13.01.2021

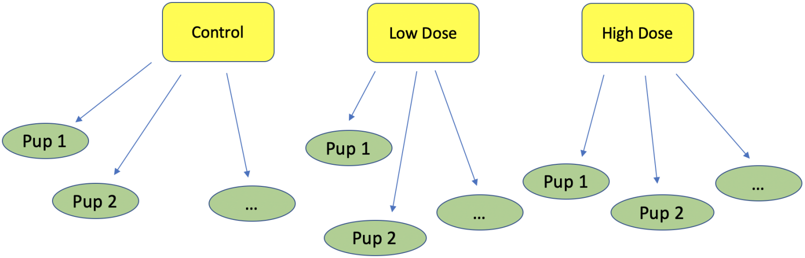
Advanced Topics in Biostatistics 2020/2021
Linear Mixed Models

Dr. Nicholas Schreck
Biostatistics - C060
nicholas.schreck@dkfz-heidelberg.de

Outline

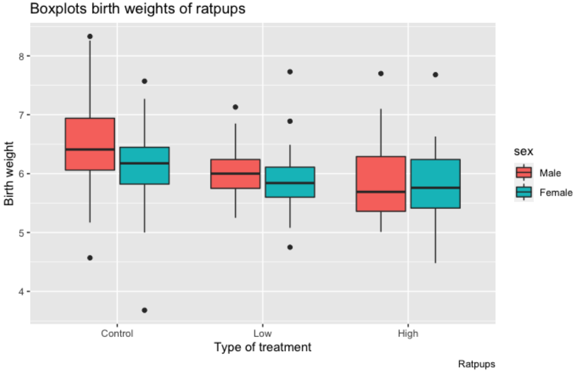
Outline

Ratpup example



1. Ratpups associated with treatment (3 levels: Control, Low Dose, High Dose)
2. Observation of birth weights of ratpups
3. Research question: Effect of treatment on birth weight of pups?

Data visualization I



1. Birth weight seems to decrease with dose
2. Differences between sexes within treatments
3. Variances within treatments equal, but not across treatments

ANOVA

See lecture “Analysis of Variance” by Silvia Calderazzo:

- Compare the means of k groups (here: $k = 3$):

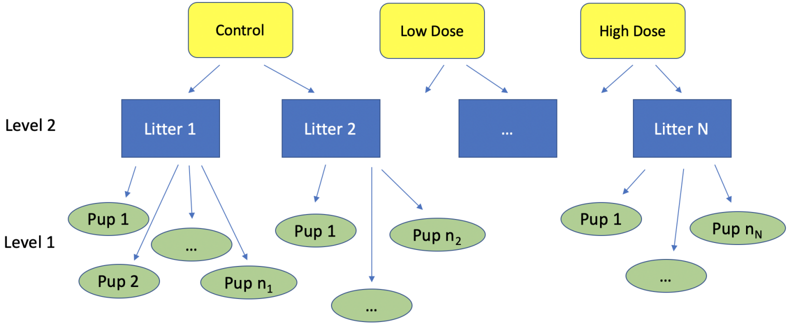
$$H_0 : \mu_{\text{control}} = \mu_{\text{low}} = \mu_{\text{high}}$$

- F-test based on sum of squares

Assumptions:

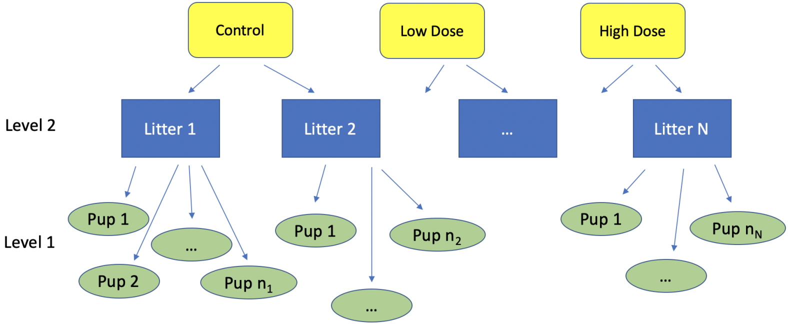
1. Measurements normally distributed
2. Measurements have equal variance
3. Measurements are independent

