ETH zürich

Mixed Effects Models: Wachstumskurven

Überblick

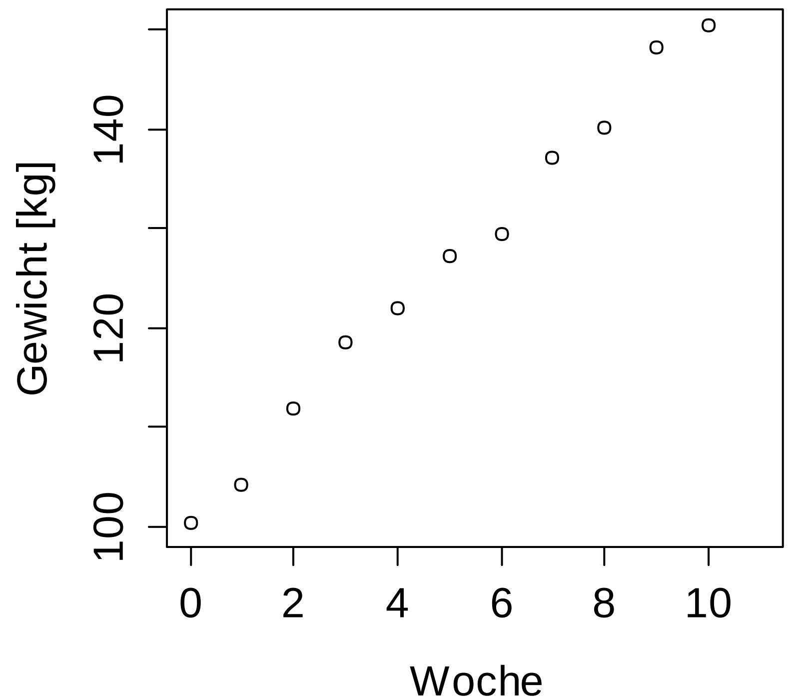
- Wiederholte Messungen (z.B. Wachstumskurven): Korrelierte Beobachtungen
- Random Intercept Model (RI)
- Random Intercept and Random Slope Model (RIRS)

Wdh: Lineare Regression

- Bsp: Kraftzuwachs durch Krafttraining
- Für eine einzelne Person:

$$y_j = \beta_0 + \beta_1 x_j + \varepsilon_j \; , \varepsilon_j \sim N(0, \sigma^2) \;\; i.i.d$$

"fixe" Effekte



ETH zürich

Viele Personen: Wiederholte Messungen

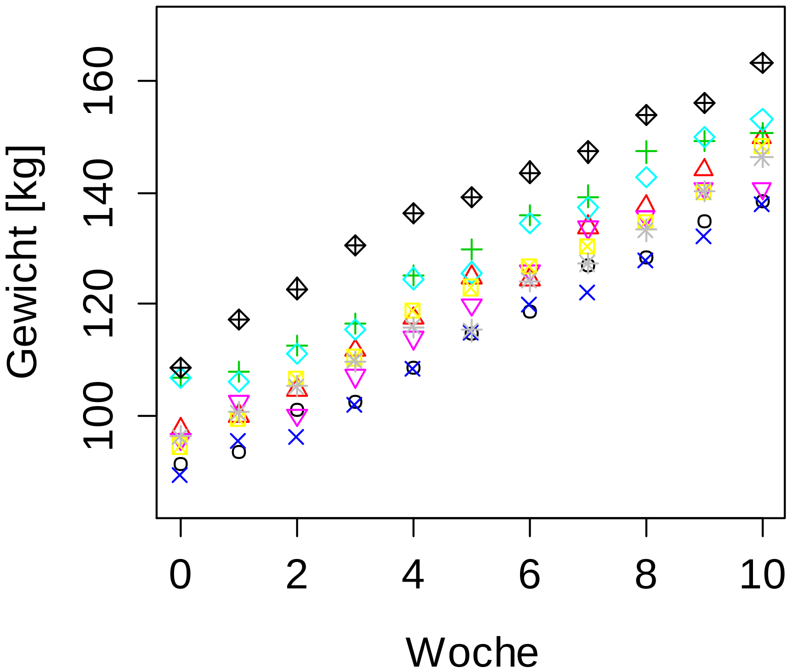
- Problem:
  Die Parameter (Achsenabschnitt und Steigung) jeder Person sind leicht unterschiedlich

- Wie beschreibt man diese Situation möglichst kompakt ?

