(Generalized) Linear Mixed Models

BAMB! Summer School
Tutorial 9

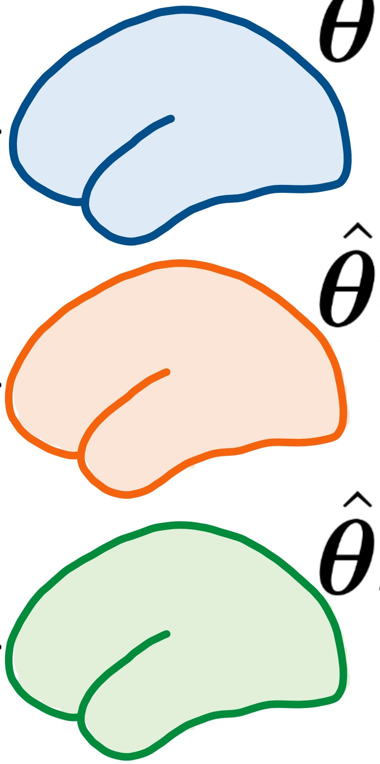
Population-level analyses

Subject 1: $(\boldsymbol{X}_1, \boldsymbol{Y}_1) \longrightarrow$ $\hat{\boldsymbol{\theta}}_1$

Subject 2: $(\boldsymbol{X}_2, \boldsymbol{Y}_2) \longrightarrow$ $\hat{\boldsymbol{\theta}}_2$

....

Subject *m*: $(\boldsymbol{X}_m, \boldsymbol{Y}_m) \longrightarrow$ $\hat{\boldsymbol{\theta}}_m$

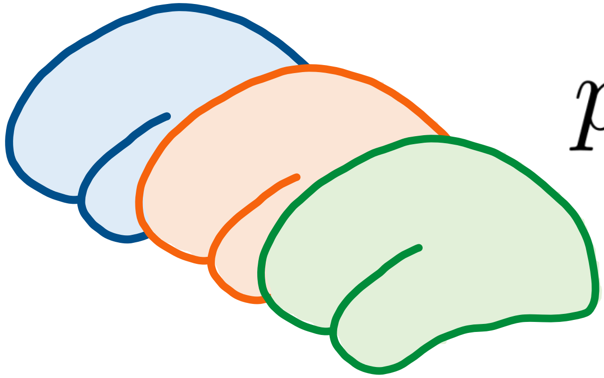


We want to infer from a sample of subjects conclusions about general population(s)

Summary statistics approach

$$\{\hat{\boldsymbol{\theta}}_j\} \longrightarrow p(\boldsymbol{\theta}|\boldsymbol{Y})$$

$$\{(\boldsymbol{X}_j, \boldsymbol{Y}_j)\} \longrightarrow$$

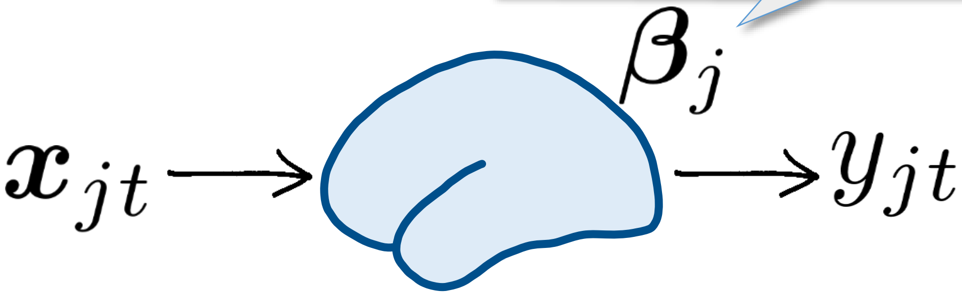


$p(\boldsymbol{\theta}|\boldsymbol{Y})$

Mixed models

Summary-statistic approach

The **parameters** of the model correspond to subject-specific weights

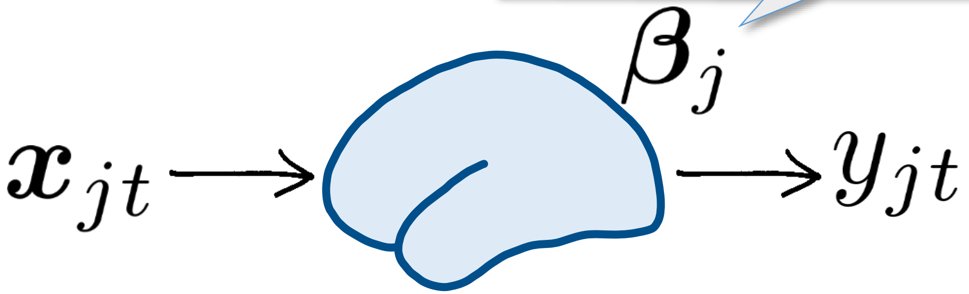


$$y_j = \beta_{j0} + \beta_{j1}x_{j1} + \ldots + \beta_{jm}x_{mj} + \eta$$

Weights are fitted separately for each subject,
Then the distribution of weights over the population is inspected