Ratpup example: Clustered data



1. Observations on subjects within same randomly selected litter
2. Units of the analysis nested within cluster (litter)

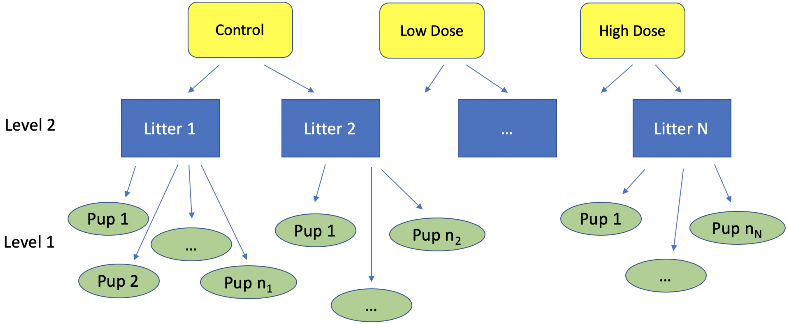
Ratpup example: Clustered data



1. Research question: Birth weight of pups different in litters?
2. Difference: Measurements of birth weight not independent

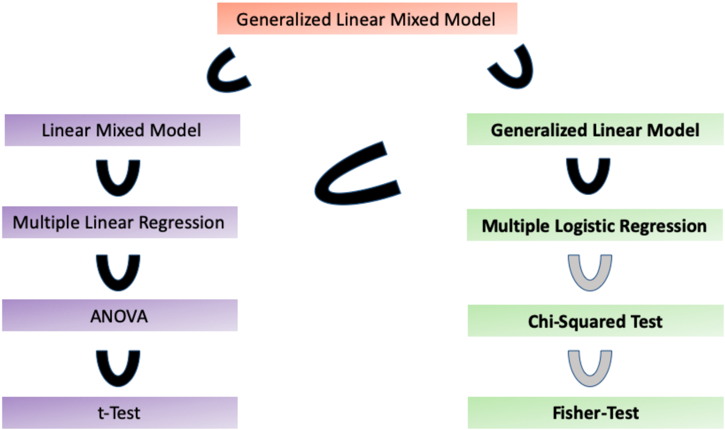
Do **NOT** apply ANOVA (t-test) to raw data!

Ratpup example: Clustered data

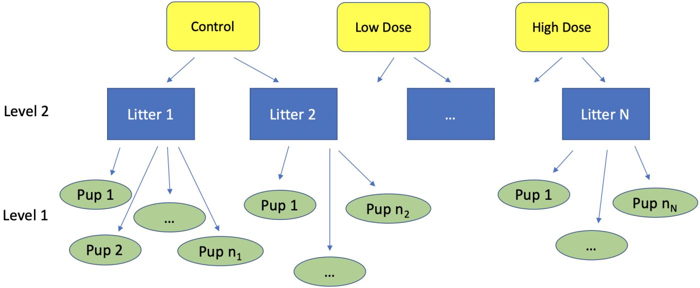


1. Research question: Birth weight of pups different in litters?
2. Difference: Measurements of birth weight not independent
3. Statistically correct analysis ("gold standard"): Linear mixed model

Placement in Context of Lecture Series



Why and when mixed effect models?



1. Modelling of (complex) hierarchical structures
2. Correlated observations (nested or crossed designs)

Why should we not use "standard" methods?

Outline

Relevance of hierarchical datasets

"To generalize conclusions to a population, we must sample its variation" (Altman and Krzywinski, 2015)

1. Biological variation: To be maintained / sampled
   - Parallel measurements of biologically distinct items
   - Capture random biological variation of population of interest
   - External validity
2. Technical variation: To be controlled
   - Repeated measurements of the same items
   - Represent independent measures of random noise associated with protocols or equipment (obtain independent measurements of random noise)
   - Internal validity

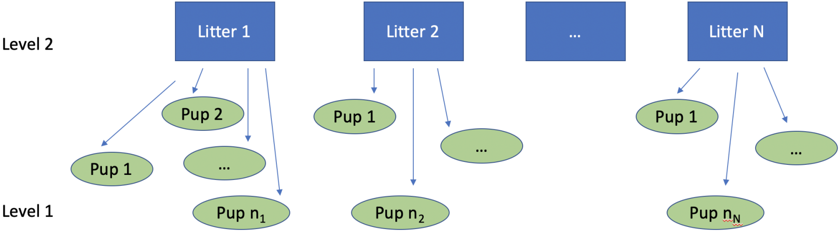
Types of hierarchical data (West et al., 2015)

