NCRM National Centre for Research Methods

Multilevel Models

Session 2: Random intercept models

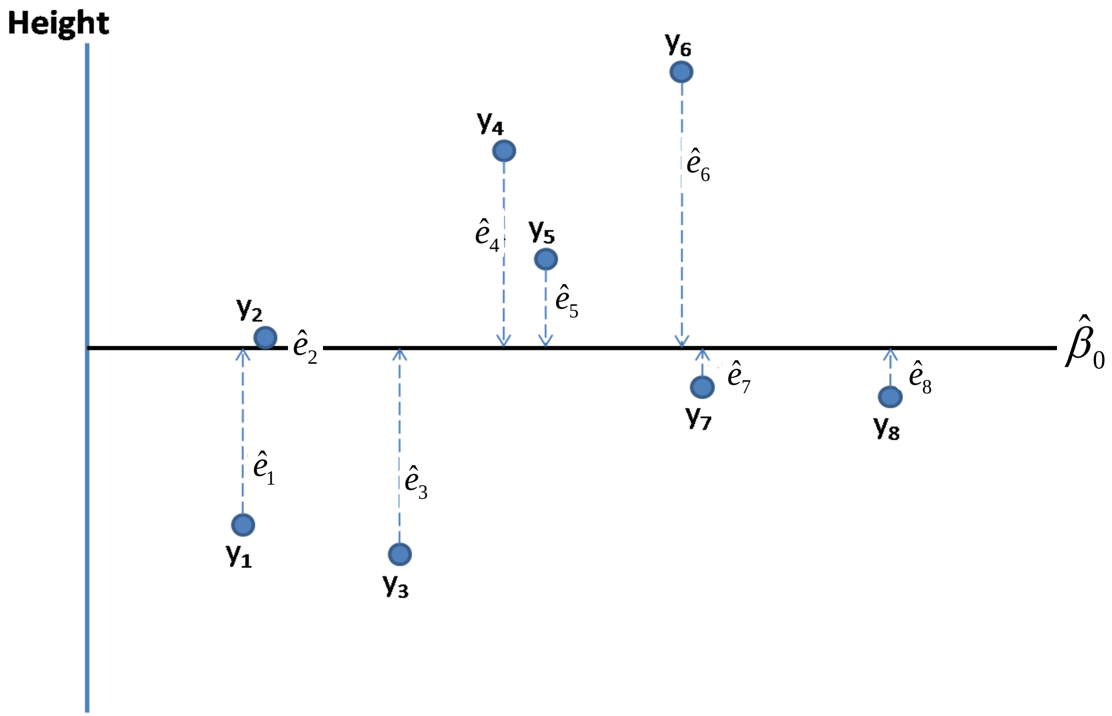
Outline

- Two level random intercept models
  - Comparing groups – the variance components model
  - Quantifying group differences – the variance partition coefficient
  - Adding predictors at the individual and group level – the random intercept model

Two level random intercept models for continuous data

- Simplest form of multilevel models in wide use
- Extends standard linear regression models by partitioning the residual error between individual and group components
- But assumes same relationship between x and y across groups
- Can provide an initial assessment of importance of groups (when no explanatory variables are included)

