Statistik III

Advanced Statistics

Florian Wickelmaier

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



Summer 2026

Repeated Measures ANOVA

Models with one repeated measures factor

- If individuals are observed at more than two time points, the analysis may be based on an ANOVA model with repeated measures
- In this case the repeated measures factor is time
- This setup is called a *randomized block design*, with the persons serving as blocks

Data structure

	Time point					
Person	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}

Model

- If $i = 1, \ldots, N$ individuals are observed at $j = 1, \ldots, n$ time points, we obtain

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

with

μ_0	total expected value
τ_j	effect of time point j (equal for all persons)
υ_i	effect of person i (constant across time)
ε_{ij}	error for person i at time point j

- Assumptions
  - $\upsilon_i \sim N(0, \sigma^2_\upsilon)$ i.i.d., σ^2_υ the variance between persons
  - $\varepsilon_{ij} \sim N(0, \sigma^2)$ i.i.d., σ^2 the variance within persons
  - υ_i and ε_{ij} independent
- Constraints $\sum_{j=1}^{n} \tau_j = 0$

Model Features

- The following properties hold

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

- The correlation between different observations within a single person is given by

$$Cor(y_{ij}, y_{ij'}) = \frac{Cov(y_{ij}, y_{ij'})}{\sqrt{Var(y_{ij}) \cdot Var(y_{ij'})}} = \frac{\sigma_\upsilon^2}{\sigma_\upsilon^2 + \sigma^2}$$

(intra-class correlation, ICC)

Compound Symmetry

- This implies the so-called *compound symmetry structure* for the covariance matrix of the observations within a person

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

- The covariance matrix of all observations then is *block-diagonal*

$$Cov(\mathbf{y}) = \begin{pmatrix} \mathbf{\Sigma}_{\mathbf{y}_1} & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{\Sigma}_{\mathbf{y}_2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & \mathbf{\Sigma}_{\mathbf{y}_N} \end{pmatrix}$$

- Compare to $Cov(\mathbf{y}) = \sigma^2 \mathbf{I}$ in the general linear model

Compound Symmetry

- In general, assuming a compound symmetry structure for the covariance matrix seems to be unrealistically strict in case of longitudinal data
  - Successive observations are usually more strongly correlated than more distant ones (covariance not constant)
  - Variance often increases over time, for instance, if some individuals are more responsive to a treatment than others (variance not homogeneous)
- Statistical tests have been proposed to test for compound symmetry, or the somewhat weaker assumption of sphericity (e. g., Mauchly test), but they have poor statistical properties
  - For example, the Mauchly test is known to fail to detect departures from sphericity in small samples and to over-detect them in large samples

ANOVA Table

Sums of squares

Source	df	SS	E(MS)
Subjects	$N-1$	$SS_S = n\sum_{i=1}^{N}(\bar{y}_{i+} - \bar{y}_{++})^2$	$\sigma^2 + n\sigma_v^2$
Time	$n-1$	$SS_T = N\sum_{j=1}^{n}(\bar{y}_{+j} - \bar{y}_{++})^2$	$\sigma^2 + \frac{N}{n-1}\sum_j \tau_j^2$
Residual	$(N-1)\times(n-1)$	$SS_R = \sum_i\sum_j(y_{ij} - \bar{y}_{i+} - \bar{y}_{+j} + \bar{y}_{++})^2$	σ^2
Total	$Nn-1$	$SS_{tot} = \sum_i\sum_j(y_{ij} - \bar{y}_{++})^2$	

- Moment estimators of variance components in balanced designs

$$\hat{\sigma}_v^2 = (MS_S - MS_R)/n, \qquad \hat{\sigma}^2 = MS_R$$

Hypothesis Testing

- Testable hypotheses and related F statistics

$$\mathrm{H}_0^{\tau}: \quad \tau_j = 0 \quad \forall j \qquad F^{\tau} = \frac{MS_T}{MS_R} \sim F_{n-1,(N-1)(n-1)}$$

$$\mathrm{H}_0^{\sigma_\upsilon}: \qquad \sigma_\upsilon^2 = 0 \qquad F^{\sigma_\upsilon} = \frac{MS_S}{MS_R} \sim F_{N-1,(N-1)(n-1)}$$