Hierarchical noise sources (e. g. biological and technical variation) in nested or crossed designs (Kryzwinski et al., 2014)

1. Clustered data (blocks)
2. Repeated measurements (under different conditions)
3. Longitudinal studies (over time)

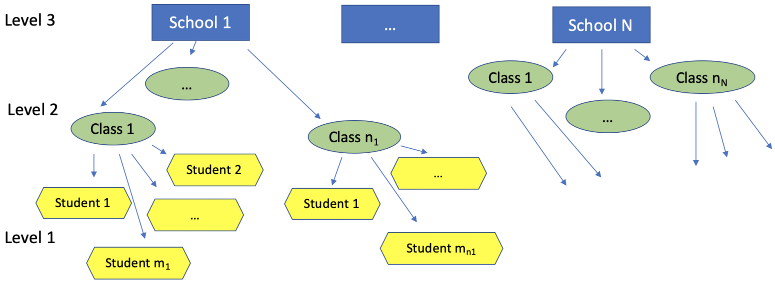
1. Nested factors: particular level of a factor measured within single level of another factor (e. g. ratpups in litters)
2. Crossed factors: particular level of a factor measured across multiple levels of another factor (e. g. treatments across sex in ratpup example)

Clustered data: Ratpup example



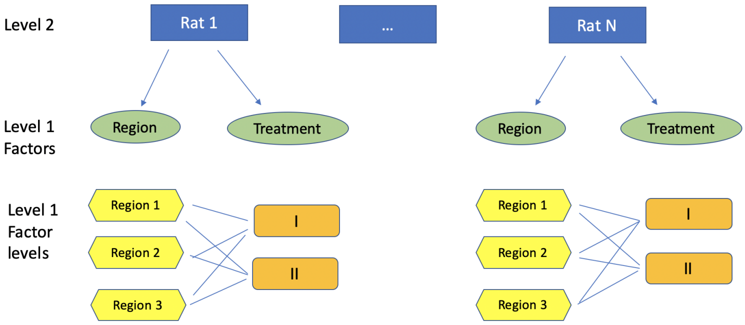
1. Observations on subjects within same randomly selected group
2. Units of the analysis nested within cluster
3. Clusters are sampled from larger population –> random effects

Clustered data: Classroom example



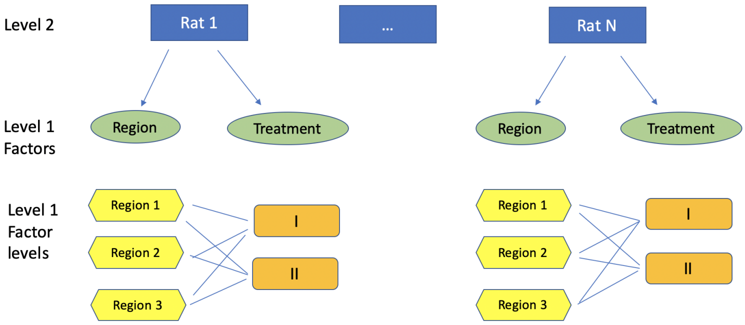
1. Observations on subjects within same randomly selected group
2. Units of the analysis nested within cluster
3. Clusters are sampled from larger population –> random effects

Repeated-measures data: Ratbrain example



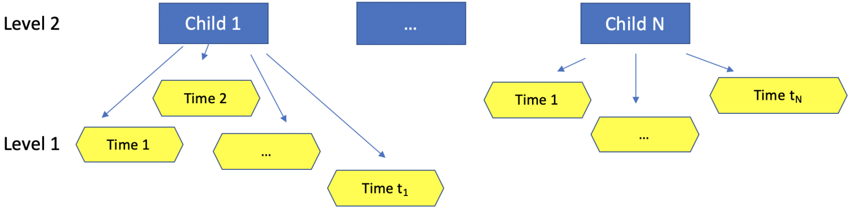
1. Measurements on same unit under changing conditions
2. Measurements for the same subject correlated

Repeated-measures data: Ratbrain example



1. Clustered data in crossed two-level design
2. Brain regions are crossed with treatments

Longitudinal studies: Example



1. Multiple measurements on same subject over time
2. Measurements for the same subject correlated
3. Dropout huge concern
4. Combination of clustered and longitudinal data possible

Outline

Fixed or Random Factor?

> Decision fixed / random depends on the aim of the study!