Single-level model for mean height

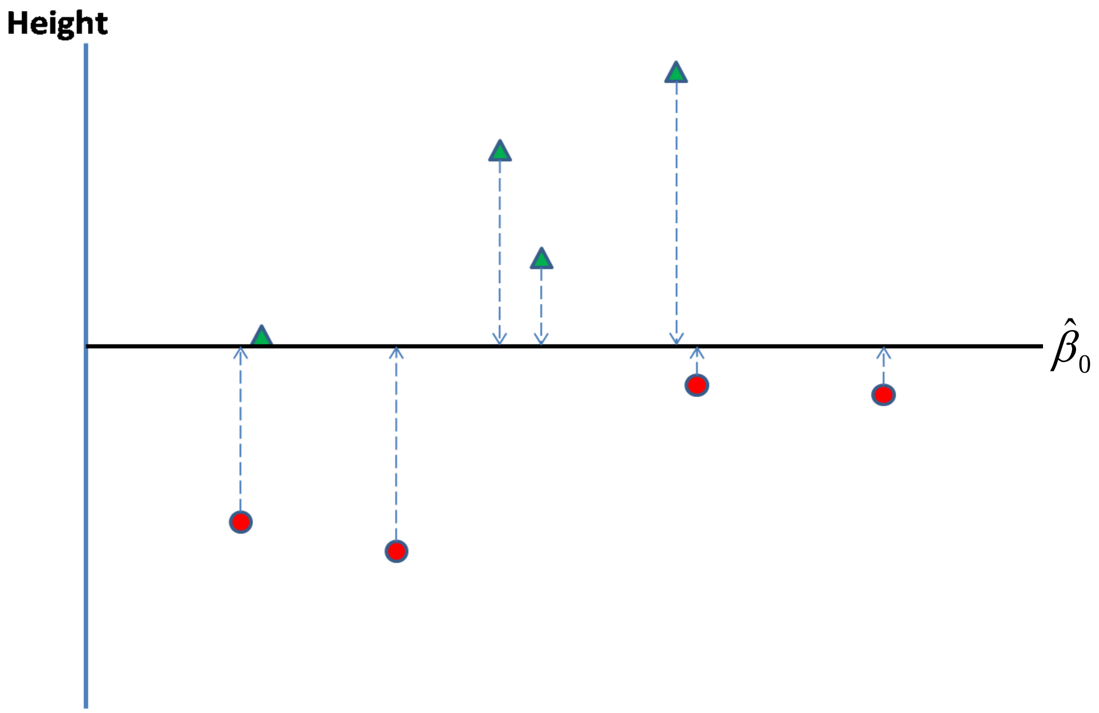


$$y_i = \beta_0 + e_i$$

$$e_i \sim N(0, \sigma_e^2)$$

- y_i = the height for the i[th] individual
- β_0 is mean height in population and e_i is residual for i[th] individual (i=1,2...,n)
- Assume e_i are approximately normal with mean 0. The variance summarises distribution around the mean.

Single-level model for mean height

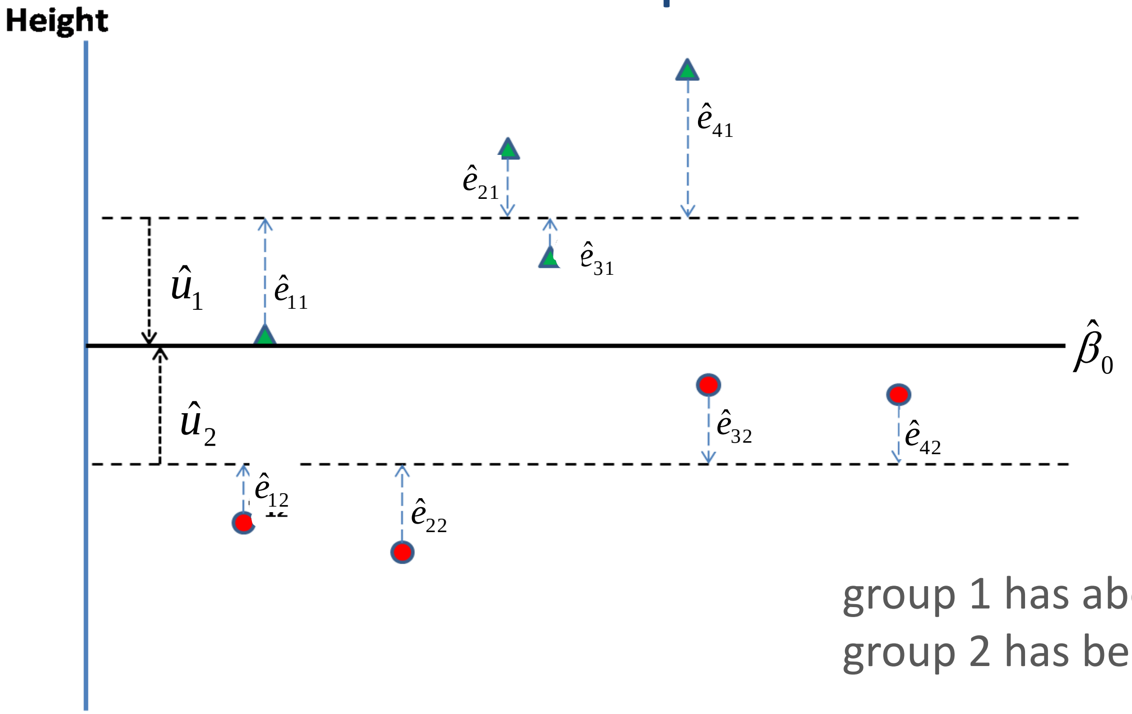


- But suppose we know our observations come from different groups (e.g. families), j=1,…, J

  shown here are two groups (in practice, there will be many more)

- We can capitalise on this additional information and improve our model

Multi-level 'empty' model for mean height: variance components model



$$y_{ij} = (\beta_0 + u_j) + e_{ij}$$

$$u_j \sim N(0, \sigma_u^2)$$

$$e_{ij} \sim N(0, \sigma_e^2)$$

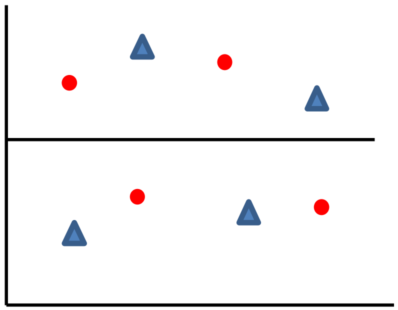
group 1 has above-average mean (positive u)
group 2 has below-average mean (negative u)

- y_{ij} - height for the i^{th} individual in j^{th} group (1,2,…n).
- β_0 - average height across all groups
- u_j - group mean deviations from overall mean height
- e_{ij} - individual deviations from group means
- $\beta_0 + u_j$ - average height in group j

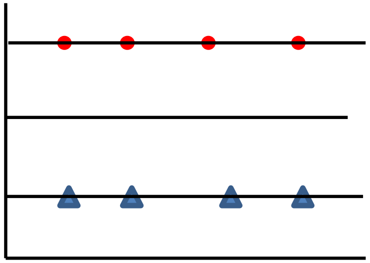
Variance Partition Coefficient

$$VPC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

- VPC tells us how important group level differences are (e.g. what proportion of variance is at the group level?)