Two-Way Repeated Measures ANOVA

Models with repeated measures and group factors

- If subjects are divided into different (experimental) groups, it is of interest whether the time effect differs across the groups (time-group interaction)
- This setup is called a *split-plot design*, in which the blocks (persons, whole plots) are each exposed to one level of the group factor and all of the levels of the repeated measures factor

Data structure

		Time point			
Group	Person	1	2	$\cdots$	n
1	1	y_{111}	y_{112}	$\cdots$	y_{11n}
1	2	y_{121}	y_{122}	$\cdots$	y_{12n}
1	$\vdots$	$\vdots$			$\vdots$
1	N_1	y_{1N_11}	y_{1N_12}	$\cdots$	y_{1N_1n}
$\vdots$	$\vdots$	$\vdots$			$\vdots$
s	1	y_{s11}	y_{s12}	$\cdots$	y_{s1n}
s	2	y_{s21}	y_{s22}	$\cdots$	y_{s2n}
s	$\vdots$	$\vdots$			$\vdots$
s	N_s	y_{sN_s1}	y_{sN_s2}	$\cdots$	y_{sN_sn}

Model

- With $h = 1, \ldots, s$ groups, $i = 1, \ldots, N$ persons, and $j = 1, \ldots, n$ time points, one obtains

$$y_{hij} = \mu_0 + \gamma_h + \tau_j + (\gamma\tau)_{hj} + \upsilon_{i(h)} + \varepsilon_{hij}$$

and the additional effects

γ_h	effect of group h
$(\gamma\tau)_{hj}$	interaction effect of group h and time point j

- If there is only one observation per time point (most often the case), then the interaction between person and time can serve as an error term, $\varepsilon_{hij} = (\upsilon\tau)_{i(h)j}$
- Constraints

$$\sum_{h=1}^{s} \gamma_h = \sum_{j=1}^{n} \tau_j = \sum_{h} (\gamma\tau)_{hj} = \sum_{j} (\gamma\tau)_{hj} = 0$$

ANOVA Table

Sums of squares

Source	df	SS	E(MS)
Group	$s-1$	$SS_G = n\sum_{h=1}^{s} N_h(\bar{y}_{h++} - \bar{y}_{+++})^2$	$\sigma^2 + n\sigma_v^2 + \frac{Nn}{s-1}\sum_j \gamma_h^2$
Time	$n-1$	$SS_T = N\sum_{j=1}^{n}(\bar{y}_{++j} - \bar{y}_{+++})^2$	$\sigma^2 + \frac{Ns}{n-1}\sum_j \tau_j^2$
Gr × T	$(s-1)\times$ $(n-1)$	$SS_{GT} = \sum_h \sum_j N_h(\bar{y}_{h+j} - \bar{y}_{h++} - \bar{y}_{++j} + \bar{y}_{+++})^2$	$\sigma^2 + \frac{N}{(s-1)(n-1)} \times \sum_j (\gamma\tau)_{hj}^2$
Subjects	$N-s$	$SS_S = n\sum_h \sum_i (\bar{y}_{hi+} - \bar{y}_{h++})^2$	$\sigma^2 + n\sigma_v^2$
Residual	$(N-s)\times$ $(n-1)$	$SS_R = \sum_h \sum_i \sum_j (y_{hij} - \bar{y}_{hi+} - \bar{y}_{h+j} + \bar{y}_{h++})^2$	σ^2
Total	$Nn-1$	$SS_{tot} = \sum_h \sum_i \sum_j (y_{hij} - \bar{y}_{+++})^2$	

