.0452 | |
| week*treat | 0 | 1 | 0 | 0 | -2.0547 | 2.8999 | 48.9 | -0.71 | 0.4820 | |

Tests of Effect Slices

| Effect | week | Num | Den | F Value | Pr > F |
|------------|------|-----|------|---------|--------|
| week*treat | 12 | 1 | 45.5 | 1.24 | 0.2707 |
| week*treat | 24 | 1 | 46 | 3.68 | 0.0612 |
| week*treat | 0 | 1 | 48.9 | 0.50 | 0.4820 |



Analysis of response profiles - drawbacks

- Can only handle balanced designs.
- Is no good with many groups or many time points, due to too many model parameters.
- Do not make use of a priori known data patterns, e.g.
 - Correlation decreasing with time.
 - Steadily increasing response to treatment.

Lecture 2: Models for the mean and the covariance.



Outline

About the course

What are repeated measurements?

Warm up: Paired data

Basics of longitudinal data

The multivariate normal distribution

Analysis of response profiles

SAS proc mixed

Baseline adjustment

Preparing data for analysis

To fit a **linear mixed model** with **any statistical software** data must be in the so-called **long format**:

- ▶ Each dataline contains only one observation of the outcome.
- ▶ A time-variable identifies the time of measurement.
- ▶ The id identifies measurements from the same person.

Contrary to the **wide format** where the dataset contains one outcome variable (column) for each occasion:

| id | sex | age | treat | aix0 | aix1 | aix2 |
|----|-----|-----|-------|------|------|------|
| 1 | 1 | 57 | 0 | 10.5 | 17.5 | 25.0 |
| 2 | 1 | 48 | 0 | -2.5 | 8.0 | 8.5 |
| 3 | 2 | 54 | 1 | 18.0 | 24.0 | 23.5 |

...

Exercises: transform data from the wide to the long format.

Long format

| Obs | id | sex | age | treat | aix0 | week | aix |
|-----|----|-----|-----|-------|------|------|------|
| 1 | 1 | 1 | 57 | 0 | 10.5 | 0 | 10.5 |
| 2 | 1 | 1 | 57 | 0 | 10.5 | 12 | 17.5 |
| 3 | 1 | 1 | 57 | 0 | 10.5 | 24 | 25.0 |
| 4 | 2 | 1 | 48 | 0 | -2.5 | 0 | -2.5 |
| 5 | 2 | 1 | 48 | 0 | -2.5 | 12 | 8.0 |
| 6 | 2 | 1 | 48 | 0 | -2.5 | 24 | 8.5 |
| 7 | 3 | 2 | 54 | 1 | 18.0 | 0 | 18.0 |
| 8 | 3 | 2 | 54 | 1 | 18.0 | 12 | 24.0 |
| 9 | 3 | 2 | 54 | 1 | 18.0 | 24 | 23.5 |
| 10 | 4 | 2 | 46 | 1 | 26.0 | 0 | 26.0 |
| 11 | 4 | 2 | 46 | 1 | 26.0 | 12 | 11.0 |
| 12 | 4 | 2 | 46 | 1 | 26.0 | 24 | 16.5 |

Syntax: Analysis of response profiles

```
PROC MIXED DATA=ckd;
CLASS id week (ref='0') treat (ref='0');
MODEL aix = week treat treat*week
      / SOLUTION DDFM=KR OUTPM=ckdfit;
REPEATED week / SUBJECT=id TYPE=UN R RCORR;
RUN;
```

- ▶ Syntax is similar to PROC GLM with a MODEL-statement specifying the (lin) relation between outcome and covariates.
- ▶ Categorical variable must be declared with CLASS.
- ▶ The model for the covariance (UN=unstructured) is specified in a separate REPEATED-statement.
- ▶ Predicted means are saved in a dataset ckdfit.

The option DDFM=KENWARDROGERS (aka KR)

(or DDFM=SATTERTHWAITE).

A technical option intended to improve the statistical performance of the F-tests.

- ▶ It has no effect on balanced data.
- ▶ In unbalanced situations (i.e for almost all observational designs and in case of missing observations) degrees of freedom are computed by a more complicated formulae.
- ▶ The computations may require a little more time, but in most cases this will not be noticeable.

When in doubt, use it!

