Mixed Linear Effects Models

Peter von Rohr

2023-04-24

Recap 2023-04-24:

So far: Fixed linear effect models (FLEM)
* FLEM are called fixed, because apart from the random residual effect, they only contain, so-called fixed effects
* In statistical modelling (Frequentist way), there are two type of effects

1. fixed effects
2. random effects

Random effects
* Mostly interested in variation

$$cov(e_i, e_j) = 0$$

$$E(e_i) = 0$$

$$var(e_i) = \sigma_e^2$$

Concept of infinite sampling

Random residuals: Properties:
Follow a certain distribution with a given expected value and a given variance

$$y_i = x_i^T * b + e_i$$

$$b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

Intercept

Breast Circumference

fixed effects, that means in the model, we are not assuming that they show any variation.

Predictor Breast circumference

Body weight of animal i
as response

fixed effects:
* Mostly interested in the effect size,

Side Note (not relevant for exam)

Bayesian Way of Statistical Modelling
* no separation between fixed and random effects
* separation of variables into "known" and "unknown" quantitties, e.g. Body Weight and Breast Circumference, they are all known, and intercept and regression slope are unknown
* all variables are assumed to follow a certain distribution...
* Estimation principle for unkown quantities is based on the posterior density of the unknowns given the knowns

Repeated Measurements

- Same characteristics (`Body Weight`, `BC`, . . .) measured multiple times for the same animal

Nr	Animal	Breast Circumference	Body Weight
1	2	177.0000	463.0000
2	2	177.3129	468.8940
	2	177.3292	467.8753
	5	179.0000	496.0000
	5	178.6501	495.0033
	5	178.7485	493.6563
	7	181.0000	518.0000
8	7	180.9819	509.3221

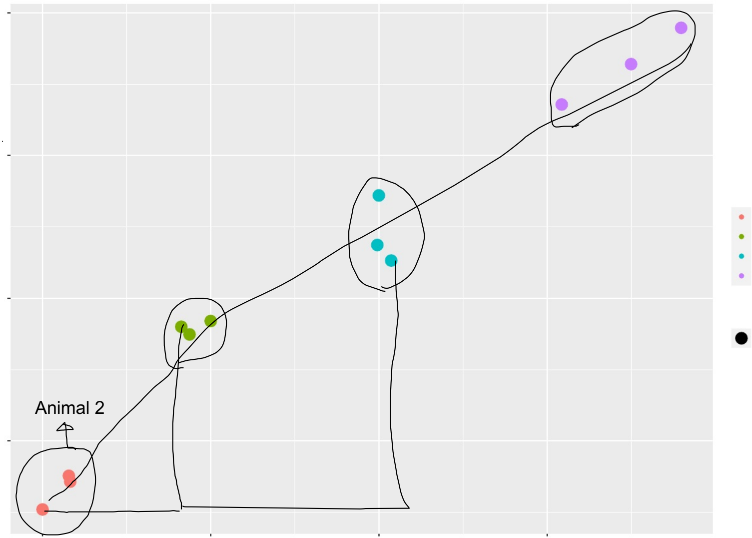
- Column `Animal` no longer just a counter, it becomes a model factor

Properties

- What is the benefit of repeated measurements?
- Equivalent to more animals in the dataset?
- Assumption: repeated observations of the same animal are more "similar"
- No longer independence of observations

Repeated observations for the same animal are
more alike than observations for different animals

Data Scatterplot



Statistical Analysis

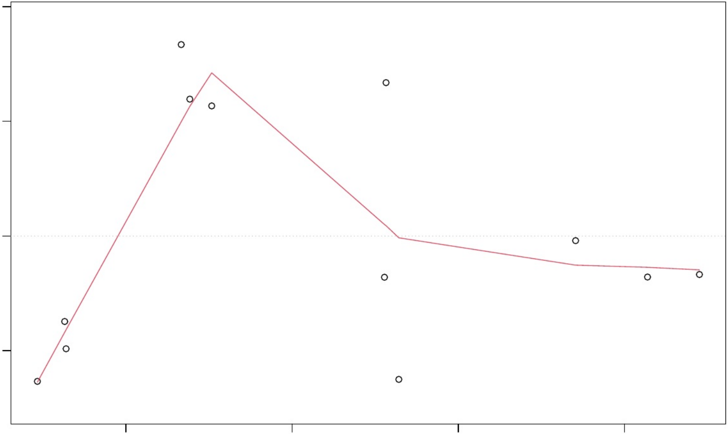
- Technically, same regression analysis is possible
- But assumption of independence is violoted
- Consequence of independence

$$var(\mathbf{e}) = \mathbf{I} * \sigma_e^2$$

* Check residuals plot

Diagnostics Plot

```
plot(lm_rep_obs)
```



ANOVA

Analysis of Variance

* rather "old" or "established" type of statistical analysis
* focus on variability, more specifically which model components can explain what percentage of variability of the data.