Hypothesis Testing

- Testable hypotheses and related F statistics

$$\mathrm{H}_0^{(\gamma\tau)}: \quad (\gamma\tau)_{hj} = 0, \forall(h,j) \quad F^{(\gamma\tau)} = \frac{MS_{GT}}{MS_R} \sim F_{(s-1)(n-1),(N-s)(n-1)}$$

$$\mathrm{H}_0^{\tau}: \quad \tau_j = 0, \forall j \quad F^{\tau} = \frac{MS_T}{MS_R} \sim F_{n-1,(N-s)(n-1)}$$

$$\mathrm{H}_0^{\gamma}: \quad \gamma_h = 0, \forall h \quad F^{\gamma} = \frac{MS_G}{MS_S} \sim F_{s-1,N-s}$$

$$\mathrm{H}_0^{\sigma_\upsilon}: \quad \sigma_\upsilon^2 = 0 \quad F^{\sigma_\upsilon} = \frac{MS_S}{MS_R} \sim F_{N-s,(N-s)(n-1)}$$

Interpretation

- If the null hypothesis regarding the group-time interaction is rejected, then
  - the differences between groups do not remain constant over time
  - the group-specific curves are not parallel over time
  - group and time effects are confounded and cannot be tested and estimated separately

  Thus, neither are there time-independent group effects nor group-independent time effects
- If there is no significant interaction, then main effects for group and time can be tested independently, but usually are not interesting
  - Group effect: randomization error?
  - Time effect: regression to the mean?

Two Time Points per Group

- Regression representation with indicator variables for time point ($t_{i1} = 0$, $t_{i2} = 1$) and group (x_i)

$$y_{ij} = \beta_0 + \beta_1\, t_{ij} + \beta_2\, x_i + \beta_3\,(t_{ij} \cdot x_i) + \upsilon_i + \varepsilon_{ij}$$

  with $\upsilon_i \sim N(0, \sigma_\upsilon^2)$ i.id., and $\varepsilon_{ij} \sim N(0, \sigma^2)$ i.i.d., and independent from each other

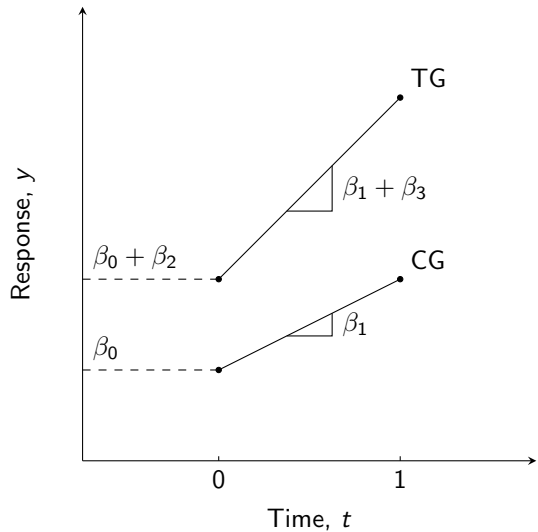
- Interpretation

β_0	mean initial (baseline) value of the reference group
β_1	time effect (slope) in reference group
β_2	effect of treatment group
β_3	effect on the slope in treatment group

Two Time Points per Group

Hypotheses

- $H_0\colon \beta_3 = 0$
  Does the change differ across groups?
- The hypotheses $H_0\colon \beta_1 = 0$ and $H_0\colon \beta_2 = 0$ are not interpretable for $\beta_3 \neq 0$, or uninteresting anyway



Two Time Points per Group

Repeated measures ANOVA is equivalent to change-score analysis

- For $j = 1, 2$, and drawing upon the regression representation of the analysis of variance, we obtain

$$y_{i1} = \beta_0 + \beta_2 x_i + \upsilon_i + \varepsilon_{i1}$$
$$y_{i2} = \beta_0 + \beta_1 + \beta_2 x_i + \beta_3 x_i + \upsilon_i + \varepsilon_{i2}$$