1. Fixed factor:
   - Levels / conditions constitute entire population of interest
   - Stay the same when repeating experiment
   - Typically used in ANOVA setting
   - Qualitative covariate (gender, treatments, . . .)
2. Random factor:
   - Levels, not observations, randomly sampled from population
   - Not all possible levels of the factor present in data
   - Levels change when repeating experiment
   - Levels itself not of intrinsic interest (but variability of factor, noise source)
   - Mitigate noise to detect effects
   - Typically: classification variables of higher levels in hierarchical data (litters, schools, . . .)
   - Random intercept (cluster-specific deviation from overall mean)

Fixed or Random Effect?

Decision fixed / random depends on the aim of the study!

1. Fixed effect:
    - Unknown, fixed quantitative population parameter (e. g. height, . . .)
    - Typically: Regression coefficients
2. Random effect:
    - Random deviations from relationships defined by fixed effects
    - Quantitative variable specific to random factor (e. g. age in longitudinal studies, . . .)
    - Typically: Random slope (deviations from population-wide regression coefficients)

Illustration: Fixed effects

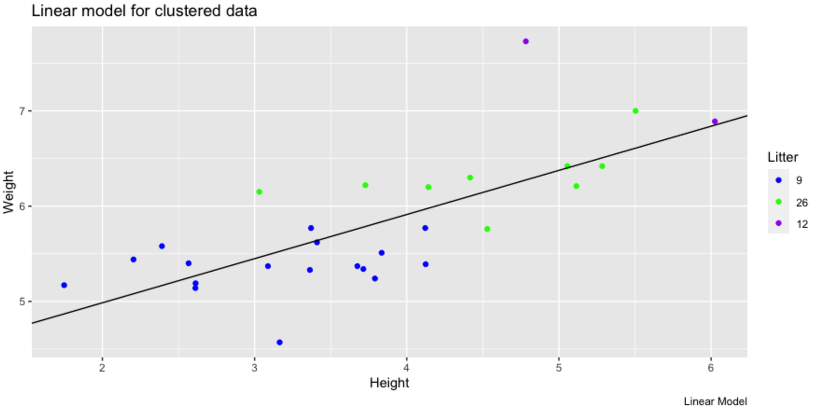


Illustration: Random intercept

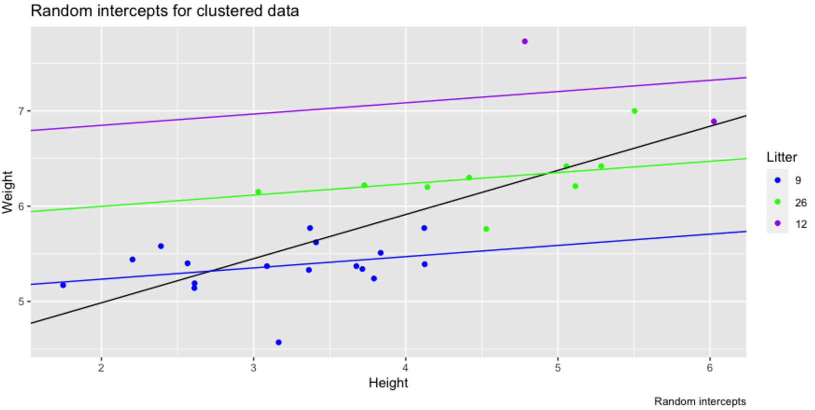
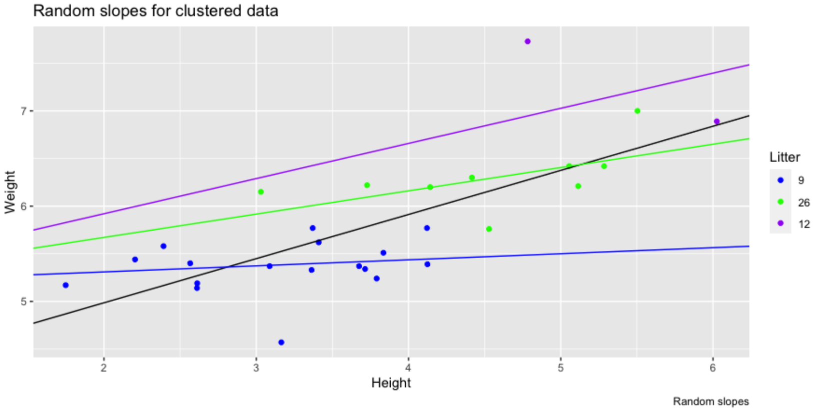
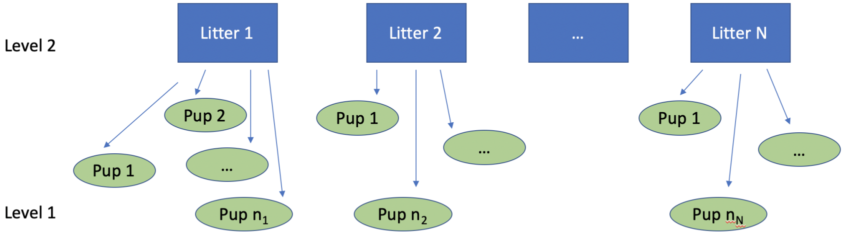