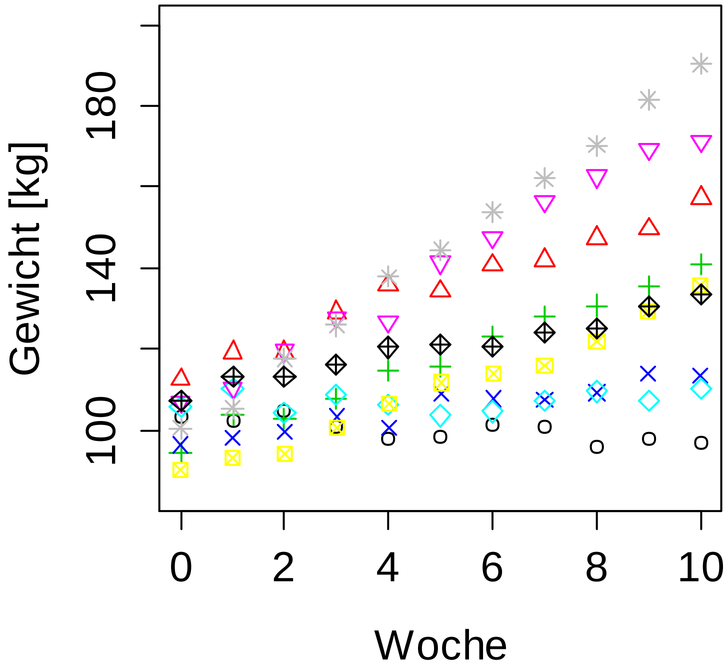
ETH zürich

Zurück zum Krafttraining

Jede Person hat eine unterschiedliche Kraft zu Beginn



Unterschiedliche Kraft zu Beginn
&
Spricht unterschiedlich auf Training an



ETH zürich

i: Person
j: Woche

Wiederholte Messungen 1/3: Block Effekte

- Möglichkeit 1: Block Effekte

$$y_{ij} = \left(\beta_0 + \beta_{0,i}\right) + \beta_1 x_j + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim N(0, \sigma^2)\ i.i.d$$

"fixe" Effekte

- Schätze: $\beta_0, \beta_{0,i}, \beta_1, \sigma$
- Erlaubt Aussagen über Individuen: Z.B. "Herr Meier hatte eine signifikant grössere Anfangskraft als Herr Müller"
- Erlaubt keine direkte Aussage über Population: Z.B. "Die typische Streuung der Anfangskraft in der Bevölkerung ist ca. 20 kg"

Wiederholte Messungen 2/3: Random Intercept (RI)

i: Person
j: Woche

- Möglichkeit 2: Mixed Effects Model

$$y_{ij} = (\beta_0 + u_i) + \beta_1 x_j + \varepsilon_{ij},$$
$$\varepsilon_{ij} \sim N(0, \sigma^2), u_i \sim N(0, \sigma_u^2) \quad i.i.d$$

"zufälliger" Effekt

"fixe" Effekte

Fixed + Random
=
Mixed

- Schätze: $\beta_0, \beta_1, \sigma, \sigma_u$
- Erlaubt keine direkten Aussagen über Individuen: Z.B. "Herr Meier hatte eine signifikant grössere Anfangskraft als Herr Müller"
- Erlaubt direkte Aussage über Population: Z.B. "Die typische Streuung der Anfangskraft in der Bevölkerung ist ca. 20 kg"

i: Person
j: Woche

Wiederholte Messungen 3/3: Random Slope and Random Intercept (RIRS)

“fixe” Effekte

“zufällige” Effekte

- Möglichkeit 2: Mixed Effects Model

$$y_{ij} = (\beta_0 + u_{1,i}) + (\beta_1 + u_{2,i})x_j + \varepsilon_{ij},$$
$$\varepsilon_{ij} \sim N(0, \sigma^2) \; i.i.d$$
$$u_{1,i} \sim N(0, \sigma_1^2), u_{2,i} \sim N(0, \sigma_2^2), cor(u_1, u_2) = \rho$$