- For the change score this means

$$y_{i2} - y_{i1} = \beta_1 + \beta_3 x_i + (\varepsilon_{i2} - \varepsilon_{i1})$$

- Due to independence of the ε_{ij} we obtain the change-score model

$$y_{i2} - y_{i1} = \beta_1 + \beta_3 x_i + \varepsilon_i$$

with

$$\varepsilon_i = \varepsilon_{i2} - \varepsilon_{i1} \sim N(0, \sigma_d^2 = 2\sigma^2)$$

Summary

- In principle, ANOVA models with time as a repeated measures factor may be considered for longitudinal data with more than two time points
  - The focus in this analysis is the interaction between the group and time factors
- However, the assumption of a compound symmetrical covariance matrix is almost always unrealistic
- Many authors (e. g., McCulloch, 2005) therefore discourage the use of ANOVA models for the analysis of longitudinal data and recommend *linear mixed models*
  - These models allow for a flexible and parsimonious modeling of the covariance structure

References

Fahrmeir, L., Hamerle, A., & Tutz, G. (1996). Multivariate statistische Verfahren. Berlin: Walter de Gruyter. Kapitel 5

Hedeker, D. R., & Gibbons, R. D. (2006). Longitudinal data analysis. Hoboken, NJ: Wiley. Chap. 2

Keppel, G., & Wickens, T. D. (2004). Design and analysis: A researcher's handbook. Upper Saddle River: Prentice Hall.

Lohr, S. L. (1995). Hasse diagrams in statistical consulting and teaching. *American Statistician*, 49, 376–381.

McCulloch, C. E. (2005). Repeated measures ANOVA, R.I.P.? *Chance*, 18, 29–33.

Mixed Models

Linear mixed models

- The assumptions of ANOVA models (compound symmetry, equal number of observations per subject) are often too restrictive for analysis of longitudinal or hierarchical data
- Linear mixed models are a class of statistical models that include both fixed and random effects
- The random effects imply marginal correlation structures that allow for describing clustered data in a highly flexible way

Linear Mixed Models

- A linear mixed model is of the general form

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \boldsymbol{\upsilon}_i + \boldsymbol{\varepsilon}_i$$

with the fixed effects $\boldsymbol{\beta}$, the random effects $\boldsymbol{\upsilon}_i$, the design matrices $\mathbf{X}_i$ and $\mathbf{Z}_i$, and the distributional assumptions

$$\boldsymbol{\upsilon}_i \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\upsilon) \text{ i.i.d.}, \qquad \boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i}) \text{ i.i.d.}$$

and independent

- For the marginal covariance matrix this implies

$$Cov(\mathbf{y}_i) = \boldsymbol{\Sigma}_i = \mathbf{Z}_i \boldsymbol{\Sigma}_\upsilon \mathbf{Z}_i' + \sigma^2 \mathbf{I}_{n_i}$$

- The following models are particular instances of a linear mixed model

Example: Depression and Imipramine

- Reisby et al. (1977) studied the effect of imipramine in 66 patients hospitalized for depressive disorder
- Depressive symptoms were measured with the Hamilton depression rating scale (HDRS) for six weeks
- The concentration of imipramine and its metabolite desipramine in blood plasma was assayed from week 2 on for four weeks
- Furthermore, patients were classified as suffering from endogenous or non-endogenous depression

