

Analyse av repeterte målinger

(mixed models, generalized estimating equation, paired t-test ...)

Harald Weedon-Fekjær

Oslo Center for Biostatistics and Epidemiology

Oslo Centre for Biostatistics and Epidemiology [OCBE]

<http://www.med.uio.no/imb/english/research/centres/ocbe/>



Universitet i Oslo,
Det medisinske fakultet



Arnoldo Frigessi,
leder OCBE

- Undervisning i biostatistikk
- ***Veiledning i biostatistikk***
- Biostatistikk forskning

Planla å reise til Sørlandet Sykehus

- Hadde gjort avtaler om veiledning
- Planlagt overnattingen
- Planlagt middagen

...

...

...

Og så kom det en pandemi 😞

... men BioNTech sin løsning er på vei



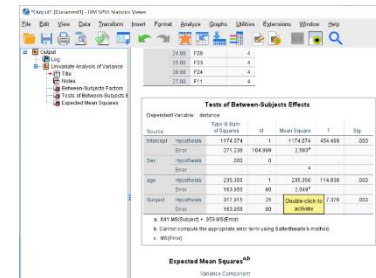
Jeg har ikke gitt opp, jeg kommer igjen;
Middagen er ikke spist og skogen står der stadig 😊



How do we analyze repeated measurements?

ANOVA ? (analysis of variance)

- Old classical approach
- Easy to calculate by hand
- ... but:
 - > Gives p-values, not estimated (clinical relevant) properties with confidence interval
 - > No unbalanced missing data
 - > Not very flexible



Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Total	1174.274	1	1174.274	454.459	.000
Model	271.238	104.889	2.587		
Error	250	0			
Total					
Age	235.355	1	235.355	114.838	.000
Sex	163.885	1	163.885	79.947	.000
Height	377.813	25	15.112	7.319	.000
Error	163.885	89	1.831		
Total					

a. R Squared = .305 (Adjusted R Squared = .300)

b. Contrast computes the appropriate error term using Scheffé's method.

c. Model(s)

Expected Mean Squares^{a,b}

Variance Component



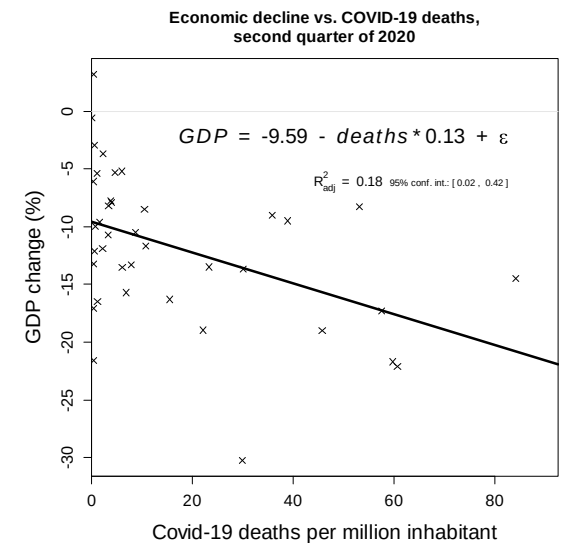
- ***An endangered method for repeated measurements?***

How do we analyze repeated measurements?

Mixed models regression



- The modern flexible approach to repeated measurements
- *Works in Stata, R and SPSS*
- *An extension of ordinary regression analysis*
- *Estimates effects with confidence intervals*



Random effects (Modelling individual effects)

Classic basic statistical model: (linear regression etc.)

$$Y = \text{“covariate effects”} + \text{“random noise”}$$

A mixed effects model with random effects:

$$Y = \text{“covariate effects”} + \text{“individual variation”} + \text{“random noise”}$$

=> We take into account individual variation by random effects
(random effects might be used for all sorts of cluster data)

Random effects (Modelling individual effects)

Classic basic statistical model: (linear regression etc.)

$$Y_i = \alpha + \beta * X + \varepsilon_i$$

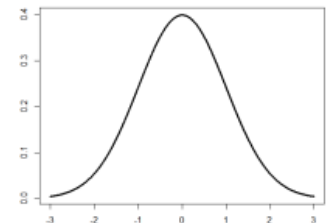
A mixed effects model with random effects:

$$Y_{i,j} = \alpha + \beta * X + \mathbf{u}_{j.} + \varepsilon_{i,j}$$

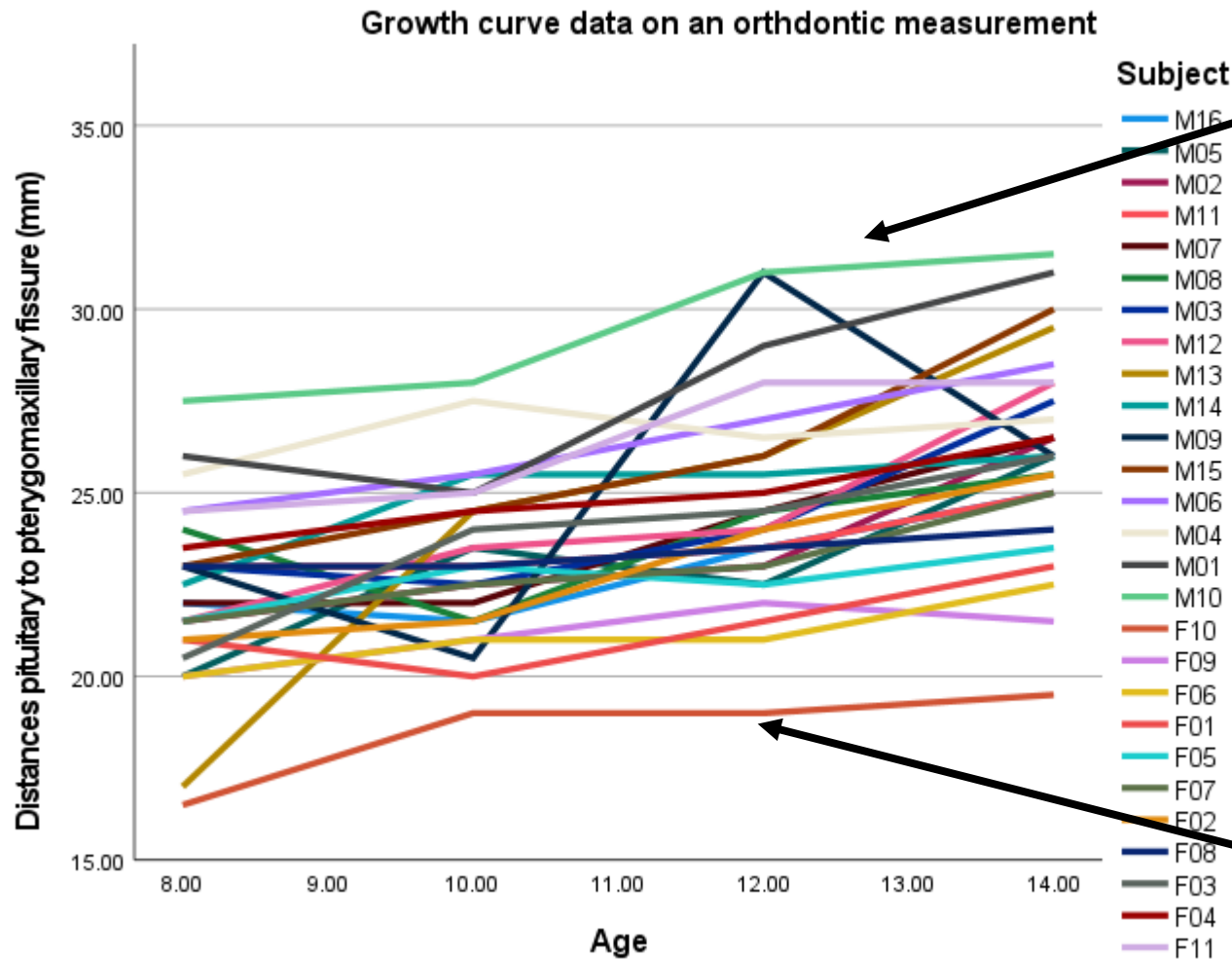
Assumptions;

The individual level comes from
a normal distribution,

(like the normal distributed random noise)



Random (individual) effects



M10 have general **long** distance between pituitary and pterygomaxillary fissure

F04 have general **short** distance between pituitary and pterygomaxillary fissure

Mixed effects models

- > Mixed effects models might be complex with:
 - > Many different levels/clusters
 - > Advanced covariance structures
 - > Non-linear expected (mean) values
- > Individual with more observations influence the estimates more, but new individuals is usually more important than one more repeated observation
- Mixed effects models could be used in both:
 - > linear regression
 - > logistic regression
 - > poisson regression
 - > Survival analysis (=“frailty models”)

Slide excluded due to non published data

Slide excluded due to non published data

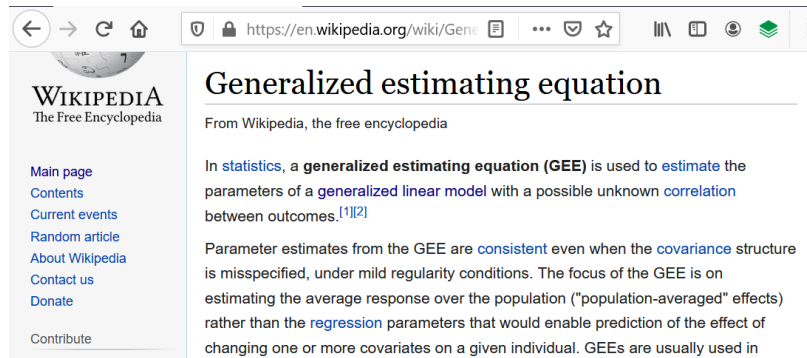
Do we always need mixed effects models?