- Schätze: $\beta_0, \beta_1, \sigma, \sigma_1, \sigma_2, \rho$

i: Person
j: Woche

Wiederholte Messungen 3/3: Random Slope and Random Intercept (RIRS)

- $y_{ij} = \left(\beta_0 + u_{1,i}\right) + (\beta_1 + u_{2,i})x_j + \varepsilon_{ij},$
  $\varepsilon_{ij} \sim N(0, \sigma^2)\ i.i.d$
  $u_{1,i} \sim N(0, \sigma_1^2), u_{2,i} \sim N(0, \sigma_2^2), cor(u_1, u_2) = \rho$
- Angenommen folgende Werte werden geschätzt: $\beta_0 = 100, \beta_1 = 5, \sigma = 1, \sigma_1 = 10, \sigma_2 = 1, \rho = -0.7$
- Stimmt folgende Aussage (gegeben diese Schätzwerte): "Personen mit einer unterdurchschnittlichen Anfangskraft profitieren überdurchschnittlich von dem Trainingsprogramm."

Zusammenfassung: Wiederholte Messungen

- Block Effekte (fixe Effekte):
  Statistische Aussage für Individuen, aber nicht Bevölkerung
- Mixed effects:
  Statistische Aussage für Bevölkerung, aber nicht Individuen
  - Random Intercept (RI): Individueller Achsenabschnitt
  - Random Intercept and Random Slope (RIRS): Individueller Achsenabschnitt und Steigung

  Komplexere Modelle sind möglich, aber schwieriger zu fitten

Fix oder Random ?

- Wie wirkt Krafttraining bei den 11 Spielern der Fussball-Nati? → fixe Effekte, da Information über genau diese 11 Spieler gewünscht wird
- 11 zufällige Probanden; wie stark streut der Kraftzuwachs durch unser Trainingsprogramm in der Bevölkerung → zufällige Effekte (mixed models), da Information über die zu Grunde liegende Bevölkerung gewünscht wird

Schätzen von Mixed Effects Modellen

- Maximum Likelihood (ML):
  - Varianzschätzungen haben Bias
  + Tests zw. Modellen mit verschiedenen fixen Effekten möglich

- Restricted Maximum Likelihood (REML):
  + Varianzschätzungen haben keinen Bias
  - Kann nur Modelle mit gleichen fixen Effekten vergleichen

Empfohlen für den endgültigen Fit (default in R)