Example: Depression and Imipramine

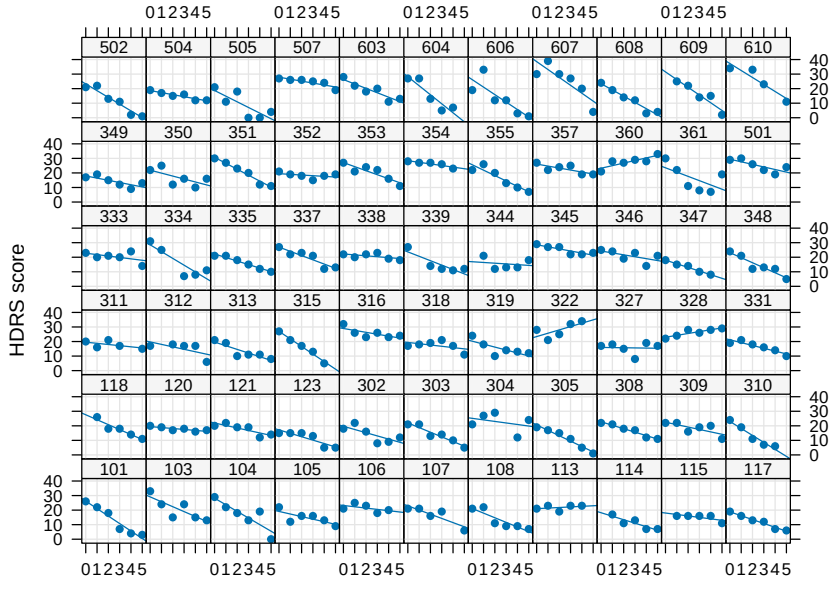
HDRS score

t	week 0	week 1	week 2	week 3	week 4	week 5
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 the HDRS score

	week 0	week 1	week 2	week 3	week 4	week 5
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

Example: Individual Trajectories



Models with Random Intercepts

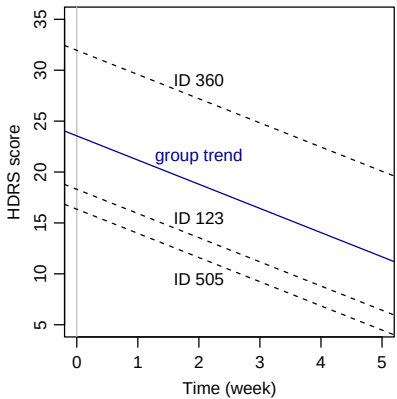
- Model equations

$$
\begin{aligned}
\text{(Level 1)} \quad & y_{ij} = b_{0i} + b_{1i}\, t_{ij} + \varepsilon_{ij} \\
\text{(Level 2)} \quad & b_{0i} = \beta_0 + \upsilon_{0i} \\
& b_{1i} = \beta_1 \\
\text{(2) in (1)} \quad & y_{ij} = \beta_0 + \beta_1\, t_{ij} + \upsilon_{0i} + \varepsilon_{ij}
\end{aligned}
$$

with

$$
\begin{aligned}
& \upsilon_{0i} \sim N(0, \sigma_\upsilon^2) \text{ i.i.d.} \\
& \varepsilon_{ij} \sim N(0, \sigma^2) \text{ i.i.d.} \\
& \upsilon_{0i} \text{ and } \varepsilon_{ij} \text{ independent}
\end{aligned}
$$

Model Predictions



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

Marginal Covariance Matrix

- The random-intercept model implies a certain form of the marginal covariance matrix (of observations)
- For three time points $t_{ij} = 0, 1, 2$, $\mathbf{Z}_i = \mathbf{1}'_{n_i}$ and $\mathbf{\Sigma}_\upsilon = \sigma^2_\upsilon$ one obtains

