Repeated measures

Nora Wickelmaier

Last modified: December 16, 2024

Outline

1 Repeated measures ANOVA

Model with one repeated measurement factor

- When subjects are observed for more than two time points, we can model these data using a repeated measures ANOVA
- Data layout

	Time point					
Subject	1	2	$\cdots$	j	$\cdots$	n
1	y_{11}	y_{12}	$\cdots$	y_{1j}	$\cdots$	y_{1n}
2	y_{21}	y_{22}	$\cdots$	y_{2j}	$\cdots$	y_{2n}
$\vdots$	$\vdots$					$\vdots$
i	y_{i1}	y_{i2}	$\cdots$	y_{ij}	$\cdots$	y_{in}
$\vdots$	$\vdots$					$\vdots$
N	y_{N1}	y_{N2}	$\cdots$	y_{Nj}	$\cdots$	y_{Nn}

Statistical model

- When we observe $i = 1, \ldots, N$ subjects for $j = 1, \ldots, n$ time points, we get

$$y_{ij} = \mu_0 + \tau_j + \upsilon_i + \varepsilon_{ij}$$

with

μ_0	grand mean
τ_j	effect of time point j equal for all subjects
υ_i	effect of subject i constant over time
ε_{ij}	error term for subject i at time point j

- Assumptions
  - $\upsilon_i \sim N(0, \sigma_\upsilon^2)$ i.i.d., σ_υ^2 being the variance between subjects
  - $\varepsilon_{ij} \sim N(0, \sigma^2)$ i.i.d., σ^2 being the variance within subjects
  - υ_i and ε_{ij} are independent

Statistical model

$$y_{ij} = \mu_0 + \tau_j + \upsilon_i + \varepsilon_{ij}$$

response	grand mean	time effect	subject effect	residual
y_{11}	μ_0	τ_1	υ_1	ε_{11}
y_{12}	μ_0	τ_2	υ_1	ε_{12}
y_{13}	μ_0	τ_3	υ_1	ε_{13}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_{21}	μ_0	τ_1	υ_2	ε_{21}
y_{22}	μ_0	τ_2	υ_2	ε_{22}
y_{23}	μ_0	τ_3	υ_2	ε_{23}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_{Nn}	μ_0	τ_n	υ_N	ε_{Nn}

Statistical model

$$y_{ij} = \mu_0 + \tau_j + \upsilon_i + \varepsilon_{ij}$$

response	grand mean	time effect	subject effect	residual
y_{11}	μ_0	τ_1	υ_1	ε_{11}
y_{12}	μ_0	τ_2	υ_1	ε_{12}
y_{13}	μ_0	τ_3	υ_1	ε_{13}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_{21}	μ_0	τ_1	υ_2	ε_{21}
y_{22}	μ_0	τ_2	υ_2	ε_{22}
y_{23}	μ_0	τ_3	υ_2	ε_{23}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_{Nn}	μ_0	τ_n	υ_N	ε_{Nn}
$\sigma_\upsilon^2 + \sigma^2$			σ_υ^2	σ^2

Model properties

- For the model, we have

$$\begin{aligned} E(y_{ij}) &= \mu_0 + \tau_j \\ Var(y_{ij}) &= Var(\mu_0 + \upsilon_i + \tau_j + \varepsilon_{ij}) = \sigma_\upsilon^2 + \sigma^2 \\ Cov(y_{ij}, y_{i'j}) &= 0 \text{ for subjects } i \neq i' \\ Cov(y_{ij}, y_{ij'}) &= \sigma_\upsilon^2 \text{ for observations } j \neq j' \end{aligned}$$

- The correlation between observations and subjects is

$$Corr(y_{ij}, y_{ij'}) = \frac{\sigma_\upsilon^2}{\sigma_\upsilon^2 + \sigma^2}$$

known as intraclass correlation (ICC)