Illustration: Random slope



Outline

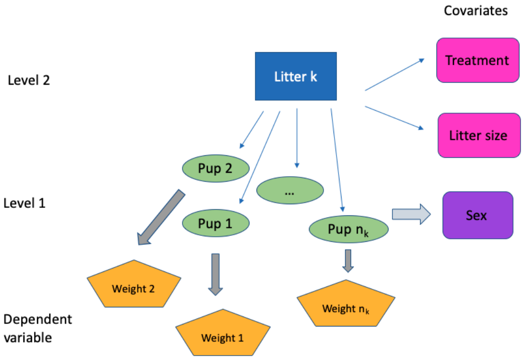
Ratpup example (clustered two-level)



1. Nested clustered randomized trial
2. 27 female rats got one of 3 doses of treatment (high, low, control)
3. $n = 322$ total observations of birth weight

> Question: Choice of treatment influence on birth weight?

Measurements in ratpup example



1. Each subject (pup) measured only once
2. Level of covariates important
3. Individual cluster-specific (litter-specific) effects

What is a LMM?

$$y_{ij} = X_{ij}^{(1)}\beta_1 + X_{ij}^{(2)}\beta_2 + \ldots + X_{ij}^{(p)}\beta_p + Z_{ij}^{(1)}u_{1j} + \ldots + Z_{ij}^{(q)}u_{qj} + \varepsilon_{ij}$$

1. Two-level data: individual subject $i = 1, \ldots n_j$ in group $j = 1, \ldots, N$
2. p (population)-wide fixed effects, q (subject-specific) random effects
3. X, Z can be continuous or indicators

Parametric statistical model

- for continuous responses
- with normally distributed residuals
- that may not be independent
- and / or may not have constant variance
- that is linear in the parameter effects
- which may involve a mix of fixed and random effects.

What is a LMM?

$$y_{ij} = X_{ij}^{(1)}\beta_1 + X_{ij}^{(2)}\beta_2 + \ldots + X_{ij}^{(p)}\beta_p + Z_{ij}^{(1)}u_{1j} + \ldots + Z_{ij}^{(q)}u_{qj} + \varepsilon_{ij}$$

$$\text{weight}_{ij} = \beta_0 + \text{Treatment-low}_j \times \beta_1 + \text{Treatment-high}_j \times \beta_2 + u_j + \varepsilon_{ij}$$

1. Ratpup $i = 1, \ldots n_j$ in litter $j = 1, \ldots, 27$
2. $p = 3$, $q = 1$
3. $u_j \sim \mathcal{N}(0, \sigma_u^2)$: random intercept for litter j
4. $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$: independent errors (for all weight measurements)

Some comments (Advanced)

Fixed effects:

1. Test statistic follows a F statistic only approximately
2. Degrees of freedom have to be estimated (e. g. Satterthwaite's method)

Variance components:

1. Estimated using maximum-likelihood (ML) or restricted maximum-likelihood (REML)
2. REML often preferred (less biased)
3. Model comparisons (different fixed effects): ML estimation

Model fit in R

```r
### Linear mixed model
library(lme4)
library(lmerTest)
data("RatPupWeight")
attach(RatPupWeight)
ratpup <- RatPupWeight
ratpup$treat <- relevel(factor(ratpup$Treatment, ordered = F), ref = "Control")
rm(RatPupWeight)

linmixmod <- lmer(weight ~ treat + (1 | Litter), data = ratpup, REML = T)
summary(linmixmod)
anova(linmixmod)
```

