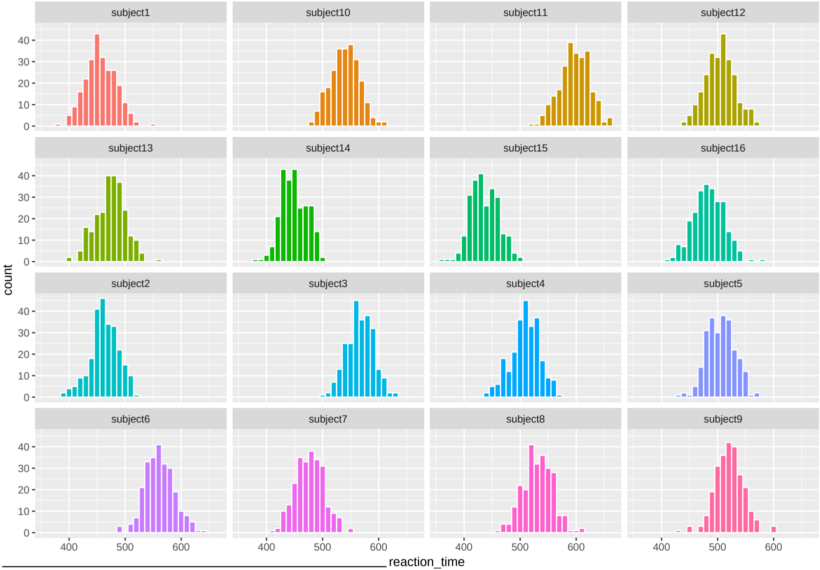
Multilevel Models

Mark Andrews
Psychology Department, Nottingham Trent University

mark.andrews@ntu.ac.uk

Multilevel data: Example 1

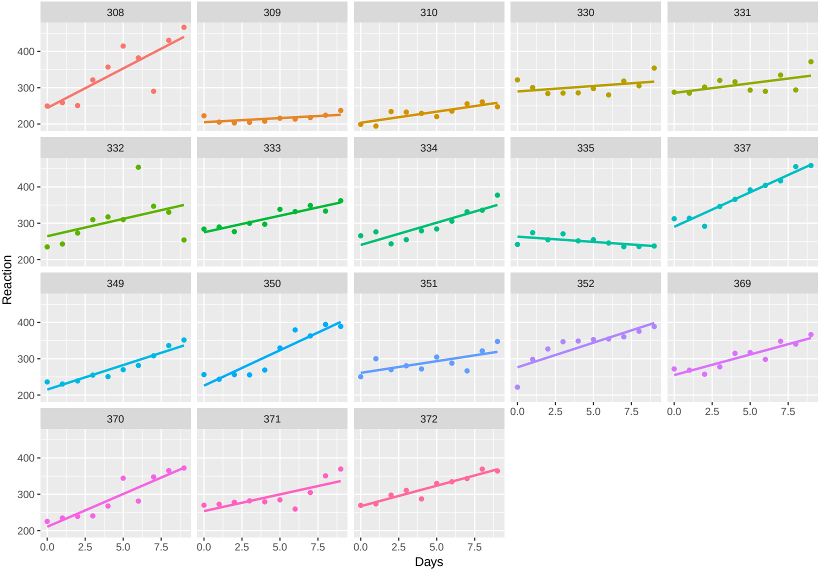
Reaction times[1] for 16 subjects.



[1] This is fake data. Real reaction times would not look so normal.