Compound Symmetry

- For the covariance matrix of the observations for one subject we get the so-called compound symmetry structure

$$\mathbf{\Sigma}_{\mathbf{y}_i} = \sigma_v^2 \mathbf{1}\mathbf{1}' + \sigma^2 \mathbf{I} = \begin{pmatrix} \sigma_v^2 + \sigma^2 & \sigma_v^2 & \sigma_v^2 & \cdots & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 + \sigma^2 & \sigma_v^2 & \cdots & \sigma_v^2 \\ \vdots & & \ddots & & \vdots \\ \sigma_v^2 & \sigma_v^2 & \sigma_v^2 & \cdots & \sigma_v^2 + \sigma^2 \end{pmatrix}$$

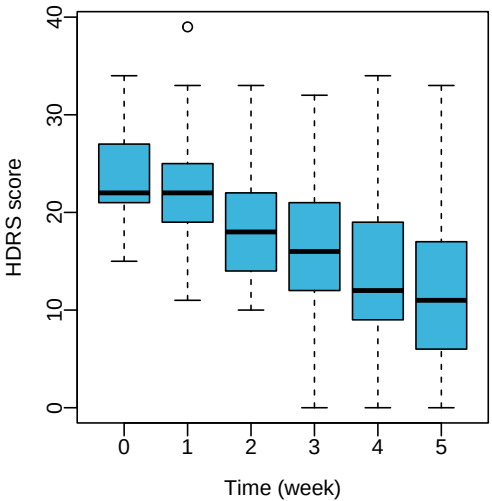
- The assumption of a compound symmetry structure is usually unrealistic for longitudinal data
- In general, successive observations are more strongly correlated than observations being farther apart (covariance is not constant)
- Variance increases with time, e. g., when some subjects are more responsive to a certain treatment than others

2 Example: Depression and Imipramin

Depression and Imipramin (Reisby et al., 1977)

- Reisby et al. (1977) studied the effect of Imipramin on 66 inpatients treated for depression
- Depression was measured with the Hamilton depression rating scale (HDRS)
- Additionally, the concentration of Imipramin and its metabolite Desipramin was measured in their blood plasma
- Patients were classified into endogenous and non-endogenous depressed
- Depression was measured weekly for 6 time points; the effect of the antidepressant was observed starting at week 2 for four weeks

Descriptive statistics



HDRS score

t	W0	W1	W2	W3	W4	W5
M	23.44	21.84	18.31	16.42	13.62	11.95
SD	4.53	4.70	5.49	6.42	6.97	7.22
n	61	63	65	65	63	58

Empirical correlation matrix of HDRS score

	W0	W1	W2	W3	W4	W5
Week 0	1	.49	.41	.33	.23	.18
Week 1	.49	1	.49	.41	.31	.22
Week 2	.41	.49	1	.74	.67	.46
Week 3	.33	.41	.74	1	.82	.57
Week 4	.23	.31	.67	.82	1	.65
Week 5	.18	.22	.46	.57	.65	1

Depression and Imipramin

```
dat      <- read.table("data/reisby.dat", header = TRUE)
dat$id   <- factor(dat$id)
dat$diag <- factor(dat$diag, levels = c("nonen", "endog"))
dat      <- na.omit(dat)      # drop missing values

# descriptive statistics
aggregate(hamd ~ week, dat, mean)
aggregate(hamd ~ week, dat, sd)
aggregate(hamd ~ week, dat, length)
dat[, c("hamd", "id", "week")] |>
  reshape(direction = "wide", timevar = "week") |>
  dplyr::select(-id) |>
  cor(use = "pairwise.complete.obs") |>
  round(3)
```

Fitting repeated-measures ANOVA

```r
# week needs to be a factor when computing an ANOVA
dat$week2 <- factor(dat$week)
contrasts(dat$week2) <- "contr.sum"   # effect coding
summary(aov(hamd ~ week2 + Error(id + id:week2), dat))
# --> ??

library(ez)   # "SPSS"-style
ezANOVA(data = dat, dv = hamd, wid = id, within = week2, type = 3)

# check data
ezDesign(data = dat, x = week2, y = id, col = diag)
replications(hamd ~ week2 + id, dat)
```