SAS: proc mixed output

Dimensions

| | |
|-----------------------|----|
| Covariance Parameters | 6 |
| Columns in X | 12 |
| Columns in Z | 0 |
| Subjects | 51 |
| Max Obs Per Subject | 3 |

Number of Observations

| | |
|---------------------------------|-----|
| Number of Observations Read | 153 |
| Number of Observations Used | 144 |
| Number of Observations Not Used | 9 |

Iteration History

| Iteration | Evaluations | -2 Res Log Like | Criterion |
|-----------|-------------|-----------------|------------|
| 0 | 1 | 1070.85454941 | |
| 1 | 2 | 982.86560047 | 0.00144735 |
| 2 | 1 | 982.26253864 | 0.00009905 |
| 3 | 1 | 982.22468047 | 0.00000061 |
| 4 | 1 | 982.22445749 | 0.00000000 |

Convergence criteria met.

Always check that the numerical optimisation has converged.

SAS: proc mixed output

The Mixed Procedure

Model Information

| | |
|---------------------------|---------------|
| Data Set | WORK.CKD_LONG |
| Dependent Variable | aix |
| Covariance Structure | Unstructured |
| Subject Effect | id |
| Estimation Method | REML |
| Residual Variance Method | None |
| Fixed Effects SE Method | Kenward-Roger |
| Degrees of Freedom Method | Kenward-Roger |

Class Level Information

| Class | Levels | Values |
|-------|--------|---|
| id | 51 | 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 45 46 47 48 49 51 52 53 54 |
| week | 3 | 12 24 0 |
| treat | 2 | 1 0 |



SAS: proc mixed output

Options R and RCORR makes SAS print the estimated covariance and correlation matrices.

The Mixed Procedure

Estimated R Matrix for id 1

| Row | Col1 | Col2 | Col3 |
|-----|---------|---------|---------|
| 1 | 106.23 | 96.3802 | 80.1893 |
| 2 | 96.3802 | 159.64 | 106.48 |
| 3 | 80.1893 | 106.48 | 106.38 |

Estimated R Correlation Matrix for id 1

| Row | Col1 | Col2 | Col3 |
|-----|--------|--------|--------|
| 1 | 1.0000 | 0.7401 | 0.7544 |
| 2 | 0.7401 | 1.0000 | 0.8171 |
| 3 | 0.7544 | 0.8171 | 1.0000 |



SAS: proc mixed output

Fit Statistics

| | |
|--------------------------|--------|
| -2 Res Log Likelihood | 982.2 |
| AIC (smaller is better) | 994.2 |
| AICC (smaller is better) | 994.9 |
| BIC (smaller is better) | 1005.8 |

Null Model Likelihood Ratio Test

| DF | Chi-Square | Pr > ChiSq |
|----|------------|------------|
| 5 | 88.63 | <.0001 |

Used for comparison of different models*.

* Make sure to use the PROC MIXED METHOD=ML-option if you want to use this to test nested models for the mean-structure (lecture 2).

SAS: proc mixed output

At last: Parameter estimates and tests.

Solution for Fixed Effects

| Effect | week | treat | Estimate | StdError | DF | t Value | Pr > t |
|------------|------|-------|----------|----------|------|---------|---------|
| Intercept | | | 24.3431 | 2.0793 | 49.4 | 11.71 | <.0001 |
| week | 12 | | 1.0887 | 1.7694 | 46.2 | 0.62 | 0.5414 |
| week | 24 | | 3.0895 | 1.4995 | 44.5 | 2.06 | 0.0452 |
| week | 0 | | 0 | . | . | . | . |
| treat | | 1 | -2.0547 | 2.8999 | 48.9 | -0.71 | 0.4820 |
| treat | | 0 | 0 | . | . | . | . |
| week*treat | 12 | 1 | -1.9493 | 2.4871 | 45.8 | -0.78 | 0.4372 |
| week*treat | 12 | 0 | 0 | . | . | . | . |
| week*treat | 24 | 1 | -3.6078 | 2.1298 | 45.3 | -1.69 | 0.0971 |
| week*treat | 24 | 0 | 0 | . | . | . | . |
| week*treat | 0 | 1 | 0 | . | . | . | . |
| week*treat | 0 | 0 | 0 | . | . | . | . |