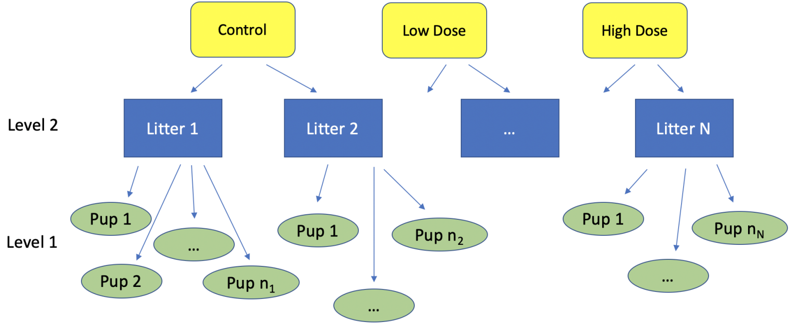
R Output

```
Random effects:
 Groups   Name        Variance Std.Dev.
 Litter   (Intercept) 0.2770   0.5263
 Residual             0.1966   0.4433
Number of obs: 322, groups:  Litter, 27

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)
(Intercept)   6.4533     0.1716 21.1614  37.598   <2e-16 ***
treatLow     -0.4287     0.2435 21.3452  -1.761   0.0926 .
treatHigh    -0.3944     0.2696 21.8031  -1.463   0.1577
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> anova(linmixmod)
Type III Analysis of Variance Table with Satterthwaite's method
       Sum Sq Mean Sq NumDF  DenDF F value Pr(>F)
treat 0.72091 0.36045     2 21.678  1.8339 0.1837
```

Fill-in solution



Fill-in: Average over level 1 measurements and compare means with ANOVA (t-test) (Holland-Letz and Kopp-Schneider, 2020)

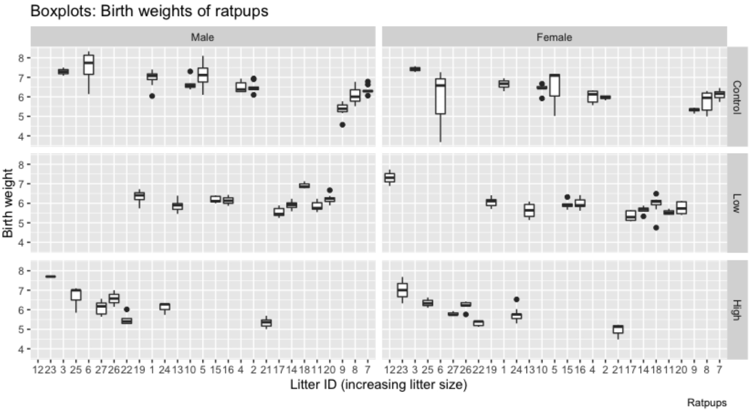
1. If number of level 1 observations (approximately) equal in groups and
2. If no other (confounding) covariates to consider and
3. If no interest in the specific sources of randomness

Comparison of approaches

1. ANOVA to raw data (irrespective of hierarchy): $p < 0.0001$ Wrong, underestimation of variance
2. Mean method: $p = 0.2418$ Valid, but lower power
3. Linear mixed model (F-test, REML): $p = 0.1837$ Valid, optimal approach

ANOVA approach leads to over-optimistic p-values!

Data visualization II



- Birth weight decreases with litter size for all combinations of sex and treatment

Building the model

$$y_{ij} = X_{ij}^{(1)}\beta_1 + X_{ij}^{(2)}\beta_2 + \ldots + X_{ij}^{(p)}\beta_p + Z_{ij}^{(1)}u_{1j} + \ldots + Z_{ij}^{(q)}u_{qj} + \varepsilon_{ij}$$

$$\text{weight}_{ij} = \beta_0 + \text{Treatment-low}_j \times \beta_1 + \text{Treatment-high}_j \times \beta_2 + \\ + \text{Sex-female}_{ij} \times \beta_3 + \text{Litter-size}_j \times \beta_4 + u_j + \varepsilon_{ij}$$

1. Ratpup $i = 1, \ldots n_j$ in litter $j = 1, \ldots, 27$
2. $p = 3,\ q = 1$
3. $u_j \sim \mathcal{N}(0, \sigma_u^2)$: random intercept for litter j
4. $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$: independent errors (for all weight measurements)
5. Add level 1 (sex) and level 2 (litter size) covariables
6. Interaction effect possible (as in every linear model)