Fitting repeated-measures ANOVA

```
# remove IDs with missing observations
ids <- names(which(replications(hamd ~ id, dat)$id == 6))
dat_val <- dat[dat$id %in% ids, ]

# fit ANOVAs again
aov1 <- aov(hamd ~ week2 + Error(id/week2), dat_val)
summary(aov1)

ez1 <- ezANOVA(data = dat_val, dv = hamd, wid = id, within = week2)
ez1$ANOVA
```

Fitting repeated-measures ANOVA

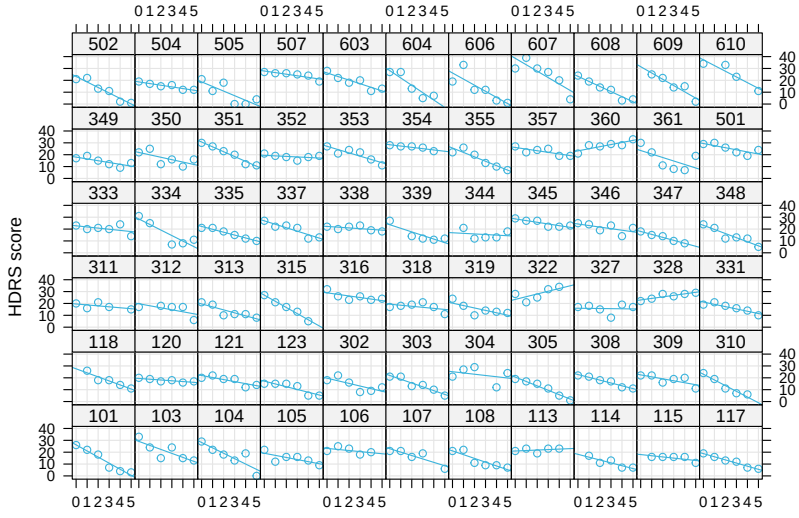
```
# How close can we get with a mixed-effects model?
library(lme4)

lme1 <- lmer(hamd ~ week2 + (1 | id), dat_val)
anova(lme1)

# calculate mean sum of squares for id by hand
sp <- attr(VarCorr(lme1)$id, "stddev")
se <- sigma(lme1)
se^2 + 6 * sp^2
```

❸ Mixed-effects regression

Depression and Imipramin – individual processes



Alternative model with a constant time term

$$y_{ij} = \beta_0 + \beta_1 time + \upsilon_{0i} + \varepsilon_{ij}$$

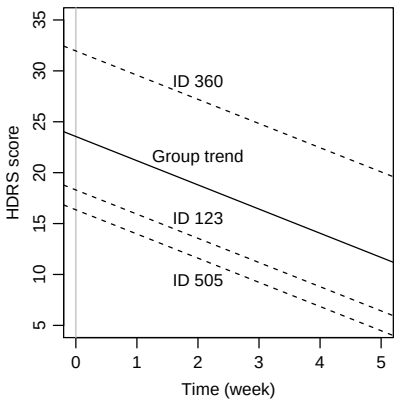
with $\upsilon_{0i} \sim N(0, \sigma_\upsilon^2)$ i.i.d., $\varepsilon_{ij} \sim N(0, \sigma^2)$ i.i.d., υ_{0i} and ε_{ij} i.i.d.

response	intercept	time effect	time	subject effect	residual
y_{11}	β_0	β_1	0	υ_1	ε_{11}
y_{12}	β_0	β_1	1	υ_1	ε_{12}
y_{13}	β_0	β_1	2	υ_1	ε_{13}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_{21}	β_0	β_1	0	υ_2	ε_{21}
y_{22}	β_0	β_1	1	υ_2	ε_{22}
y_{23}	β_0	β_1	2	υ_2	ε_{23}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_{Nn}	β_0	β_1	n	υ_N	ε_{Nn}

Random intercept model

$$y_{ij} = \beta_0 + \beta_1 time + \upsilon_{0i} + \varepsilon_{ij}$$

with $\upsilon_{0i} \sim N(0, \sigma_\upsilon^2)$ i.i.d., $\varepsilon_{ij} \sim N(0, \sigma^2)$ i.i.d., υ_{0i} and ε_{ij} i.i.d.