Mixed Effects Modelle in R

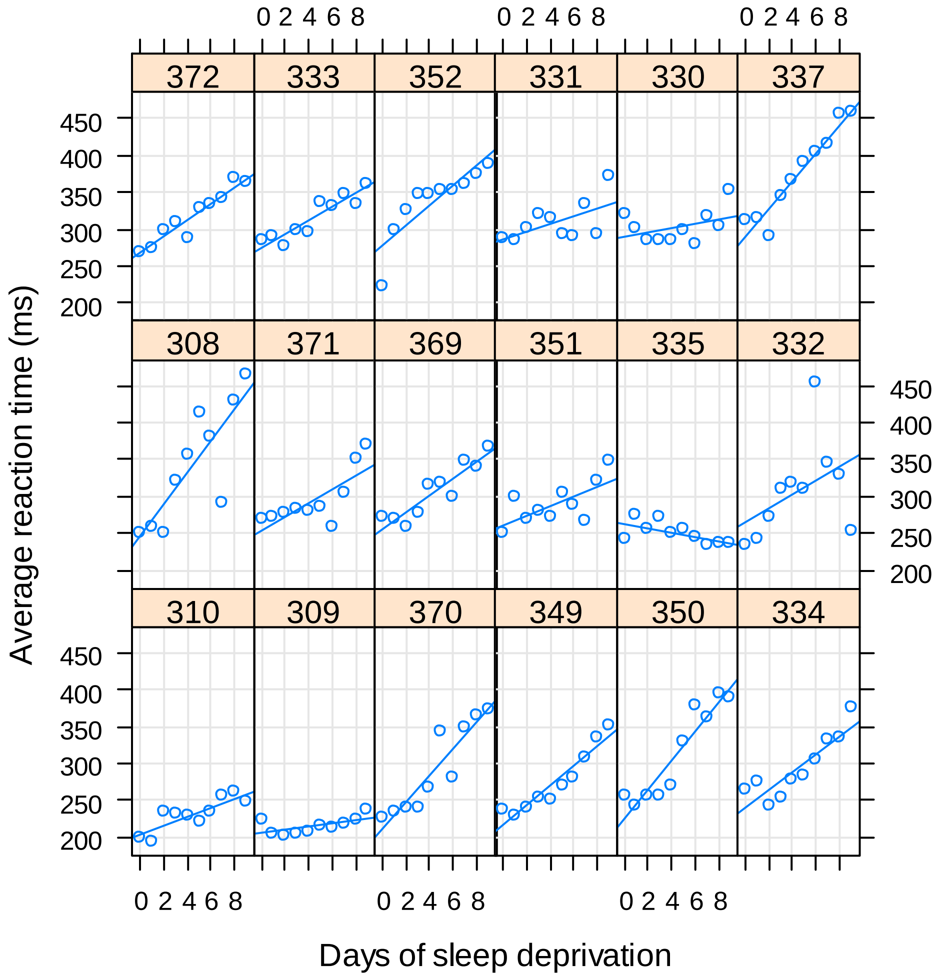
- Funktion “lmer” in Paket “lme4”
- Paket “lmerTest” enthält verbesserte Routinen zum Berechnen von p-Werten (der fixen Effekte).
- Paket “lattice” hilft beim plotten von wiederholten Messungen

Bsp: Reaktionszeit

- 18 Fernfahrer mit Schlafentzug (3h Schlaf pro Nacht)
- Wie ändert sich Reaktionszeit im Verlauf der Tage ?
- Siehe “?sleepstudy” in R

ETH zürich

Reaktionszeit - Überblick



RIRS Modell in R: Input

Zufällige Schwankung in Achsenabschnitt und Steigung pro Person

```
fm1 <- lmer(Reaction ~ Days + (Days | Subject), sleepstudy)
```

Achsenabschnitt und Steigung für Gesamtbevölkerung

RIRS Modell in R: Output

```
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Subject  (Intercept) 612.09   24.740
          Days         35.07    5.922   0.07
 Residual             654.94   25.592
Number of obs: 180, groups:  Subject, 18

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)
(Intercept)  251.405      6.825  16.998  36.838  < 2e-16
Days          10.467      1.546  16.995   6.771 3.27e-06
```

$$y_{ij} = \left(251.4 + u_{1,i}\right) + (10.5 + u_{2,i})x_j + \varepsilon_{ij},$$
$$\varepsilon_{ij} \sim N(0, 25.6^2\ )\ i.i.d$$
$$u_{1,i} \sim N(0, 24.7^2\ ), u_{2,i} \sim N(0, 5.9^2), cor(u_1, u_2) = 0.07$$

Residuenanalyse

- Residuenanalyse wie in Linearer Regression:
  - Tukey-Anscombe Plot
  - QQ-Plot der Residuen

- Zusätzlich: Zufällige Schwankungen des Achsenabschnitts und der Steigung müssen normalverteilt sein
  → QQ-Plots der zufälligen Schwankungen
  (mit Funktion “ranef”)

RIRS oder RI ?