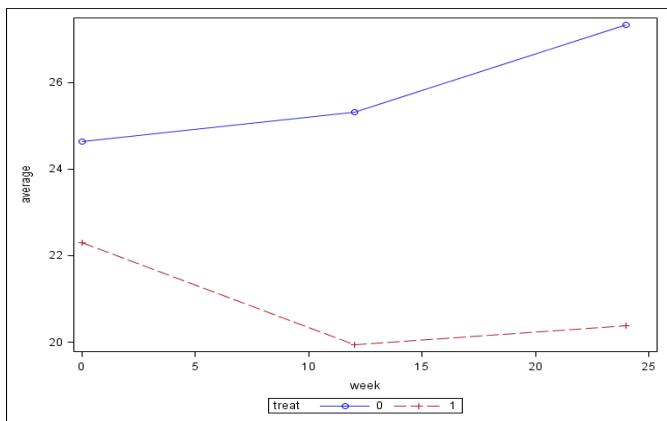
Type 3 Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|------------|--------|--------|---------|--------|
| week | 2 | 44.5 | 0.99 | 0.3794 |
| treat | 1 | 47 | 1.84 | 0.1817 |
| week*treat | 2 | 44.5 | 1.43 | 0.2490 |

Plotting the estimated response profiles

```
PROC SORT DATA=ckdfit; BY treat week id; RUN;

PROC SGPLOT DATA=ckdfit;
SERIES x = week y = pred / GROUP = treat MARKERS;
RUN;
```



Alternative syntax: LSMeans

```
PROC MIXED DATA=ckd;
CLASS id week treat;
MODEL aix = treat*week / NOINT DDFM=KR;
LSMEANS treat*week / DIFF SLICE=week CL;
REPEATED week / SUBJECT=id TYPE=UN R RCORR;
RUN;
```

- ▶ Estimates the means for all times and treatments,
- ▶ ... and **all possible differences** between them (DIFF-option).
- ▶ NOINT means that the model does not include an intercept (so there is no need to specify reference groups)
- ▶ Use SLICE=week to test for overall differences between multiple groups at each time separately (one-way ANOVA).



Outline

About the course

What are repeated measurements?

Warm up: Paired data

Basics of longitudinal data

The multivariate normal distribution

Analysis of response profiles

SAS proc mixed

Baseline adjustment

Hypothetical comparison of two treatment groups

What happens if we ignore the baseline problem?

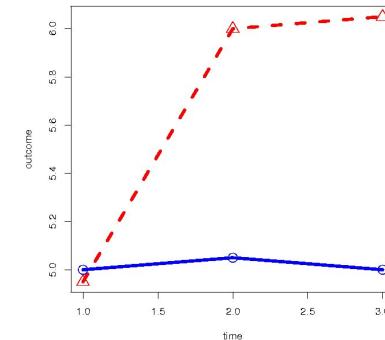
- ▶ We are spending information on estimating a non-existing difference. Thus, the power of the test is reduced.

So should we leave out the baseline measurement?

- ▶ We loose information about change over time and again the power of the test of treatment effect is reduced.

Baseline measurements

In randomized clinical trials, the first measurement is often a **baseline measurement**.



- ▶ The group means **must** be equal at baseline.

Constrained model (cLMM)

| | group = Control | group = Eplerenone |
|---------|---------------------|-----------------------------------|
| time=0 | β_1 | $\beta_1 + 0$ |
| time=12 | $\beta_1 + \beta_2$ | $\beta_1 + \beta_2 + 0 + \beta_4$ |
| time=24 | $\beta_1 + \beta_3$ | $\beta_1 + \beta_3 + 0 + \beta_5$ |

- ▶ Intercept.
- ▶ Time effect with standard treatment
- ▶ Difference between groups at baseline = 0!
- ▶ Interactions (differences in time-effects)



Baseline adjustment: Constrained model

Constrained linear mixed model (cLMM):

- ▶ Analysis of response profiles.
- ▶ Leave out the main-effect of treat **and** redefine the treatment variable to get identical group means at baseline.

```
DATA ckd;
SET ckd;
treatadj = treat;
IF week = 0 THEN treatadj = 0;
RUN;

PROC MIXED DATA=ckd;
CLASS id week (ref='0') treat_adj (ref='0');
MODEL aix = week treat_adj*week / SOLUTION CL DDFM=KR;
REPEATED week / SUBJECT=id TYPE=UN R RCORR;
RUN;
```

69 / 80

Eplerenone: cLMM output

Note: Covariance estimates hardly change when the constraint is put on the mean parameters at baseline.