- The estimated mean baseline HDRS score is $\hat{\beta}_0 = 23.55$
- However, the estimated standard deviation between patients is $\hat{\sigma}_\upsilon = 4.02$
- The mean improvement per week is $\hat{\beta}_1 = -2.38$

Implied marginal covariance matrix

- For the three time points $t_{ij} = 0, 1, 2$, $\mathbf{Z}_i = \mathbf{1}'_{n_i}$ and $\mathbf{\Sigma}_v = \sigma_v^2$ we get

$$\begin{aligned} Cov(\mathbf{y}_i) &= \mathbf{Z}_i \mathbf{\Sigma}_v \mathbf{Z}_i' + \sigma^2 \mathbf{I}_{n_i} \\ &= \sigma_v^2 \mathbf{1}_{n_i} \mathbf{1}'_{n_i} + \sigma^2 \mathbf{I}_{n_i} \\ &= \begin{pmatrix} \sigma_v^2 + \sigma^2 & \sigma_v^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 + \sigma^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 & \sigma_v^2 + \sigma^2 \end{pmatrix} \end{aligned}$$

- The random intercept model implies the compound symmetry structure

Random slope model

$$y_{ij} = \beta_0 + \beta_1 time + \upsilon_{0i} + \upsilon_{1i} time + \varepsilon_{ij}$$

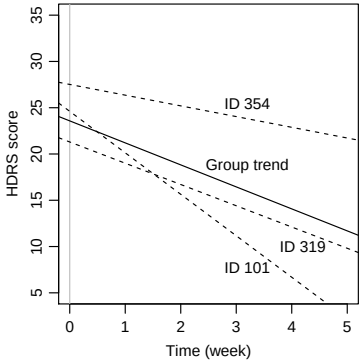
with

$$\begin{pmatrix} \upsilon_{0i} \\ \upsilon_{1i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{\Sigma}_\upsilon = \begin{pmatrix} \sigma^2_{\upsilon_0} & \sigma_{\upsilon_0\upsilon_1} \\ \sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_1} \end{pmatrix} \right) \text{ i.i.d.}$$

$$\varepsilon_i \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i}) \text{ i.i.d.}$$

- Individual intercepts and slopes each have a unique variance component and correlate with $\varrho_{\upsilon_0\upsilon_1} = \frac{\sigma_{\upsilon_0\upsilon_1}}{\sigma_{\upsilon_0}\,\sigma_{\upsilon_1}}$

Model predictions



- The estimated mean baseline HDRS score is $\hat{\beta}_0 = 23.58$
- The estimated standard deviation between patients is $\hat{\sigma}_{v_0} = 3.55$
- The mean improvement per week is $\hat{\beta}_1 = -2.38$
- The estimated standard deviation between patients is $\hat{\sigma}_{v_1} = 1.44$

- The estimated correlation between individual intercepts and slopes is $\hat{\varrho}_{v_0 v_1} = -0.28$
- Patients with higher (that means worse) baseline scores improve more strongly than patients with smaller baseline scores

Implied marginal covariance matrix

- For the three time points $t_{ij} = 0, 1, 2$,

$$\mathbf{Z}_i = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \text{ und } \mathbf{\Sigma}_\upsilon = \begin{pmatrix} \sigma^2_{\upsilon_0} & \sigma_{\upsilon_0\upsilon_1} \\ \sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_1} \end{pmatrix}$$