aov() function in R that computes ANOVA

```
aov_bw_breed <- aov(`Body Weight` ~ Breed,
                                  data = tbl_rep_obs_breed)
summary(aov_bw_breed)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## Breed        2   7356    3678   22.29 0.000327 ***
## Residuals    9   1485     165
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.
```

ANOVA I: Breed and Animal

```r
aov_bw_breed_animal <- aov(`Body Weight` ~ Breed + Animal,
                           data = tbl_rep_obs_breed)
summary(aov_bw_breed_animal)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## Breed        2   7356    3678  160.06 3.53e-07 ***
## Animal       1   1302    1302   56.64 6.76e-05 ***
## Residuals    8    184      23
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.
```

ANOVA II

```r
aov_bw_no_breed_rep <- aov(`Body Weight` ~ Error(Animal),
                          data = tbl_rep_obs_no_breed)
summary(aov_bw_no_breed_rep)
```

```
##
## Error: Animal
##           Df Sum Sq Mean Sq F value Pr(>F)
## Residuals  3   8658    2886
##
## Error: Within
##           Df Sum Sq Mean Sq F value Pr(>F)
## Residuals  8  183.8   22.98
```

ANOVA III

```
aov_bw_breed_rep <- aov(`Body Weight` ~ Breed +
                                 Error(Animal),
                               data = tbl_rep_obs_breed)
summary(aov_bw_breed_rep)
```

```
##
## Error: Animal
##           Df Sum Sq Mean Sq F value Pr(>F)
## Breed      2   7356    3678   2.826  0.388
## Residuals  1   1302    1302
##
## Error: Within
##           Df Sum Sq Mean Sq F value Pr(>F)
## Residuals  8  183.8   22.98
```

lme4

```r
lme_bw_no_breed_rep <- lme4::lmer(`Body Weight` ~ (1|Anima
                                  data = tbl_rep_obs_no_bre
summary(lme_bw_no_breed_rep)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: ‘Body Weight‘ ~ (1 | Animal)
##    Data: tbl_rep_obs_no_breed
##
## REML criterion at convergence: 82.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.36360 -0.50301  0.06086  0.26850  1.43838
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Animal   (Intercept) 954.34   30.892
##  Residual              22.98    4.794
```

With Breed

Animal	Body Weight	Breed
2	463.0000	Angus
2	468.8940	Angus
2	467.8753	Angus
5	496.0000	Simmental
5	495.0033	Simmental
5	493.6563	Simmental
7	518.0000	Limousin
7	509.3221	Limousin
7	506.5958	Limousin
10	541.0000	Limousin
10	547.3609	Limousin
10	533.9288	Limousin

ANOVA

```r
aov_bw_breed_rep <- aov(`Body Weight` ~ Breed +
                          Error(Animal),
                           data = tbl_rep_obs_breed)
summary(aov_bw_breed_rep)
```

```
##
## Error: Animal
##           Df Sum Sq Mean Sq F value Pr(>F)
## Breed      2   7356    3678   2.826  0.388
## Residuals  1   1302    1302
##
## Error: Within
##           Df Sum Sq Mean Sq F value Pr(>F)
## Residuals  8  183.8   22.98
```

lme4

```r
lme_bw_breed_rep <- lme4::lmer(`Body Weight` ~ Breed +
                                 (1|Animal),
                                  data = tbl_rep_obs_breed
summary(lme_bw_breed_rep)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: ‘Body Weight‘ ~ Breed + (1 | Animal)
##    Data: tbl_rep_obs_breed
##
## REML criterion at convergence: 61.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.3714 -0.5383  0.0640  0.3213  1.4305
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Animal   (Intercept) 426.21   20.645
```