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

- This means that the random-intercept model implies compound symmetry

Models with Random Slopes

- Model equations

$$
\begin{aligned}
\text{(Level 1)} \quad & y_{ij} = b_{0i} + b_{1i}\, t_{ij} + \varepsilon_{ij} \\
\text{(Level 2)} \quad & b_{0i} = \beta_0 + \upsilon_{0i} \\
& b_{1i} = \beta_1 + \upsilon_{1i} \\
\text{(2) in (1)} \quad & y_{ij} = \beta_0 + \beta_1\, t_{ij} + \upsilon_{0i} + \upsilon_{1i}\, t_{ij} + \varepsilon_{ij}
\end{aligned}
$$

with

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

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

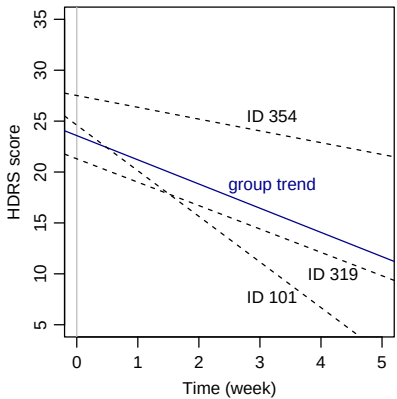
and independent

Models with Random Slopes

- Individual intercepts and slopes both come with their own variance component, and their correlation is given by

$$\rho_{v_0 v_1} = \frac{\sigma_{v_0 v_1}}{\sigma_{v_0}\, \sigma_{v_1}}$$

Model Predictions



- The estimated mean baseline HDRS score is $\hat{\beta}_0 = 23.58$ (estimated standard deviation between patients $\hat{\sigma}_{\upsilon_0} = 3.55$)
- The mean improvement per week is $\hat{\beta}_1 = -2.38$ ($\hat{\sigma}_{\upsilon_1} = 1.44$)
- The estimated correlation between individual intercepts and slopes is $\hat{\rho}_{\upsilon_0\upsilon_1} = -0.28$
  - Patients with higher (worse) baseline scores tend to improve more than patients with lower baseline scores

Marginal Covariance Matrix

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

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

one obtains

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

- This means that the random-slope model implies a more flexible covariance structure compared to compound symmetry

Wald Test

- For the fixed effects, the covariance matrix

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

  can be used to construct Wald tests analogous to the general linear model with approximate normally or t-distributed test statistics
  - Since $\boldsymbol{\Sigma}_i$ must be estimated first, the quality of the approximation depends critically on the precision of the estimation of the variance and covariance components
- Wald tests for random effects are generally discouraged (e. g., Hedeker & Gibbons, 2006, p. 52)

Likelihood Ratio Test

- Hypotheses about fixed or random effects can be tested using a likelihood ratio test
  - Let M_0 be a model that emerges from a more general model M_1 by restricting parameters, 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 degrees of freedom equal to (number of parameters in M_1) − (number of parameters in M_0)
- Known problems
  - For fixed effects, the test can lead to progressive decisions for small samples (H_0 rejected too often)
  - In contrast, for random effects and hypotheses of the form $H_0 \colon \sigma_v^2 = 0$, the test is conservative because the hypothesis lies on the border of the parameter space

Parametric Bootstrapping

- To validate the results of the likelihood ratio test, a *parametric bootstrap* procedure can be used
  - Parameter estimation for the general (M_1) and constrained model (M_0) and calculation of the G^2 statistic
  - Simulation of B bootstrap samples $\mathbf{x}^{*B}$ based on the stochastic component of M_0 (H_0 holds in these samples)
  - For each bootstrap sample: parameter estimation in M_1 and M_0 and computation of a bootstrap replication of G^2
  - Determine the significance of G^2 based on the empirical sampling distribution of the bootstrap replications

Models with Time-Varying Covariates

- Model equations

$$
\begin{aligned}
\text{(Level 1)} \quad & y_{ij} = b_{0i} + b_{1i}\, t_{ij} + b_{2i}\, x_{ij} + \varepsilon_{ij} \\
\text{(Level 2)} \quad & b_{0i} = \beta_0 + \upsilon_{0i} \\
& b_{1i} = \beta_1 + \upsilon_{1i} \\
& b_{2i} = \beta_2 \\
\text{(2) in (1)} \quad & y_{ij} = \beta_0 + \beta_1\, t_{ij} + \beta_2\, x_{ij} + \upsilon_{0i} + \upsilon_{1i}\, t_{ij} + \varepsilon_{ij}
\end{aligned}
$$