Estimated R Matrix for id 1

| Row | Col1 | Col2 | Col3 |
|-----|---------|---------|---------|
| 1 | 105.29 | 95.5196 | 79.4628 |
| 2 | 95.5196 | 158.77 | 105.76 |
| 3 | 79.4628 | 105.76 | 105.77 |

Estimated R Correlation Matrix for id 1

| Row | Col1 | Col2 | Col3 |
|-----|--------|--------|--------|
| 1 | 1.0000 | 0.7388 | 0.7530 |
| 2 | 0.7388 | 1.0000 | 0.8161 |
| 3 | 0.7530 | 0.8161 | 1.0000 |

Eplerenone: cLMM output

| Effect | week | treatadj | Estimate | StdError | DF | t Value | Pr > t | Alpha |
|---------------|------|----------|----------|----------|------|---------|---------|-------|
| Intercept | | | 23.2879 | 1.4430 | 50 | 16.14 | <.0001 | 0.05 |
| week | 12 | | 1.2017 | 1.7816 | 46.8 | 0.67 | 0.5033 | 0.05 |
| week | 24 | | 3.3608 | 1.4643 | 48.3 | 2.30 | 0.0261 | 0.05 |
| week | 0 | | 0 | | . | . | . | . |
| week*treatadj | 12 | 1 | -2.1552 | 2.5240 | 46 | -0.85 | 0.3976 | 0.05 |
| week*treatadj | 12 | 0 | 0 | . | . | . | . | . |
| week*treatadj | 24 | 1 | -4.1247 | 2.0436 | 45.9 | -2.02 | 0.0494 | 0.05 |
| week*treatadj | 24 | 0 | 0 | . | . | . | . | . |
| week*treatadj | 0 | 0 | 0 | . | . | . | . | . |

| Effect | week | treatadj | Lower | Upper |
|---------------|------|----------|---------|----------|
| Intercept | | | 20.3895 | 26.1862 |
| week | 12 | | -2.3828 | 4.7861 |
| week | 24 | | 0.4171 | 6.3045 |
| week | 0 | | . | . |
| week*treatadj | 12 | 1 | -7.2356 | 2.9253 |
| week*treatadj | 12 | 0 | . | . |
| week*treatadj | 24 | 1 | -8.2385 | -0.01085 |
| week*treatadj | 24 | 0 | . | . |
| week*treatadj | 0 | 0 | . | . |

Conclusion: Significant difference at last follow-up. Estimated difference in change over time -4.12% (95% CI: -8.14% to -0.11% P=0.0442) in favor of Eplerenone.



Classical approaches for handling baseline

Vickers & Altman, *Analysing controlled clinical trials with baseline follow-up measurements*, BMJ 323, 1123–1124.

Three possibilities: 1. End point, 2. Change, 3. ANCOVA

1. Do a two-sample t-test on the end point measurements.
 - ▶ Discarding baseline is ok if the correlation is small.
2. Do a two-sample t-test on the changes.
 - ▶ Subtracting baseline is ok if the correlation is large.
3. Use baseline as a covariate; do an ANCOVA.
 - ▶ Conditioning on baseline is always ok and most powerful.

Conclusion: ANCOVA is most efficient.



Why ANCOVA is superior (technical)

Consider a comparison of treatment vs placebo with only one time of follow-up, where for simplicity we assume*

- ▶ Same variance σ^2 at both time points and correlation ρ .

Implied residual variances for the three models.

1. $\text{Var}(Y_2) = \sigma^2$
2. $\text{Var}(Y_2 - Y_1) = 2\sigma^2(1 - \rho)$
3. $\text{Var}(Y_2 | Y_1) = \sigma^2(1 - \rho^2)$

ANCOVA has the smallest residual variance regardless of ρ .

* The assumption that the variance do not depend on either group or time could be dropped.



ANCOVA with multiple times of follow-up

- ▶ Only include post baseline outcomes.
- ▶ Include baseline as a covariate in the data and include the baseline*time interaction in the model.
- ▶ Cannot use week=0 as reference, switch to last time.

Note: Different effects of baseline at different time points due to stronger correlation between baseline and early follow up.