Model fit in R II

```
library(lme4)
library(lmerTest)
data("RatPupWeight")
attach(RatPupWeight)
ratpup <- RatPupWeight
ratpup$treat <- relevel(factor(ratpup$Treatment, ordered = F), ref = "Control")
rm(RatPupWeight)

model_full <- lmer( weight ~ treat + sex + Lsize + (1 | Litter),
                    data = ratpup, REML=T)
summary(model_full)
anova(model_full)
```

R Output II

```
Random effects:
 Groups   Name        Variance Std.Dev.
 Litter   (Intercept) 0.0974   0.3121
 Residual             0.1628   0.4035
Number of obs: 322, groups:  Litter, 27

Fixed effects:
             Estimate Std. Error        df t value Pr(>|t|)
(Intercept)   8.30987    0.27371  32.61126  30.360  < 2e-16 ***
treatLow     -0.42850    0.15040  22.90425  -2.849   0.0091 **
treatHigh    -0.85870    0.18181  24.97854  -4.723 7.65e-05 ***
sexFemale    -0.35908    0.04749 301.82484  -7.562 4.81e-13 ***
Lsize        -0.12900    0.01879  31.67409  -6.864 9.63e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> anova(model_full)
Type III Analysis of Variance Table with Satterthwaite's method
      Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
treat 3.7752  1.8876     2  24.240  11.595  0.000293 ***
sex   9.3093  9.3093     1 301.825  57.182 4.814e-13 ***
Lsize 7.6708  7.6708     1  31.674  47.117 9.630e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Outline

Take-home messages

Linear mixed models

1. are used to model dependent, normally distributed responses
2. enable analysis for dependent observations (clustered design, repeated measurements, longitudinal data)
3. enable (approximately) correct inference on fixed effects

1. are computationally intensive (specialized software necessary)
2. are more complicated to use and interpret
3. have to be used carefully when testing fixed effects
4. have to be used extremely carefully when testing random effects

Linear Mixed Model vs. Mixed Model ANOVA

LMM

- Likelihood-based
- Asymptotically well-behaved (as number of levels of random effects become large)
- provides basis for optimal results (appropriate inference possible)
- Competent modelling necessary

ANOVA

- Based on fixed effect models (ordinary least squares)
- Lack of model-based principles to form guidelines in applications with mixed effects (Littell, 2002)
- Difficulties occur especially for unbalanced data
- Easy to apply

Software Notes

1. Mixed Models computationally intensive
2. Software includes R, SAS, SPSS, Stata, HLM
3. R packages
   1. "nmle" (Pinheiro et al., 2020)
   2. "lme4" (Bates et al., 2015)
   3. "lmerTest" (Kuznetsova et al., 2017)
4. Usage of the R packages depends on specific application (use nlme only for complex covariance structures)
5. Graphpad Prism: Repeated-measures ANOVA, no LMM
6. Examples and explanations: West et al. (2015) https://www.academia.edu/37093545/LINEAR_MIXED_MODELS_A_Practical_Guide_Using_Statistical_Software

References

Naomi Altman and Martin Krzywinski. Sources of variation. *Nature Methods*, 12(1):5–6, 2015.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015.

Tim Holland-Letz and Annette Kopp-Schneider. Drawing statistical conclusions from experiments with multiple quantitative measurements per subject. *Radiotherapy and Oncology*, 152:30–33, 2020.

Martin Kryzwinski, Naomi Altman, and Paul Blainey. Nested designs. *Nature Methods*, 11(10): 977–978, 2014.

Alexandra Kuznetsova, Per B. Brockhoff, and Rune H. B. Christensen. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13):1–26, 2017.

Ramon C. Littell. Analysis of unbalanced mixed model data: A case study comparison of ANOVA versus REML/GLS. *Journal of Agricultural, Biological, and Environmental Statistics*, 7: 472–490, 2002.

Jose Pinheiro, Douglas Bates, Saikat DebRoy, Deepayan Sarkar, and R Core Team. *nlme: Linear and Nonlinear Mixed Effects Models*, 2020. URL https://CRAN.R-project.org/package=nlme. R package version 3.1-150.

Brady West, Kathleen Welch, and Andrzej Galecki. *Linear Mixed Models - A Practical Guide Using Statistical Software*. CRC Press, 2 edition, 2015.

Next lecture

Topic: Introduction to Bayesian thinking
Lecturer: Silvia Calderazzo
Date: 20 January