- Time-varying covariates enter the model as Level 1 effects

Models with Group Effect d_i

- Model equations

$$
\begin{aligned}
\text{(Level 1)} \quad & y_{ij} = b_{0i} + b_{1i}\, t_{ij} + \varepsilon_{ij} \\
\text{(Level 2)} \quad & b_{0i} = \beta_0 + \beta_2\, d_i + \upsilon_{0i} \\
& b_{1i} = \beta_1 + \beta_3\, d_i + \upsilon_{1i} \\
\text{(2) in (1)} \quad & y_{ij} = \beta_0 + \beta_1\, t_{ij} + \beta_2\, d_i + \beta_3\, (t_{ij} \times d_i) + \upsilon_{0i} + \upsilon_{1i}\, t_{ij} + \varepsilon_{ij}
\end{aligned}
$$

with

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

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

- Variables that remain constant over time enter as Level 2 effects, because they vary across individuals only

Summary

- Linear models with fixed and random effects are flexible models for analyzing longitudinal and hierarchical data
- Advantages compared to ANOVA models
  - Flexible covariance structures through the introduction of random effects
  - Modeling of individual trajectories
  - Automatic handling of unequal observation times and unequal number of observations (missing values)
  - Investigating the effects of time-varying predictors

Mixed Model Software

- Different packages are available in R for dealing with mixed models (e. g., `nlme`, `lme4`)
- In the following we will confine consideration to the `lme4` package
  - It provides functions that implement generalized linear mixed models (GLMM), where the data structure does not have to be exclusively hierarchical (nested structure), but can also include crossed factors (crossed structure)
  - The implemented methods for parameter estimation in crossed factors are much more efficient compared to `nlme`
  - In psychological tests, persons and items are generally crossed (all of the items administered to each of the persons)

Mixed Model Software

`lme4` package

```
library(lme4)
```

- Loads the `matrix` package and needs the `lattice` package (for graphics)

Linear mixed models are specified in the `lmer` function by

```
lmer(y ~ fixed effects + (random effects), data frame, options)
```

- With the option `REML = TRUE` (default) parameters are estimated by restricted maximum likelihood (REML), setting `REML = FALSE` by maximum likelihood (ML)
- REML can produce unbiased estimates of variances and covariances, but likelihood ratio tests of fixed effects require ML estimates

Application of `lmer()` is illustrated for the Reisby et al. (1977) data

Example: Depression and Imipramine

Data

```
library(lattice)
library(lme4)

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

Variables

`id`	Patient ID
`hamd`	Score on the Hamilton depression rating scale
`week`	Time point, week $0, \ldots, 5$
`diag`	Diagnosis of endogenous vs. non-endogenous depression

Example: Depression and Imipramine

Random-intercept model

```
m1 <- lmer(hamd ~ week + (1 | id), dat, REML = FALSE)
```

ML parameter estimation (`m1`)

```
Linear mixed model fit by maximum likelihood ['lmerMod']
Formula: hamd ~ week + (1 | id)
   Data: dat
      AIC       BIC    logLik  deviance
 2293.189  2308.896 -1142.594  2285.189
Random effects:
 Groups   Name        Std.Dev.
 id       (Intercept) 4.019
 Residual             4.363
Number of obs: 375, groups: id, 66
Fixed Effects:
(Intercept)         week
     23.552       -2.376
```

Example: Depression and Imipramine

Mixed model with random slopes

```
m2 <- lmer(hamd ~ week + (week | id), dat, REML = FALSE)
```

ML parameter estimation (m2)