Summary-statistic approach

The **parameters** of the model correspond to subject-specific weights

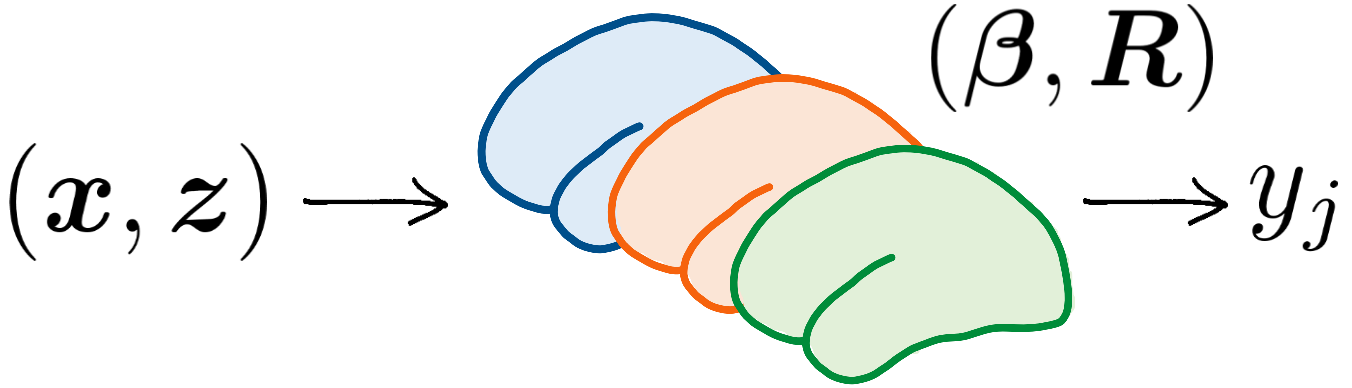


$$y_j = \beta_{j0} + \beta_{j1} x_{j1} + \ldots + \beta_{jm} x_{mj} + \eta$$

R-style formula: $'y \sim 1 + x_1 + \ldots + x_m'$

Weights are fitted separately for each subject,
Then the distribution of weights over the population is inspected

Linear mixed model

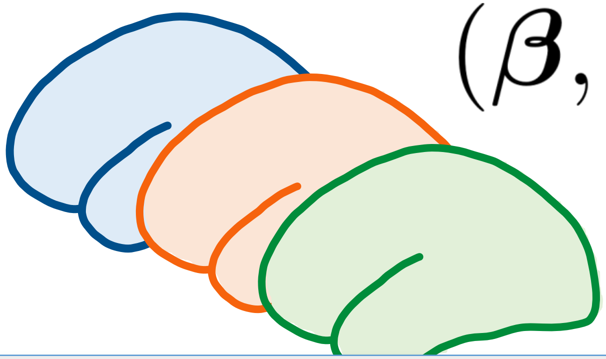


$$y_j = \underbrace{\beta_0 + \ldots + \beta_m x_m}_{\text{fixed effects}} + \underbrace{u_{j1} z_1 + \ldots + u_{jp} z_p}_{\text{random effects}} + \eta$$

fixed effects + random effects + noise

Linear mixed model

$$(\boldsymbol{\beta}, \boldsymbol{R})$$

$$(\boldsymbol{x}, \boldsymbol{z}) \longrightarrow$$  $$\longrightarrow y_j$$

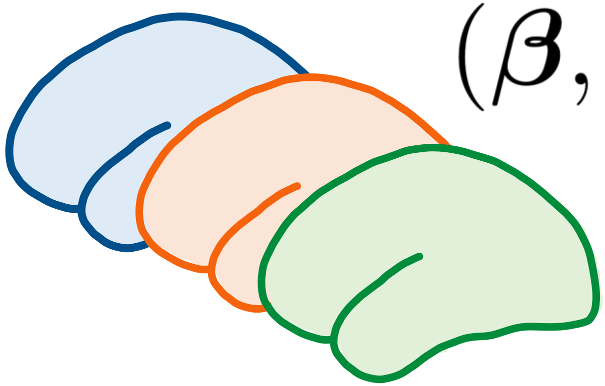
$$y_j = \beta_0 + \cdots \; u_{jp} z_p + \eta$$

This captures the general trend in the population (usually what we are interested in).

fixed effects + random effects + noise

Linear mixed model

$(\boldsymbol{\beta}, \boldsymbol{R})$

$(\boldsymbol{x}, \boldsymbol{z}) \longrightarrow$  $\longrightarrow y_j$

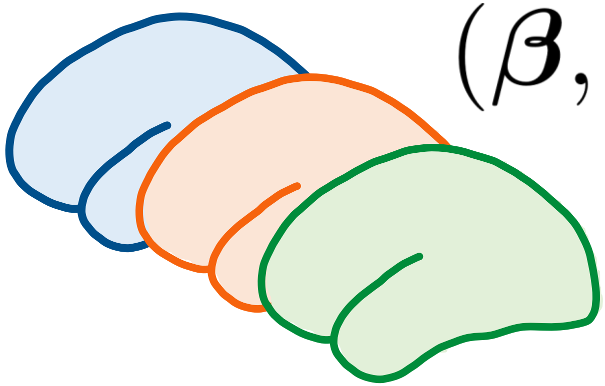
$$y_j = \underbrace{\beta_0 + \ldots + \beta_m x_m}_{\text{fixed effects}} + \underbrace{u_{j1} z \ldots}_{\text{random effects}} + \text{noise}$$

This captures the deviation in individual subjects from the average in the population

Linear mixed model

$$(\boldsymbol{\beta}, \boldsymbol{R})$$

$$(\boldsymbol{x}, \boldsymbol{z}) \longrightarrow$$



$$\longrightarrow y_j$$

$$y_j = \beta_0 + \ldots + \beta_m x_m + u_{j1} z_1 + \ldots + u_{jp} z_p + \eta$$

fixed

+ noise

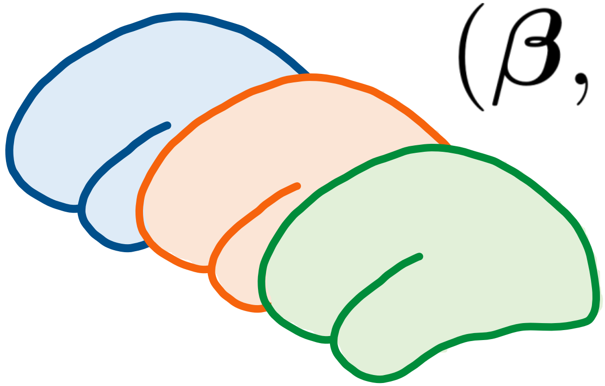
The random effects are treated as *latent* variables coming from a prior. The covariance matrix **R** defines the variability of the random effects across the population.

