Visual Inference for Graphcial Diagnostic of Linear Mixed Models
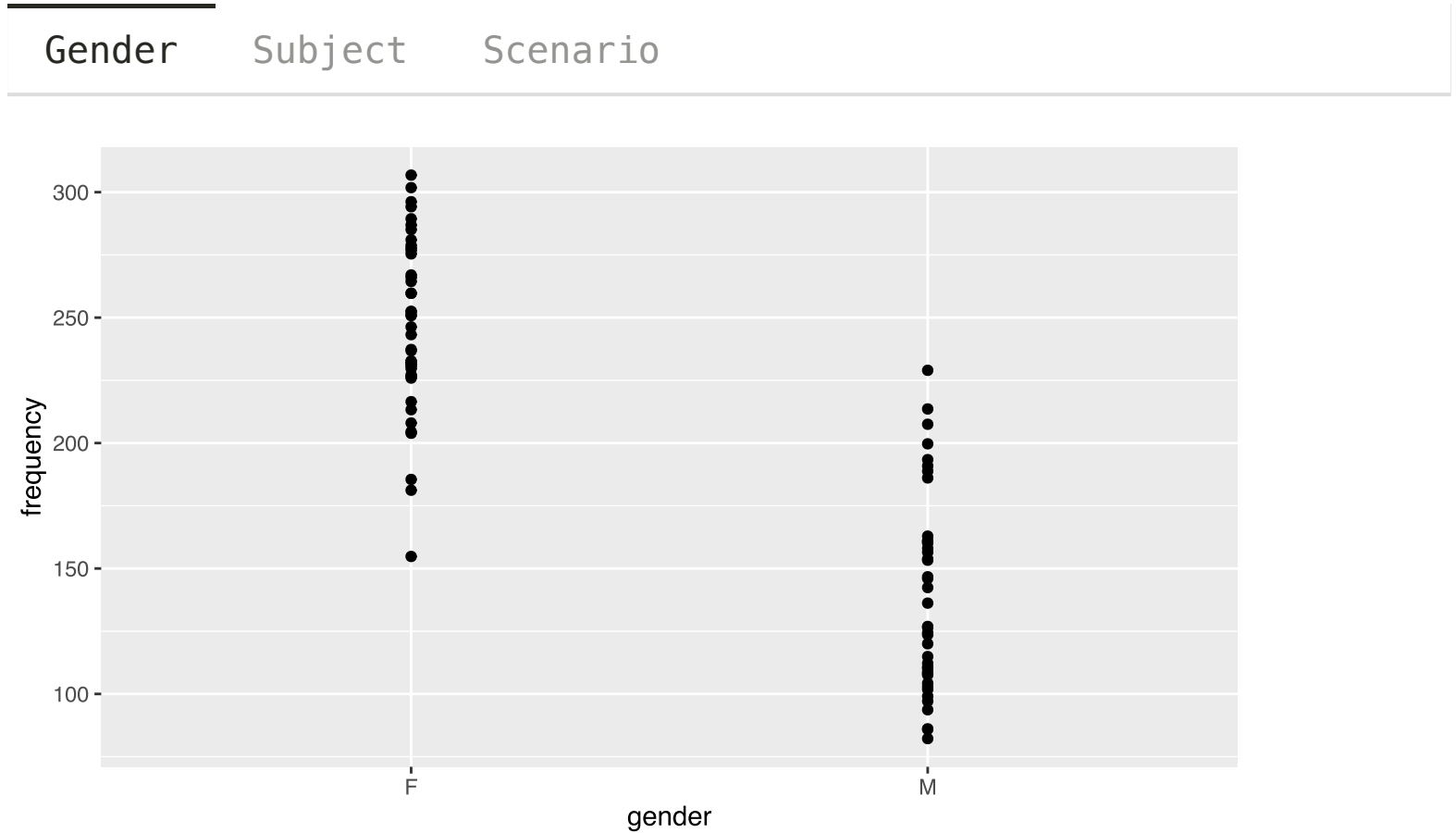
Kaiwen Jin