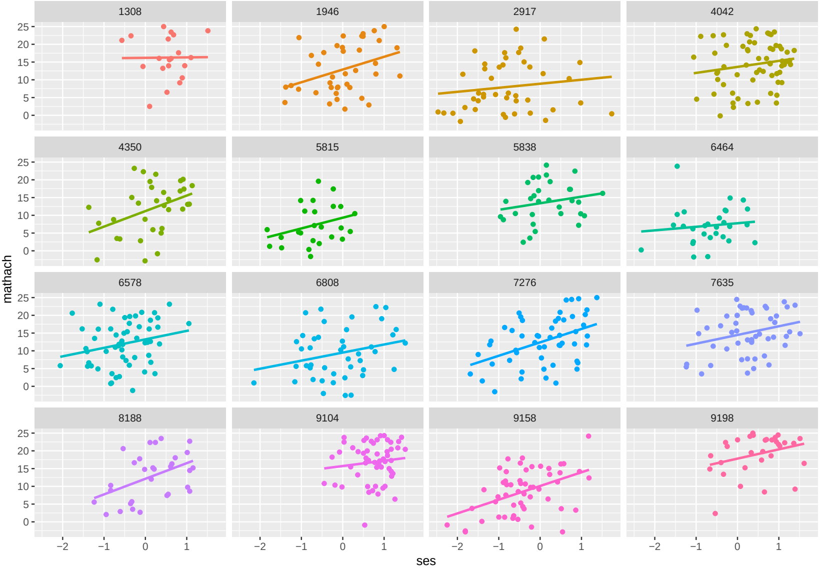
Multilevel data: Example 2

Reaction time as a function of sleep deprivation.



Multilevel data: Example 3

Mathematical achievement as function of socio-economic status.



Example: Multilevel model for reaction times

- Consider we have reaction time data from J subjects,

$$\{x_{j1}, x_{j2}, x_{j3} \ldots x_{jn_j}\}_{j=1}^{J}.$$

- A simple multilevel model for this data might be:

$$\begin{aligned} x_{ji} &\sim N(\mu_j, \sigma^2), \quad \text{for } i \in \{1 \ldots n_j\}, \\ \mu_j &\sim N(\theta, \tau^2), \quad \text{for } j \in \{1 \ldots J\}. \end{aligned}$$

- In words, each x_{ji} is drawn from a Gaussian with mean μ_j and variance σ^2, and each μ_j is drawn from a Gaussian with mean θ and variance τ^2.
- A Bayesian model will put a prior over θ and τ, and infer the posterior over θ, τ, $\mu_1 \ldots \mu_J$, σ^2.

Example: Multilevel model for reaction times

- We can re-write $x_{ji} \sim N(\mu_j, \sigma^2)$ as

$$x_{ji} = \mu_j + \epsilon_{ji}, \quad \epsilon_{ji} \sim N(0, \sigma^2).$$

- We can re-write $\mu_j \sim N(\theta, \tau^2)$ as

$$\mu_j = \theta + \eta_j, \quad \eta_j \sim N(0, \tau^2).$$

- The multilevel model can be re-written

$$x_{ji} = \theta + \eta_j + \epsilon_{ji} \quad \epsilon_{ji} \sim N(0, \sigma^2), \eta_j \sim N(0, \tau^2).$$

- This is often termed a *random-effects* model.

Example: Multilevel model for reaction times

- The variable θ denotes the global average reaction time.
- The variables $\mu_1 \dots \mu_j \dots \mu_J$ are the subjects's average reaction times.
- The variables $\eta_1 \dots \eta_j \dots \eta_J$ are the offsets of each subject's average reaction time from the global average. Each $\eta_j = \mu_j - \theta$.
- The variable σ^2 denotes the variance within any given subject.
- The variable τ^2 denotes the variance across subjects.