we get

$$\begin{aligned} Cov(\mathbf{y}_i) &= \mathbf{Z}_i \mathbf{\Sigma}_\upsilon \mathbf{Z}_i' + \sigma^2 \mathbf{I}_{n_i} \\ &= \begin{pmatrix} \sigma^2_{\upsilon_0} & \sigma^2_{\upsilon_0} + \sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_0} + 2\sigma_{\upsilon_0\upsilon_1} \\ \sigma^2_{\upsilon_0} + \sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_0} + 2\sigma_{\upsilon_0\upsilon_1} + \sigma^2_{\upsilon_1} & \sigma^2_{\upsilon_0} + 3\sigma_{\upsilon_0\upsilon_1} + 2\sigma^2_{\upsilon_1} \\ \sigma^2_{\upsilon_0} + 2\sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_0} + 3\sigma_{\upsilon_0\upsilon_1} + 2\sigma^2_{\upsilon_1} & \sigma^2_{\upsilon_0} + 4\sigma_{\upsilon_0\upsilon_1} + 4\sigma^2_{\upsilon_1} \end{pmatrix} + \sigma^2 \mathbf{I}_{n_i} \end{aligned}$$

- Hence, a more flexible covariance structure when compared to compound symmetry

Fitting mixed-effects models

```
# random intercept model
lme1 <- lmer(hamd ~ week + (1 | id), dat, REML = FALSE)
summary(lme1)

# random slope model
lme2 <- lmer(hamd ~ week + (week | id), dat, REML = FALSE)
summary(lme2)

# model comparison
anova(lme1, lme2)
```

Wald test

- For the fixed effects, based on the covariance matrix

$$Var(\hat{\beta}) = \left(\sum_{i=1}^{N} \mathbf{X}_i' \boldsymbol{\Sigma}_i^{-1} \mathbf{X}_i \right)^{-1}$$

  we can construct Wald tests analogously to the regular linear model with approximately normally or t distributed test statistics
- First, $\boldsymbol{\Sigma}_i$ must be estimated; hence, the quality of the approximation strongly depends on the quality of the estimation of the variance and covariance components
- Many authors principally discourage using Wald tests for random effects (e. g., Hedeker & Gibbons, 2006, p. 52 or `?lme4::pvalues`)

Likelihood ratio test

- Hypotheses about fixed and random effects can be tested with likelihood ratio tests
- When M_0 is a model that results from a more general model M_1 by parameter restrictions, then the test statistic

$$G^2 = 2 \log \frac{L(M_1)}{L(M_0)} = 2 \left(\log L(M_1) - \log L(M_0) \right)$$

  is approximately χ^2 distributed with
  (Number of parameters in M_1) $-$ (Number of parameters in M_0) degrees of freedom
- However, for fixed effects this test can result in progressive test decisions for small samples (H_0 is rejected too often)
- For random effects and hypotheses of the form H_0: $\sigma_v^2 = 0$, the test is rather conservative

Summary

- Assumptions of repeated measures ANOVA are often too restrictive for measurements taken over time

Summary

- Assumptions of repeated measures ANOVA are often too restrictive for measurements taken over time
- Mixed-effects models allow for a more flexible variance-covariance structure

Summary

- Assumptions of repeated measures ANOVA are often too restrictive for measurements taken over time
- Mixed-effects models allow for a more flexible variance-covariance structure
- For unequal group sizes, repeated measures ANOVA is not defined

Summary

- Assumptions of repeated measures ANOVA are often too restrictive for measurements taken over time
- Mixed-effects models allow for a more flexible variance-covariance structure
- For unequal group sizes, repeated measures ANOVA is not defined
- (For within designs where assumptions of equal variance are more plausible, repeated measures ANOVA is well suited)

Exercise

- Expand the linear mixed-effects model to more than one factor:
  - Add diagnosis ("endogenous" vs. "non-endogenous") as additional between factor
- Test if this factor interacts with `week` using a likelihood-ratio test
- Use parametric bootstrapping to get a sampling distribution for your LRT statistic
- Plot the sampling distribution and add the empirical confidence interval

References

Hedeker, D. R., & Gibbons, R. D. (2006). *Longitudinal data analysis*. John Wiley.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.

Reisby, N., Gram, L. F., Bech, P., Nagy, A., Petersen, G. O., Ortmann, J., Ibsen, I., Dencker, S. J., Jacobsen, O., Krautwald, O., Sondergaard, I., & Christiansen, J. (1977).Imipramine: Clinical effects and pharmacokinetic variability. *Psychopharmacology*, *54*, 263–272.