Monash University

2020-09-18

Linguistic Case

Linear Mixed Model

$$y = X\beta + Zb + e$$

where

- $\mathbf{y}$ is a $\mathbf{N} \times \mathbf{1}$ vector of observations, outcome variable
- $\mathbf{X}$ is a $\mathbf{N} \times \mathbf{p}$ matrix
- $\boldsymbol{\beta}$ is a $\mathbf{p} \times \mathbf{1}$ vector of the fixed effect
- $\mathbf{b}$ is a $\mathbf{N} \times \mathbf{q}$ matrix
- $\mathbf{b}$ is a $\mathbf{q} \times \mathbf{1}$ vector of the random effect
- $E[\mathbf{y}] = \boldsymbol{X\beta}$ and $V(\mathbf{y}) = \boldsymbol{\Omega} = \boldsymbol{Z\Gamma Z}^\top + \boldsymbol{R}$

How can we implement the LME?

- `lmer` function from `lme4` package
- `mmer` function from `sommer` package

Graphical diagnostic on residual analysis

- Types of residuals and corresponding residual diagnostic purpose:

  - Marginal residuals, $\hat{\boldsymbol{\xi}} = \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}$

    - Linear of the effects fixed
    - Presence of outlying observations
    - Within-units covariance matrix

  - Conditional residuals, $\hat{\boldsymbol{e}} = \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}} - \boldsymbol{Z}\hat{\boldsymbol{b}}$

    - Presence of outlying observations
    - Homoskedasticity of conditional errors
    - Normality of conditional errors

  - Random effect residuals, $\mathbf{Z}\hat{\mathbf{b}}$

    - Presence of outlying subjects
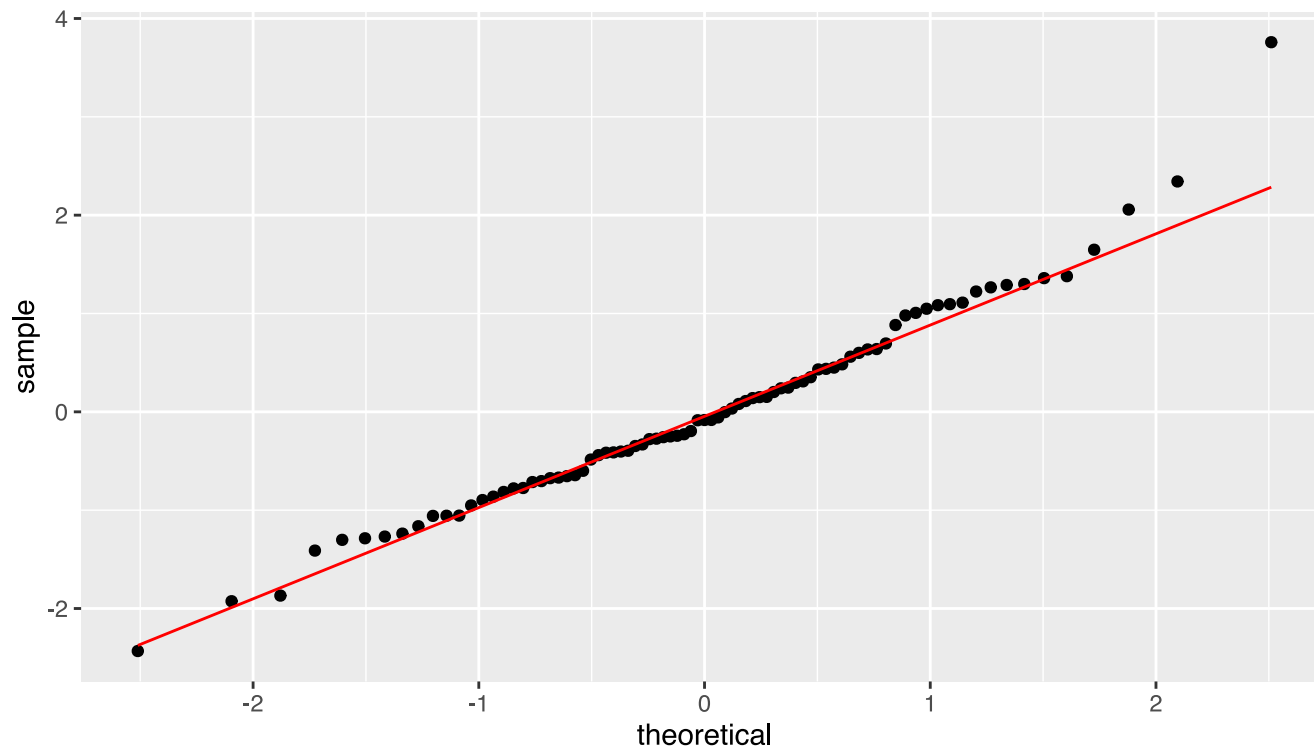   - Normality of the random effects