$$p(\boldsymbol{u}_j) = \mathcal{N}(\boldsymbol{u}; \boldsymbol{0}, \boldsymbol{R})$$

Linear mix

The **parameters** of the model correspond to the fixed-effect weights **β** and the covariance of the random-effects ***R***

$$(\boldsymbol{\beta}, \boldsymbol{R})$$

$$(\boldsymbol{x}, \boldsymbol{z}) \longrightarrow$$  $$\longrightarrow y_j$$

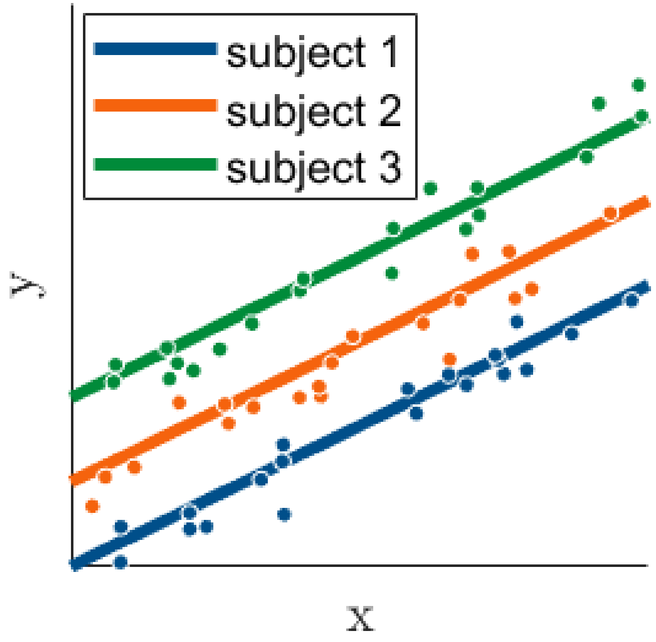
$$y_j = \underbrace{\beta_0 + \ldots + \beta_m x_m}_{\text{fixed effects}} + \underbrace{u_{j1} z_1 + \ldots + u_{jp} z_p}_{\text{random effects}} + \underbrace{\eta}_{\text{noise}}$$

fixed effects + random effects + noise

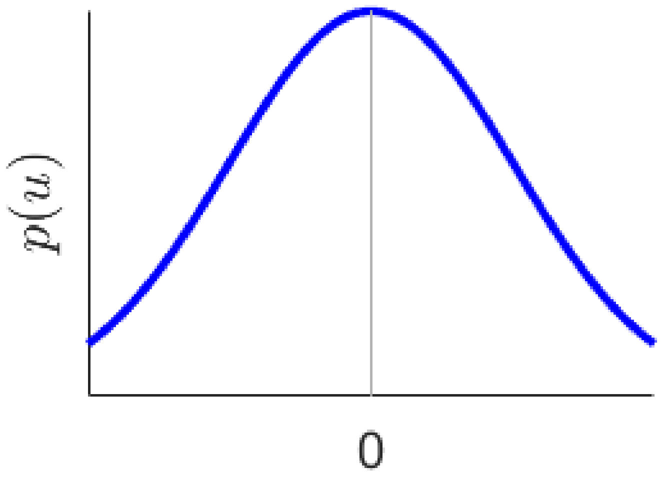
$$p(\boldsymbol{u}_j) = \mathcal{N}(\boldsymbol{u}; \boldsymbol{0}, \boldsymbol{R})$$