```
Linear mixed model fit by maximum likelihood ['lmerMod']
Formula: hamd ~ week + (week | id)
   Data: dat
      AIC      BIC   logLik  deviance
 2231.037 2254.599 -1109.519 2219.037
Random effects:
 Groups   Name        Std.Dev. Corr
 id       (Intercept) 3.554
          week        1.442    -0.28
 Residual             3.495
Number of obs: 375, groups: id, 66
Fixed Effects:
(Intercept)         week
     23.577       -2.377
```

Example: Depression and Imipramine

Group-specific trajectories

```
m3 <- lmer(hamd ~ week * diag + (week | id), dat, REML = FALSE)
```

ML parameter estimation (m3)

```
Linear mixed model fit by maximum likelihood ['lmerMod']
Formula: hamd ~ week * diag + (week | id)
   Data: dat
      AIC       BIC    logLik  deviance
 2230.929  2262.345 -1107.465  2214.929
Random effects:
 Groups   Name        Std.Dev. Corr
 id       (Intercept) 3.412
          week        1.441    -0.29
 Residual             3.495
Number of obs: 375, groups: id, 66
Fixed Effects:
 (Intercept)         week    diagendog  week:diagendog
      22.476       -2.366        1.988          -0.027
```

Example: Depression and Imipramine

Likelihood ratio test

```
anova(m2, m3)

Data: dat
Models:
m2: hamd ~ week + (week | id)
m3: hamd ~ week * diag + (week | id)
   Df    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
m2  6 2231.0 2254.6 -1109.5   2219.0
m3  8 2230.9 2262.3 -1107.5   2214.9 4.1083  2     0.1282
```

Example: Depression and Imipramine

Mean scores for diagnostic groups

```
dat3 <- aggregate(hamd ~ week + diag, dat, mean)
```

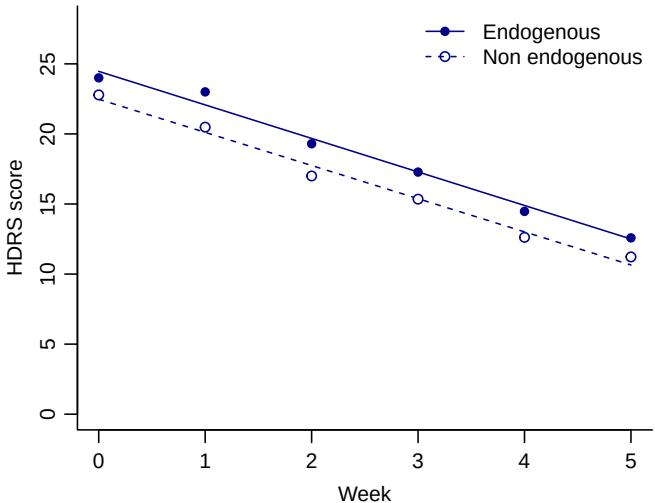
Graphical representation of mean values and predicted trends

```
dat3$m3 <- predict(m3, dat3, re.form = ~ 0)  # prediction w/o random effects

plot(m3 ~ week, dat3[dat3$diag == "endog", ], type = "l", bty = "L",
     ylim = c(0, 28), xlab = "Week", ylab = "HDRS score", col = "darkblue")
lines(m3 ~ week, dat3[dat3$diag == "nonen", ], lty = 2, col = "darkblue")
points(hamd ~ week, dat3[dat3$diag == "endog", ], pch = 16, col = "darkblue")
points(hamd ~ week, dat3[dat3$diag == "nonen", ], pch = 21, col = "darkblue",
       bg = "white")
legend("topright", c("Endogenous", "Non endogenous"),
       lty = 1:2, pch = c(16, 21), bty = "n", col = "darkblue")
```

Example: Depression and Imipramine

Graphical illustration of mean scores and predicted trends



References

Bryk, A. S., & Raudenbush, S. W. (2002). Hierarchical linear models. Thousand Oaks, CA: Sage Publications. Chap. 4

Hedeker, D. R., & Gibbons, R. D. (2006). Longitudinal data analysis. Hoboken, NJ: Wiley. Chap. 4

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.