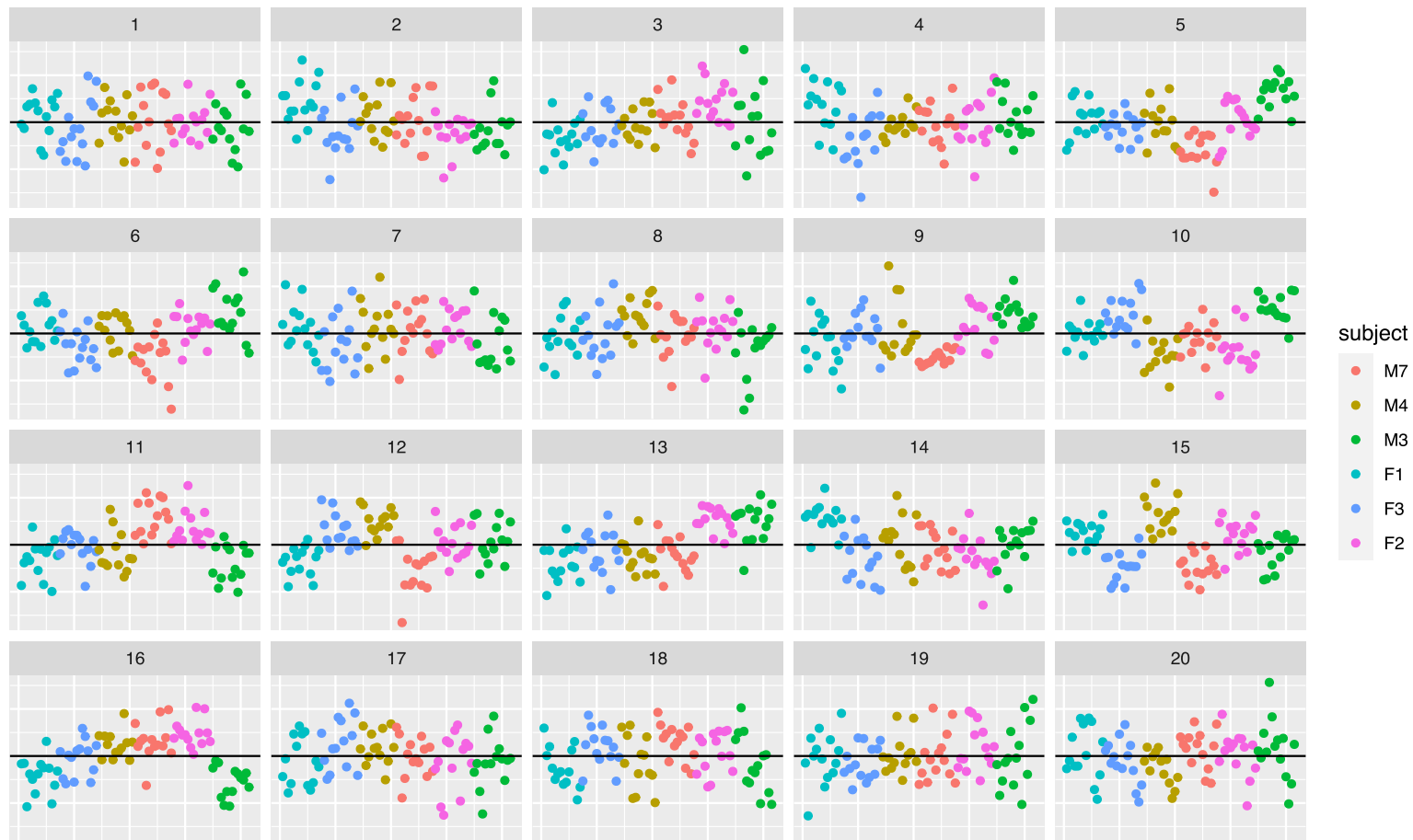