Random intercept mixed model

$$y_j = \beta_0 + \ldots + \beta_m x_m + u_j + \eta$$



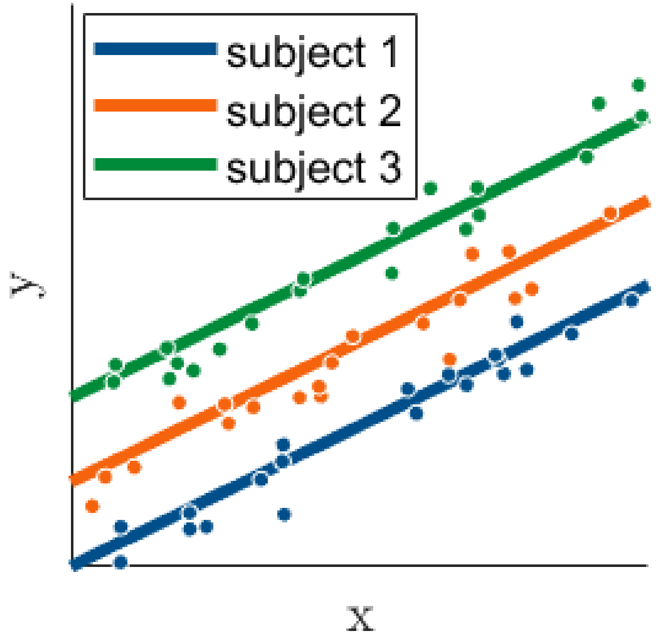
$$p(u) = \mathcal{N}(u; 0, R)$$



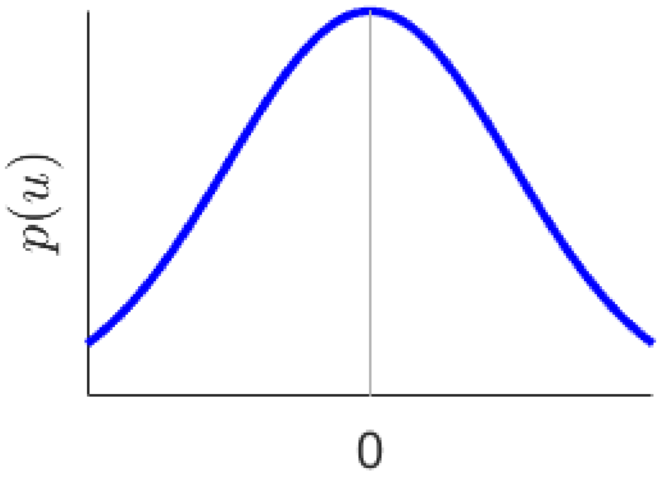
Random intercept mixed model

The intercept varies from subject-to-subject

$$y_j = \beta_0 + \ldots + \beta_m x_m + u_j + \eta$$



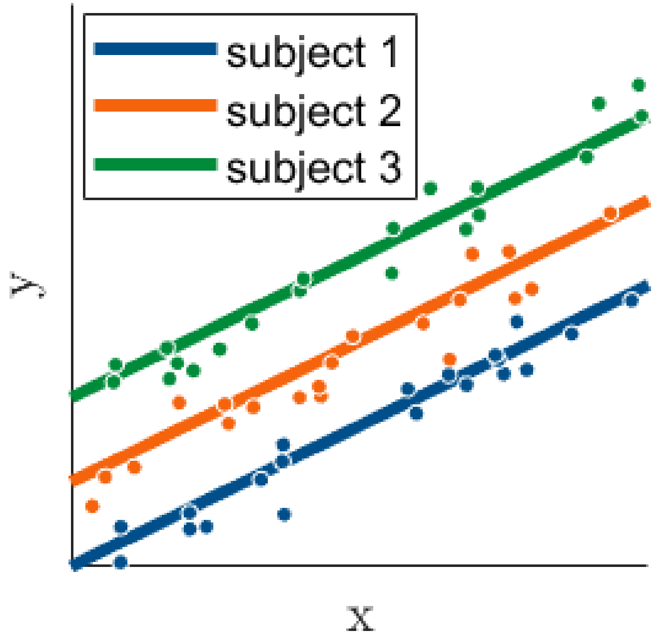
$$p(u) = \mathcal{N}(u; 0, R)$$



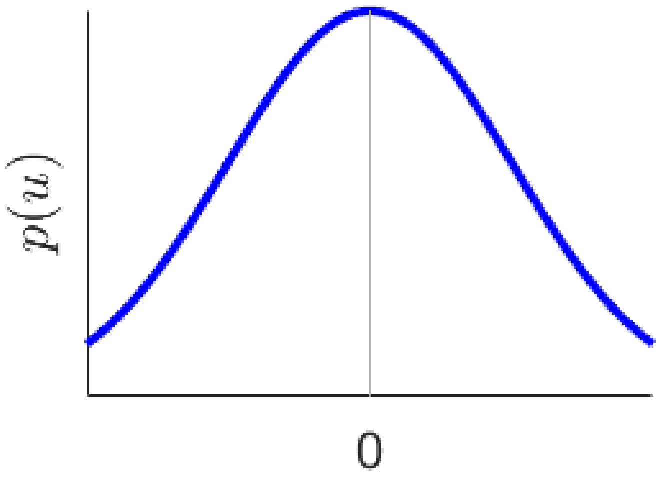
Random intercept mixed model

All other factors are constant (fixed)