- Passt das RIRS-Modell signifikant besser als das RI-Modell?

```
> fm2 <- lmer(Reaction ~ Days + (1|Subject), sleepstudy)
> anova(fm1, fm2)
refitting model(s) with ML (instead of REML)
Data: sleepstudy
Models:
..1: Reaction ~ Days + (1 | Subject)
object: Reaction ~ Days + (Days | Subject)
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
..1     4 1802.1 1814.8 -897.04   1794.1
object  6 1763.9 1783.1 -875.97   1751.9 42.139      2  7.072e-10
```

fm1 hat
tieferes (=besseres)
AIC und BIC

fm1 passt sign. besser

- Fazit: RIRS-Modell passt besser als RI-Modell

RIRS Modell in R: Zufällige Schwankungen

```
> ranef(fm1)
$Subject
    (Intercept)        Days
308   2.2585654   9.1989719
309 -40.3985769  -8.6197032
310 -38.9602458  -5.4488799
330  23.6904985  -4.8143313
331  22.2602027  -3.0698946
332   9.0395259  -0.2721707
333  16.8404311  -0.2236244
334  -7.2325792   1.0745761
335  -0.3336958 -10.7521591
337  34.8903508   8.6282840
349 -25.2101104   1.1734142
350 -13.0699567   6.6142050
351   4.5778352  -3.0152572
352  20.8635924   3.5360133
369   3.2754530   0.8722166
370 -25.6128694   4.8224646
371   0.8070397  -0.9881551
372  12.3145393   1.2840297
```

Z.B. Geradengleichung für Person 308:

$$y_{ij} = (251.4 + 2.3) + (10.5 + 9.2)x_j + \varepsilon_{ij}$$

(andere Parameter wie bisher)

RIRS Modell in R: Vertrauensintervalle

```
> confint(fm1)
Computing profile confidence intervals ...
                  2.5 %      97.5 %
.sig01       14.3817933  37.7159968
.sig02       -0.4815007   0.6849859
.sig03        3.8011642   8.7533832
.sigma       22.8982669  28.8579966
(Intercept) 237.6806955 265.1295146
Days          7.3586533  13.5759188
```

σ_1

ρ

σ_2

σ

β_0, β_1

Bessere Genauigkeit

Wir haben gesehen: Die Streuung für personenabhängige Schwankung der Steigung (σ_1) ist im Bereich ca. 14.4 bis 37.7 (95%-VI). Angenommen, wir wollen diesen Bereich in einer neuen Studie verkleinern. Welche Massnahme ist dazu sinnvoll?

1) Gleich viele Personen, mehr Wochen pro Person
2) Mehr Personen, gleich viele Wochen pro Person

RIRS Modell: Interpretation

```
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Subject  (Intercept) 612.09   24.740
          Days         35.07    5.922   0.07
 Residual             654.94   25.592
Number of obs: 180, groups:  Subject, 18

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)
(Intercept)  251.405      6.825  16.998  36.838  < 2e-16
Days          10.467      1.546  16.995   6.771 3.27e-06
```

```
> confint(fm1)
Computing profile confidence intervals ...
                  2.5 %      97.5 %
.sig01       14.3817933  37.7159968
.sig02       -0.4815007   0.6849859
.sig03        3.8011642   8.7533832
.sigma       22.8982669  28.8579966
(Intercept) 237.6806955 265.1295146
Days          7.3586533  13.5759188
```

- Die mittlere Reaktionszeit (zu Beginn des Experiments) ist 251 ms (95%-VI: [238 ms, 265 ms] – Genauigkeit der Schätzung)
- Eine typische Schwankung der (anfänglichen) Reaktionszeit in der Bevölkerung ist ca. 25 ms (95%-VI: [14 ms, 38 ms] – Streuung in der Bevölkerung)
- Pro Nacht mit Schlafentzug wird die Reaktionszeit im Mittel um 10 ms schlechter (95%-VI: [7 ms/Tag, 14 ms/Tag] – Genauigkeit der Schätzung)
- Eine typische Schwankung der Reaktion auf Schlafentzug ist ca. 6 ms/Tag (95%-VI: [3.8 ms/Tag, 8.8 ms/Tag] – Streuung in der Bevölkerung)
- Es gibt keinen signifikanten Zshg zwischen anfänglicher Reaktionszeit und Wirkung des Schlafentzugs (95%-VI für ρ: [-0.48; 0.68])

Wichtige Konzepte

- Formel und Interpretation von RI und RIRS Modell
- Residuenanalyse

Wichtige R Funktionen

- Paket “lme4” bzw. “lmerTest”:
  “lmer”, “anova”, “ranef”, “confint”