- When we compare first and last observation;
Why not just work with difference last-first?
(paired samples t-test & confidence intervals)
- Is one option to first take the mean of each group?
(typically only for balanced designs without missing)

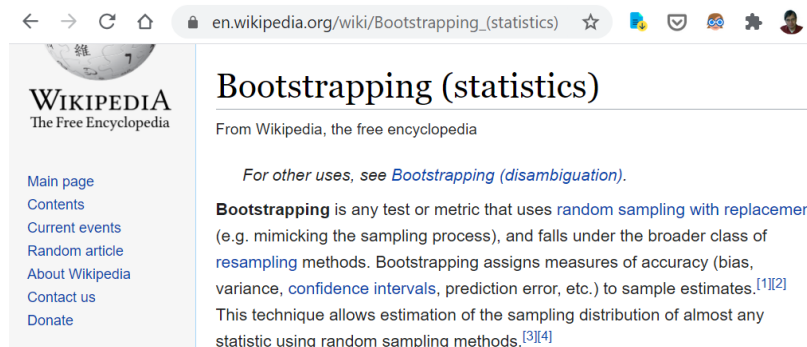
Non normal (continues variable) distributed data?

- **Generalized estimating equations (gee)**

is an option for non normal distributed data with decent number of clusters/individuals



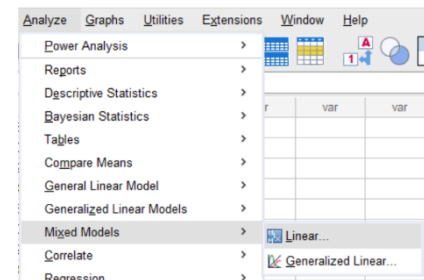
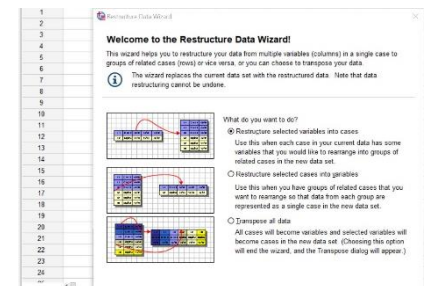
- **Bootstrap** is a good robust method, but be careful with the standard SPSS implementation (not recommended!)



Data format and short demo

- Mixed effects models take data on “long format”;
One observation per line with shared id.
- Stata/SPSS/R have routines to convert from “wide” to “long” format
- Stata/SPSS/R all have comprehensive mixed models estimation routines

distance	age	Subject	Sex
22.50	8.00	M14	Male
25.50	10.00	M14	Male
25.50	12.00	M14	Male
26.00	14.00	M14	Male
23.00	8.00	M15	Male
24.50	10.00	M15	Male
26.00	12.00	M15	Male
30.00	14.00	M15	Male
22.00	8.00	M16	Male
21.50	10.00	M16	Male
23.50	12.00	M16	Male
25.00	14.00	M16	Male
21.00	8.00	F01	Female
20.00	10.00	F01	Female
21.50	12.00	F01	Female
23.00	14.00	F01	Female



Summary:

1. (drop ANOVA analysis)
2. **First look for simplifications** (marginal means etc.)
3. **Use mixed effects models with random components for complex data taking into account individual variations and covariences**

Mixed effects models is just an extension of ordinary generalized linear models with individual variation (but can become very complex)

Spørsmål?



Steilene, Nesodden, Norway, 2013-06-14

Harald Weedon-Fekjær (statistician, PhD), OCBE, Oslo University Hospital

E-mail: <harald.weedon-fekjar@medisin.uio.no>



Referanser:

- Mixed models:

https://en.wikipedia.org/wiki/Mixed_model

<https://stats.idre.ucla.edu/other/mult-pkg/introduction-to-linear-mixed-models/>

- SPSS:

<https://www.ibm.com/analytics/spss-statistics-software>

<https://stats.idre.ucla.edu/spss/seminars/spss-mixed-command/>

- Stata:

<https://www.stata.com/>

<https://www.stata.com/features/multilevel-mixed-effects-models/>

- R:

<https://www.r-project.org/>

<https://cran.r-project.org/web/packages/lme4/vignettes/lmer.pdf>