$$y_j = \beta_0 + \ldots + \beta_m x_m + u_j + \eta$$

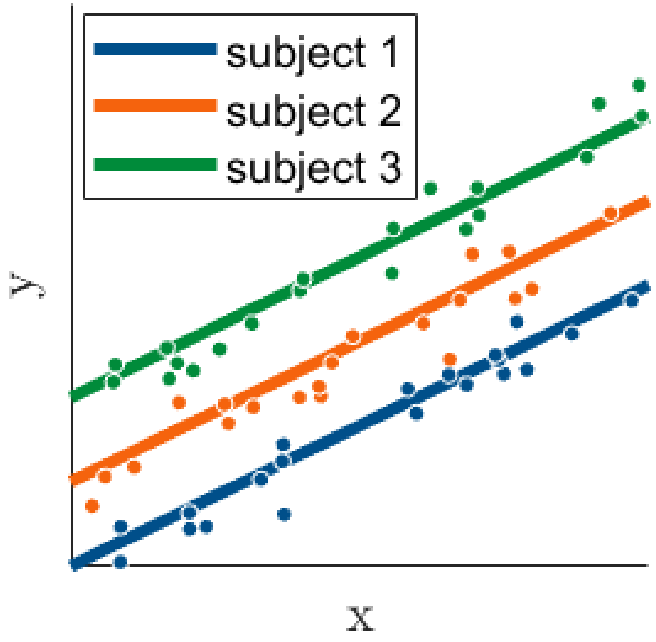


$$p(u) = \mathcal{N}(u; 0, R)$$

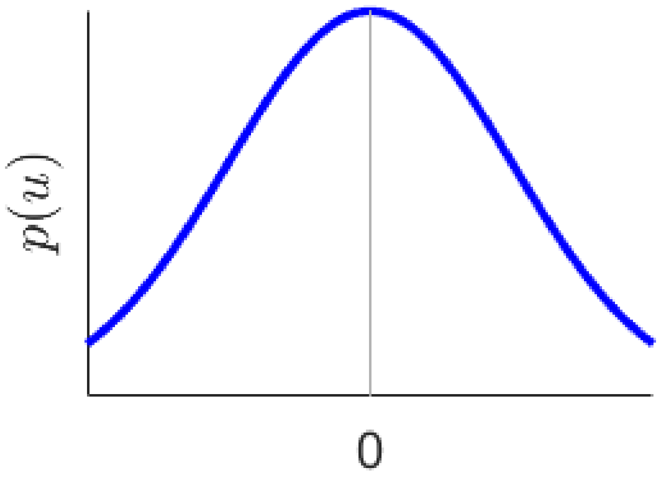


Random intercept mixed model

$$y_j = \beta_0 + \ldots + \beta_m x_m + u_j + \eta$$



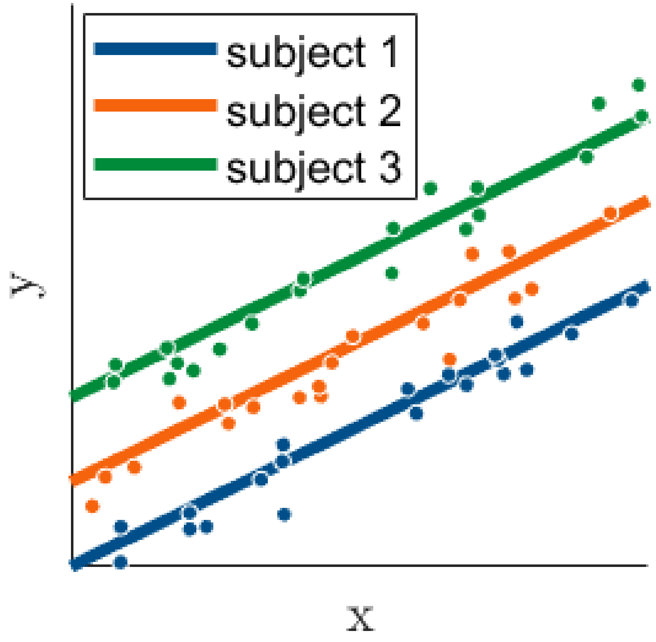
$$p(u) = \mathcal{N}(u; 0, R)$$



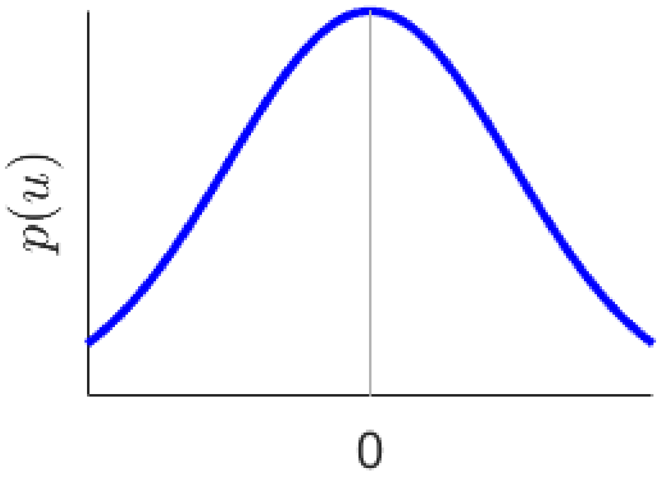
Random intercept mixed model

$$y_j = \beta_0 + \ldots + \beta_m x_m + u_j + \eta$$

R-style formula: '$y \sim 1 + x_1 + \ldots + x_m + (1|\text{subject})$'