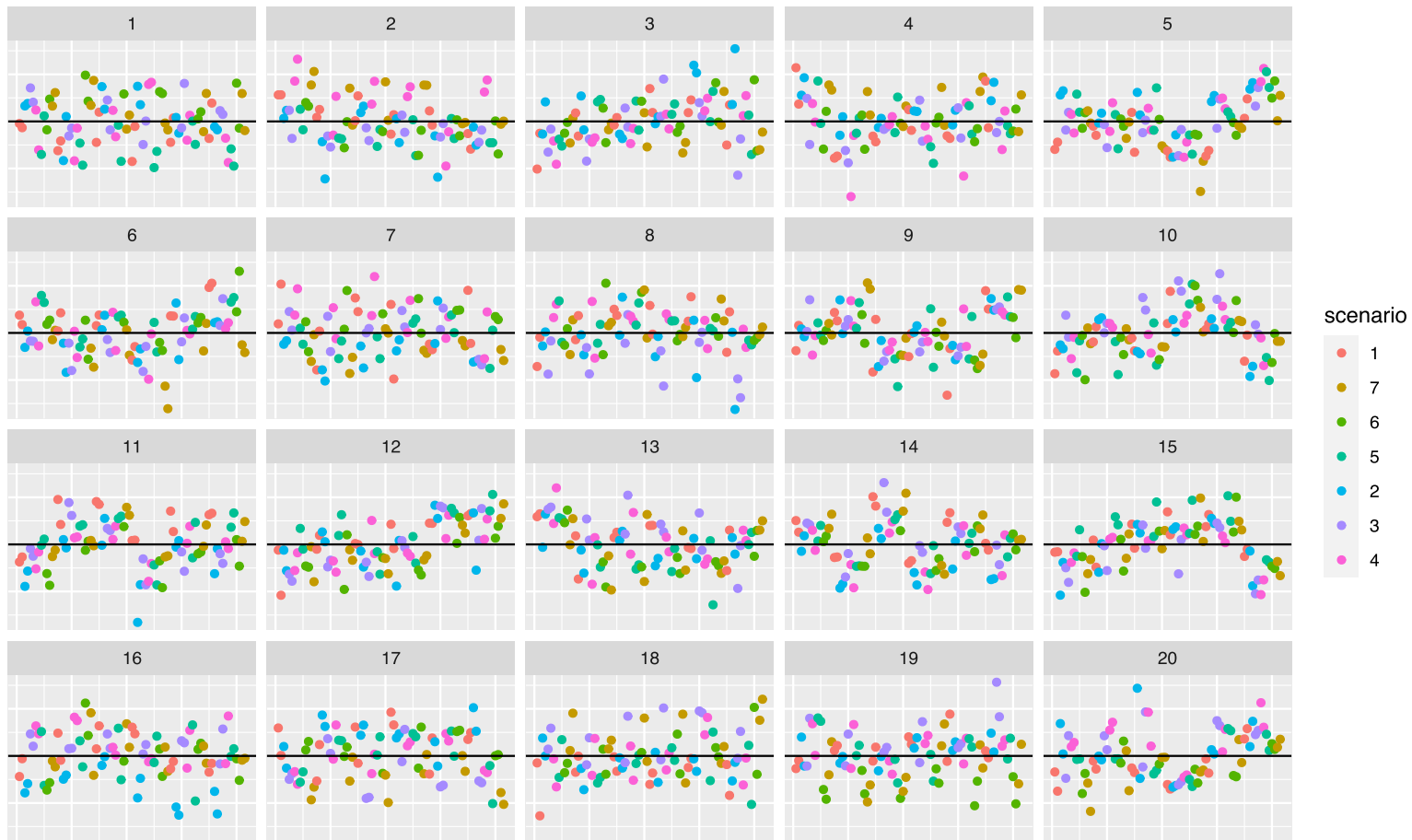# Linguistic Case