Example: Multilevel model for reaction times

- In the model just described, θ tells us the global average.
- The variance τ^2 tells us how much any given subject's average varies about θ.
- For example, 95% and 99% of the averages for individual subjects, will be in the ranges

$$\theta \pm 1.96 \times \tau, \quad \theta \pm 2.56 \times \tau,$$

  respectively.
- Likewise, 95% and 99% of any given subject's reaction times, i.e. x_{ji}, will be in the ranges

$$\theta + \eta_j \pm 1.96 \times \sigma, \quad \theta + \eta_j \pm 2.56 \times \sigma.$$

Example: Multiple drivers, multiple cars

- Let's say we want to measure the mpg of a given model of car (e.g. a Porsche 911).
- Because any one car could vary from others of the same model, we have K different examples of this model of car.
- Likewise, because any one driver could affect the recorded mpg of the car he drives, we have J different drivers.
- We get each of the J drivers to drive each of the K cars, and record the mpg as
$$y_{jk} = \text{mpg for driver } j, \text{ car } k.$$

Example: Multiple drivers, multiple cars

- A multilevel model for this mpg experiment could be

$$
\begin{aligned}
y_{jk} &\sim N(\mu_j + \nu_k, \sigma^2), \\
\mu_j &\sim N(\phi, \tau^2) \\
\nu_k &\sim N(\psi, \upsilon^2)
\end{aligned}
$$

which would work out as

$$
y_{jk} = \underbrace{\theta}_{\phi+\psi} + \eta_j + \zeta_k + \epsilon_{jk},
$$

with

$$
\eta_j \sim N(0, \tau^2),\ \zeta_k \sim N(0, \upsilon^2),\ \epsilon_{jk} \sim N(0, \sigma^2).
$$

Example: Multiple drivers, multiple cars

- In this example, we have three sources of variation

$$y_{jk} = \theta + \underbrace{\eta_j}_{\text{within driver}} + \underbrace{\zeta_k}_{\text{within car}} + \underbrace{\epsilon_{jk}}_{\text{within trial}},$$

  where τ^2 gives the within driver variance, υ^2 gives the within car variation, and σ^2 gives within trial variation.

- The variable θ provides the average mpg for the car model (i.e. the Porsche 911)
- The variables τ^2, υ^2 and σ^2 provide measures of the relative variation across in mpg drivers, cars and trials, respectively.

Example: Reaction time and math achievement

- In this problem, we have J subject. For subject j, we have n_j data points.
- In observation i from subject j, their number of days without sleep is x_{ji} and the reaction time is y_{ji}.
- A multilevel model for this data is

$$\begin{aligned} y_{ji} &\sim N(\alpha_j + \beta_j x_{ji}, \sigma^2), \\ \alpha_j &\sim N(a, \tau_a^2), \\ \beta_j &\sim N(b, \tau_b^2). \end{aligned}$$

Example: Reaction time and math achievement

- The model

$$
\begin{aligned}
y_{ji} &\sim N(\alpha_j + \beta_j x_{ji}, \sigma^2),\\
\alpha_j &\sim N(a, \tau_a^2),\\
\beta_j &\sim N(b, \tau_b^2),
\end{aligned}
$$

can be re-written

$$y_{ji} = \underbrace{(a + \eta_j)}_{\alpha_j} + \underbrace{(b + \zeta_j)}_{\beta_j} x_{ji} + \epsilon_{ji},$$

or

$$y_{ji} = \underbrace{a + b x_{ji}}_{\text{Fixed effect}} + \underbrace{\eta_j + \zeta_j x_{ji}}_{\text{Random effect}} + \epsilon_{ji},$$

where

$$\eta_j \sim N(0, \tau_a^2),\ \zeta_j \sim N(0, \tau_b^2),\ \epsilon_j \sim N(0, \sigma^2).$$

Example: Reaction time and math achievement

- In the model just described, a and b are the general regression coefficients.
- The variance τ_a^2 tells us how much variation in the intercept term there is across schools. The variance τ_b^2 tells us how much variation in the slope term there is across schools.
- For example, 95% and 99% of the intercepts for individual schools will be in the ranges

$$a \pm 1.96 \times \tau_a, \quad a \pm 2.56 \times \tau_a,$$

respectively. Likewise, 95% and 99% of the slope terms for schools will be in the ranges

$$b \pm 1.96 \times \tau_b, \quad b \pm 2.56 \times \tau_b.$$