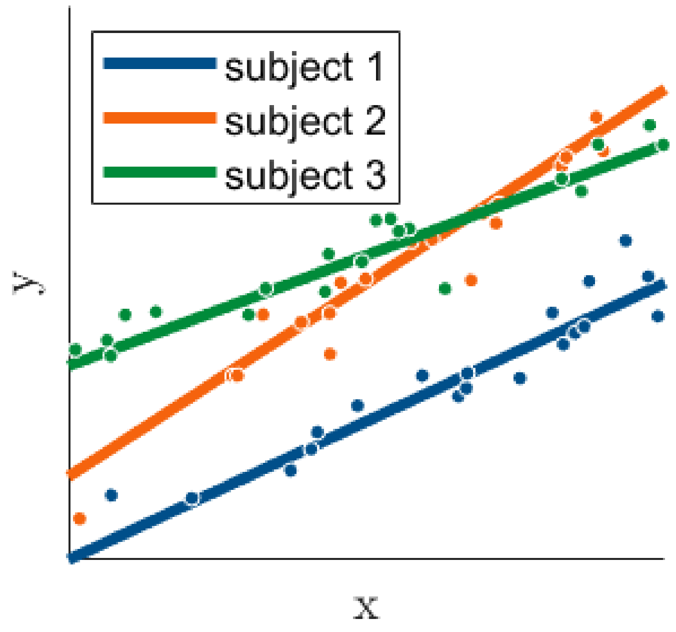


$$p(u) = \mathcal{N}(u; 0, R)$$

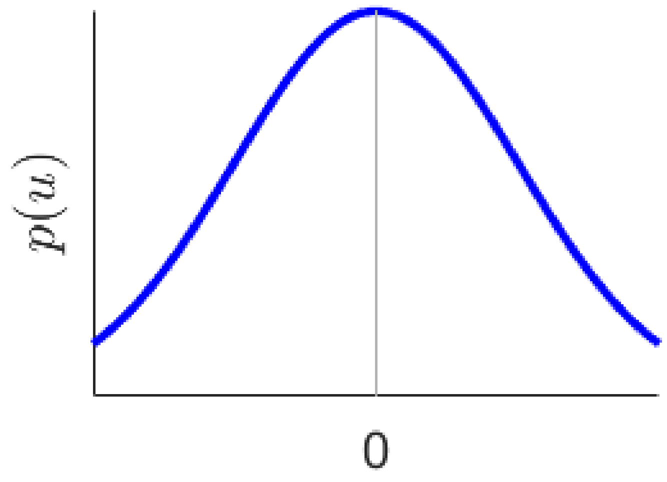


Random sensitivity mixed model

$$y_j = \beta_0 + \ldots + \beta_m x_m + u_{0j} + u_{1j} x_1 + \eta$$



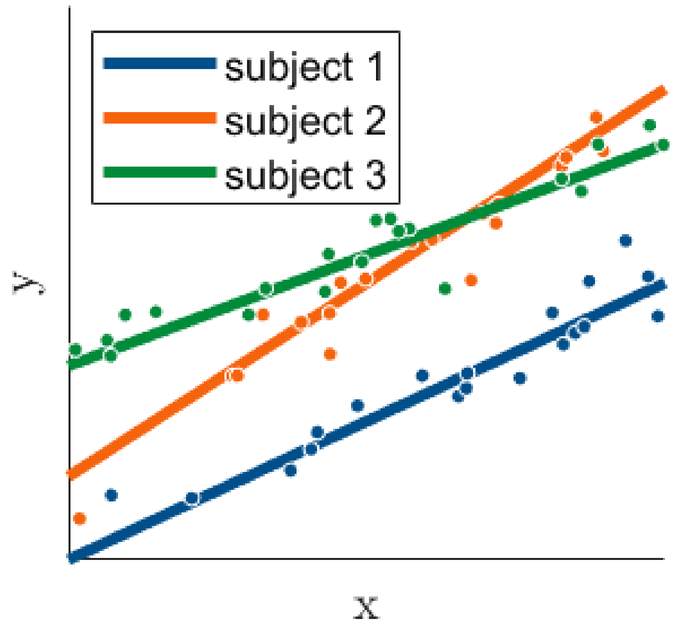
$$p(u) = \mathcal{N}(u; 0, R)$$



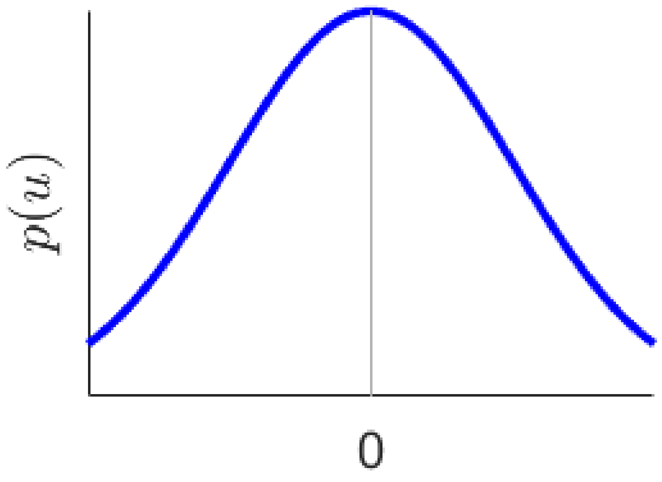
Random sensitivity mixed model

The intercept and sensitivity to x1 vary from subject-to-subject

$$y_j = \beta_0 + \ldots + \beta_m x_m + u_{0j} + u_{1j} x_1 + \eta$$

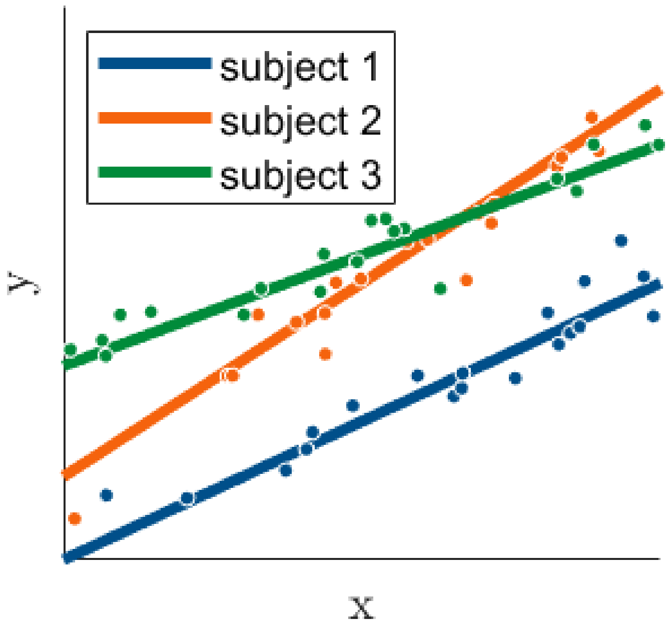


$$p(u) = \mathcal{N}(u; 0, R)$$

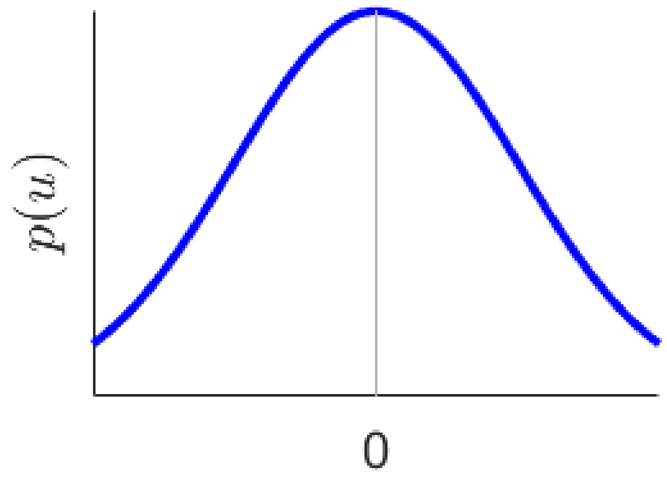


Random sensitivity mixed model

$$y_j = \beta_0 + \ldots + \beta_m x_m + u_{0j} + u_{1j} x_1 + \eta$$



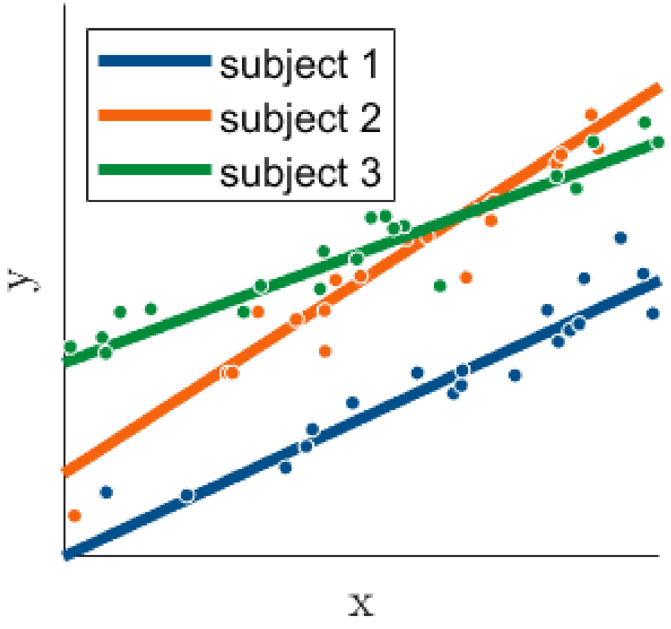