Presence of outlying observations

Subject   Scenario
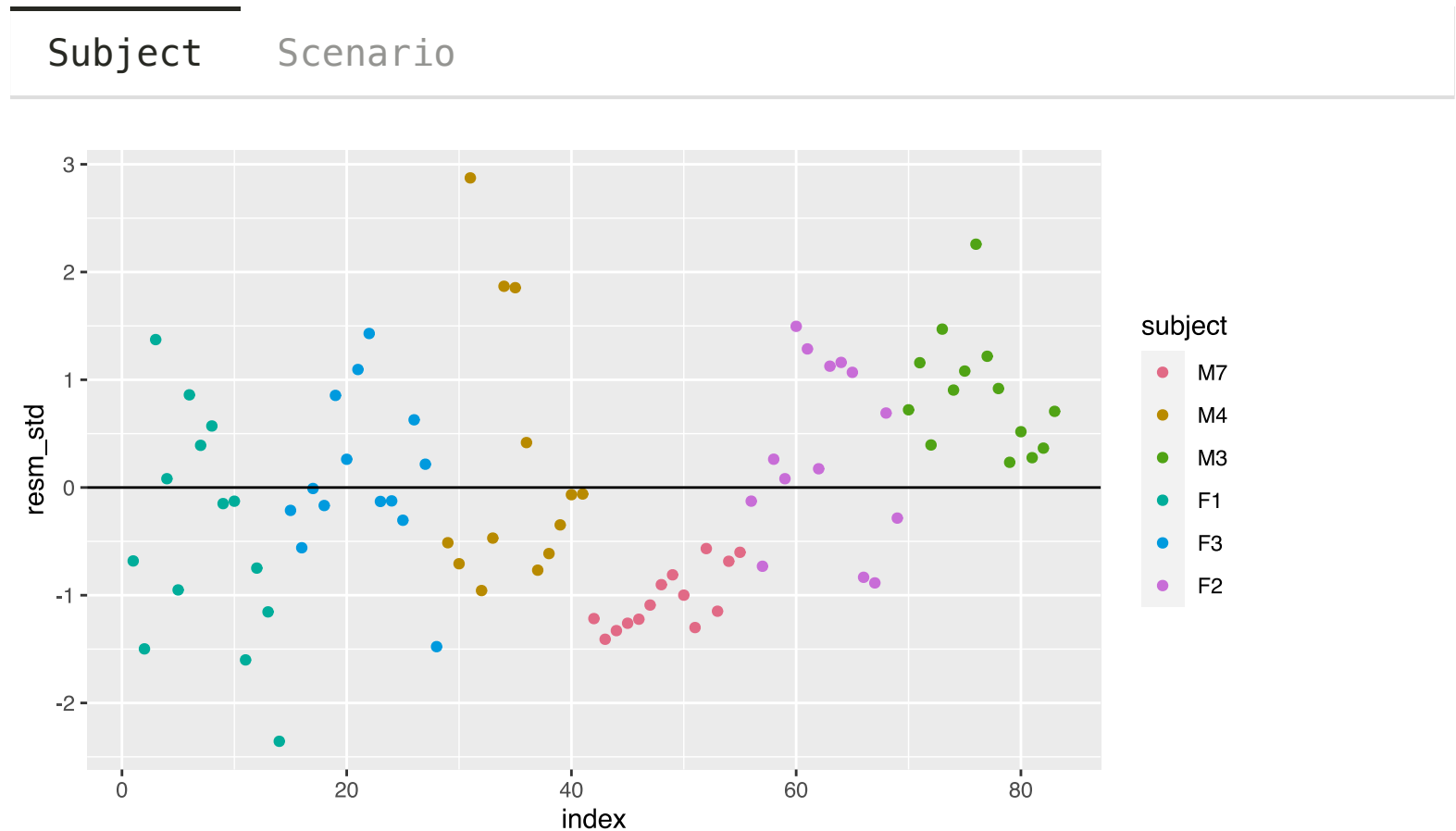
Linguistic Case

Normality of conditional errors

Lineups

Lineups

Lineups