Example:

```
PROC MIXED DATA=ckd; WHERE week > 0;
CLASS id week treat (ref='0');
MODEL aix = week treat treat*week aix0*week
/SOLUTION CL DDFM=KR;
REPEATED week / SUBJECT=id TYPE=UN R RCORR;
RUN;
```



Eplerenone: ANCOVA (FIX ME)

| Effect | week | treat | Estimate | Standard | | Pr > t | Alpha | Lower | Upper | |
|------------|--------|-------|----------|----------|------|---------|--------|-------|----------|----------|
| | | | | Error | DF | | | | | |
| Intercept | | | 9.1333 | 2.9647 | 43.1 | 3.08 | 0.0036 | 0.05 | 3.1550 | 15.1115 |
| week | 12 | | -5.5830 | 3.1340 | 43.3 | -1.78 | 0.0819 | 0.05 | -11.9020 | 0.7360 |
| week | 24 | | 0 | | | | | | | |
| treat | 1 | | -4.2329 | 2.0775 | 43.7 | -2.04 | 0.0477 | 0.05 | -8.4207 | -0.04515 |
| treat | 0 | | 0 | | | | | | | |
| week*treat | 12 | 1 | 1.8835 | 2.1936 | 44 | 0.86 | 0.3952 | 0.05 | -2.5375 | 6.3046 |
| week*treat | 12 | 0 | 0 | | | | | | | |
| week*treat | 24 | 1 | 0 | | | | | | | |
| week*treat | 24 | 0 | 0 | | | | | | | |
| aix0*week | 12 | | 0.9075 | 0.1320 | 44 | 6.88 | <.0001 | 0.05 | 0.6415 | 1.1735 |
| aix0*week | 24 | | 0.7570 | 0.1058 | 43.2 | 7.15 | <.0001 | | | |
| 0.05 | 0.5436 | | 0.9704 | | | | | | | |

Conclusion: Significant difference at last follow-up. The estimated difference for two persons with the same baseline-value is -4.23% (95% CI: -8.42% to -0.05%, P=0.0477) in favor of Eplerenone.



Eplerenone: ANCOVA

Note: Covariance estimates change substantially when baseline is included as a covariate (it explains a lot of variation in the data).

The Mixed Procedure

Estimated R Matrix for id 1

| Row | Col1 | Col2 |
|-----|---------|---------|
| 1 | 74.1578 | 34.2828 |
| 2 | 34.2828 | 47.2484 |

Estimated R Correlation Matrix for id 1

| Row | Col1 | Col2 |
|-----|--------|--------|
| 1 | 1.0000 | 0.5792 |
| 2 | 0.5792 | 1.0000 |



ANCOVA vs cLMM

Models have **different interpretations**.

- ▶ cLMM describes the population response.
- ▶ ANCOVA describes the response for subjects **having the same baseline response**.

ANCOVA and cLMMs estimate the **same treatment effect*** with similar accuracy/power

- ▶ P-values and parameter estimates are very similar.
- ▶ But the two approaches are **not identical!**

* The feature that treatment effect is the same on the subject mean and the population mean is particular to linear models (lectures 5+6).



Baseline in observational studies

Compare the outcomes for individuals from different groups (e.g. gender or illness groups):

- ▶ The groups are likely to differ in many respects . . . **including the baseline outcome value!**
- ▶ Differences in response profiles may be due to many factors, and quantifications will depend on which of these are factors included in the model.
- ▶ Adjust for the covariates that are sensible in the context.

Is the baseline measurement a sensible covariate?

Baseline in observational studies

Fitzmaurice et al. (2011)[Section 5.6]:

For example, in an observational study examining gender differences in weight gain of infants between 12 months (baseline) and 24 months (...) At baseline boys are on average 1 1/2 pounds heavier than girls, but there is no evidence of a gender effect on the 12 month change in body weight, with boys and girls both gaining approximately 5 1/4 pound. In contrast the analysis of covariance of the same data reveals a discernible gender effect with boys showing more weight gain than girls.

(...) the analysis of covariance is directed at the conditional question of whether boys are expected to gain more weight than girls given that they have the same initial weight at 12 months. (...) The reasoning is that if a boy and girl have the same initial weight at 12 months, then there are two possibilities: (1) the girl is initially overweight and is expected to gain less weight or (2) the boy is initially underweight and is expected to gain more weight over the 12 months. **We advise readers to employ the analysis of covariance approach in longitudinal settings only if the approach and its implications are fully understood.**



Suggested reading: lecture 1

- ▶ Basics of longitudinal data: FLW(2011) chapters 1-2.
- ▶ Linear models for longitudinal data: FLW chapters 3-4.
ATT: This is technical!
- ▶ Analysis of response profiles: FLW chapter 5.