$$p(u) = \mathcal{N}(u; 0, R)$$



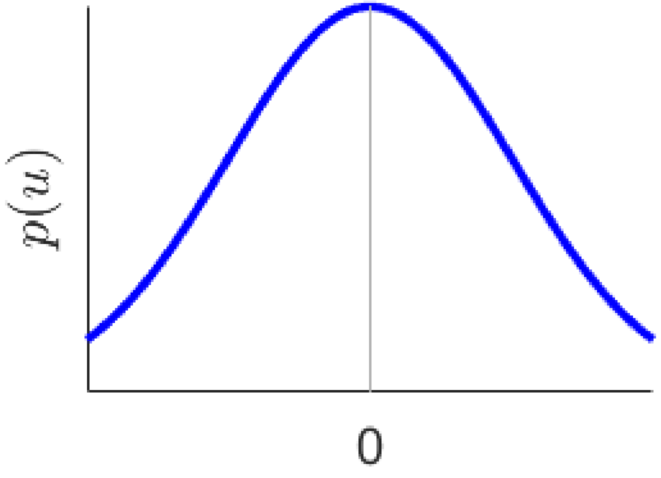
Random sensitivity mixed model

$$y_j = \beta_0 + \ldots + \beta_m x_m + u_{0j} + u_{1j} x_1 + \eta$$

R-style formula: $\text{'}y \sim 1 + x_1 + \ldots + x_m + (1 + x_1|\text{subject})\text{'}$



$$p(u) = \mathcal{N}(u; 0, R)$$

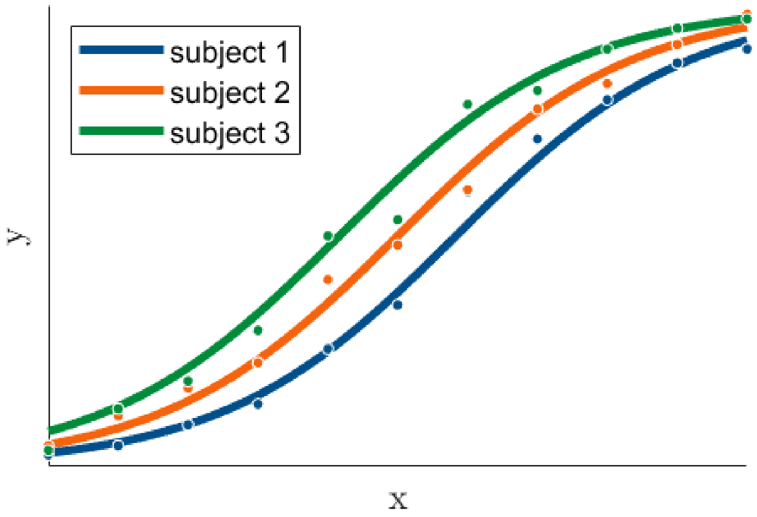


Linear Mixed Models (GLMM)

$$y_j = \underbrace{\beta_0 + \ldots + \beta_m x_m}_{\text{fixed effects}} + \underbrace{u_{j1} z_1 + \ldots + u_{jp} z_p}_{\text{random effects}} + \underbrace{\eta}_{\text{noise}}$$

fixed effects + random effects + noise

$$p(\boldsymbol{u}_j) = \mathcal{N}(\boldsymbol{u}; \boldsymbol{0}, \boldsymbol{R})$$



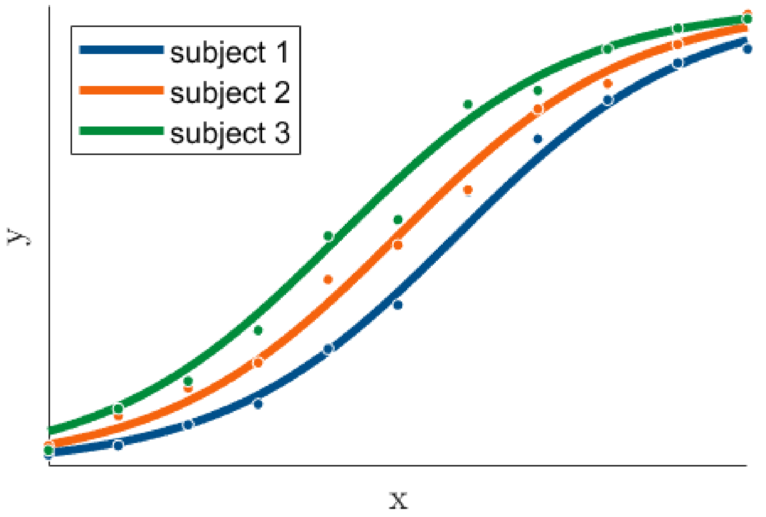
Generalized Linear Mixed Models (GLMM)

$$y_j = \underbrace{\beta_0 + \ldots + \beta_m x_m}_{\text{fixed effects}} + \underbrace{u_{j1} z_1 + \ldots + u_{jp} z_p}_{\text{random effects}} + \eta$$

fixed effects + random effects + noise

$$p(\boldsymbol{u}_j) = \mathcal{N}(\boldsymbol{u}; \boldsymbol{0}, \boldsymbol{R})$$

$$E(y_j) = \beta_0 + \ldots + \beta_m x_m + u_{j1} z_1 + \ldots + u_{jp} z_p$$



Generalized Linear Mixed Models (GLMM)

$$y_j = \underbrace{\beta_0 + \ldots + \beta_m x_m}_{\text{fixed effects}} + \underbrace{u_{j1} z_1 + \ldots + u_{jp} z_p}_{\text{random effects}} + \eta$$

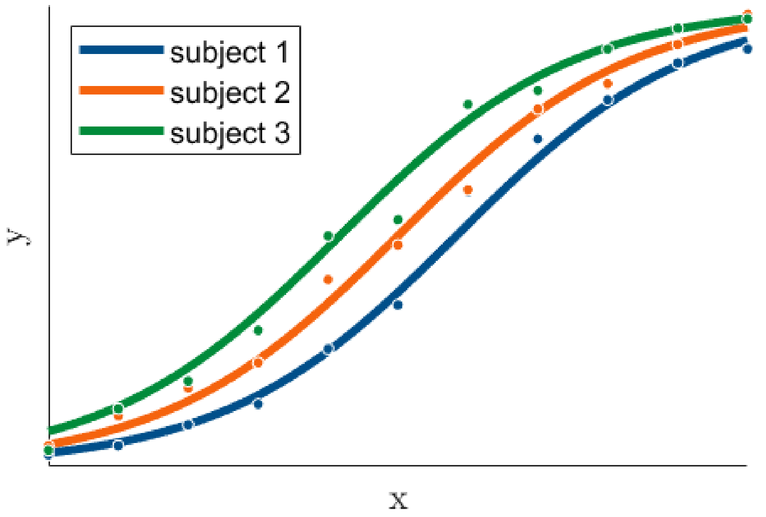
fixed effects + random effects + noise

$$p(\boldsymbol{u}_j) = \mathcal{N}(\boldsymbol{u}; \boldsymbol{0}, \boldsymbol{R})$$

$$E(y_j) = \beta_0 + \ldots + \beta_m x_m + u_{j1} z_1 + \ldots + u_{jp} z_p$$

e.g. binomial GLMM

$$p(y_j) = \sigma(\beta_0 + \ldots + \beta_m x_m + u_{j1} z_1 + \ldots + u_{jp} z_p)$$



Fitting (G)LMMs

By maximizing the likelihood over parameters (fixed effects beta, random effect covariance ***R***), marginalized over the latent variables (random effects)

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\beta}, \boldsymbol{R}) &= p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{R}) \\
&= \int_{\boldsymbol{u}} p(\boldsymbol{y}, \boldsymbol{u}|\boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{R}) d\boldsymbol{u} \\
&= \int_{\boldsymbol{u}} p(\boldsymbol{y}|\boldsymbol{u}, \boldsymbol{X}, \boldsymbol{\beta}) p(\boldsymbol{u}; \boldsymbol{R}) d\boldsymbol{u}
\end{aligned}$$

This protects us from overfitting to individual subject data.

Bates 2015, Journal of Statistical Software

Fitting (G)LMMs

By maximizing the likelihood over parameters (fixed effects beta, random effect covariance ***R***), marginalized over the latent variables (random effects)

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{\beta}, \boldsymbol{R}) &= p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{R}) \\
&= \int_{\boldsymbol{u}} p(\boldsymbol{y}, \boldsymbol{u}|\boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{R}) d\boldsymbol{u} \\
&= \int_{\boldsymbol{u}} p(\boldsymbol{y}|\boldsymbol{u}, \boldsymbol{X}, \boldsymbol{\beta}) p(\boldsymbol{u}; \boldsymbol{R}) d\boldsymbol{u}
\end{aligned}
$$

This protects us from overfitting to individual subject data.

This likelihood can be computed analytically in the case of LMM and then optimized (see Luigi's lecture). Otherwise we can apply the Expectation-Maximization algorithm.

Bates 2015, Journal of Statistical Software

Fitting (G)LMMs

By maximizing the likelihood over parameters (fixed effects beta, random effect covariance ***R***), marginalized over the latent variables (random effects)

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\beta}, \boldsymbol{R}) &= p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{R}) \\
&= \int_{\boldsymbol{u}} p(\boldsymbol{y}, \boldsymbol{u}|\boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{R}) d\boldsymbol{u} \\
&= \int_{\boldsymbol{u}} p(\boldsymbol{y}|\boldsymbol{u}, \boldsymbol{X}, \boldsymbol{\beta}) p(\boldsymbol{u}; \boldsymbol{R}) d\boldsymbol{u}
\end{aligned}$$

In practice, a package/function will take care of it

This protects us from overfitting to indi

This likelihood can be computed analytically in the case of LMM and then optimized (see Luigi's lecture). Otherwise we can apply the Expectation-Maximization algorithm.

Bates 2015, Journal of Statistical Software

Further reading

- Moscatelli, A., Mezzetti, M., & Lacquaniti, F. (2012). Modeling psychophysical data at the population-level: The generalized linear mixed model. *Journal of Vision*, *12*(11), 26–26. https://doi.org/10.1167/12.11.26
- Lo, S., & Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171. https://doi.org/10.3389/fpsyg.2015.01171
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01