- VPC = 0 if no group effect $\sigma_u^2 = 0$



- VPC =1 if no within group differences $\sigma_e^2 = 0$

Example: Fear of Crime across neighbourhoods

Fear of crime: higher scores mean more fear

- 27,764 individuals, nested in 3,390 areas
- Mean of 8 residents per area (1-47)

VARIANCE COMPONENTS MODEL

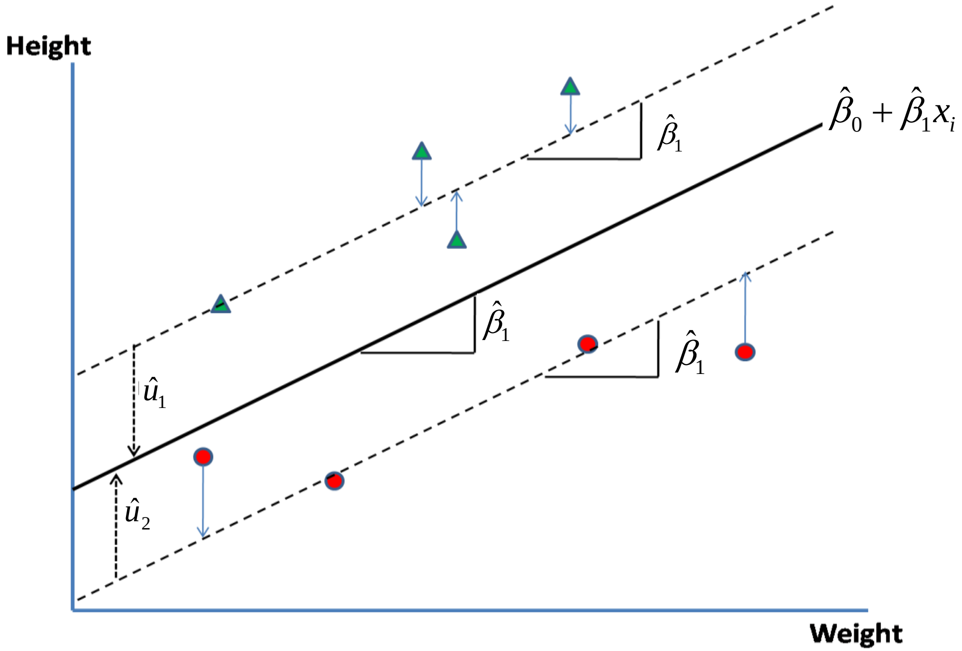
		MODEL 1
FIXED PART		
Intercept		0.027 (.009)
RANDOM PART		
σ_e^2	Individual variance	0.863 (.008)
σ_u^2	Neighbourhood variance	0.145 (.007)

Crime Survey for England and Wales, 2013/14

$$VPC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

Neighbourhood contribution = .145/(.863+.145) = 14.4%

Adding an explanatory variable: A random intercept model



$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij}$$

$$u_j \sim N(0, \sigma_u^2)$$

$$e_{ij} \sim N(0, \sigma_e^2)$$

- Overall relationship between weight and height across families is represented by intercept β_0 and slope β_1 (fixed part)
- For group j, the intercept is $\beta_0 + u_j$ (either above or below average)
- Individual deviations from group line e_{ij} and group deviations from average line u_j (random part, with means 0 and variances σ_e^2 and σ_u^2)

Group level explanatory variables

- Multilevel models enable us to explore group level variables **simultaneously** with individual
- Can be from external sources (administrative data etc), or aggregates of individual data (depending on group size)
- No need to directly identify them as group effects, this is accounted for by the group residual
- Standard errors generally underestimated if included in individual level analysis

Example: Fear of Crime across neighbourhoods

RANDOM INTERCEPT MODEL		Crime Survey for England and Wales, 2013/14
	MODEL 1	MODEL 2
FIXED PART		
Intercept	0.027 (.009)	-.005 (.009)
x_{1ij} Age (in years)		-.004 (.001)
x_{2ij} Victim in last 12 months		.248 (.014)
x_{3j} Crime Rate		.227 (.012)
RANDOM PART		
σ_e^2 Individual variance	0.863 (.008)	.850 (.008)
σ_u^2 Neighbourhood variance	0.145 (.007)	.105 (.006)

- Individual level R^2: (.863 - .850)/.863 = .015
- Neighbourhood level R^2 = (.145-.105)/.145 = .276

Summary

- In this session we have introduced the variance components model and the random intercept model
- The **variance components** model can be used to provide an initial estimate of the contribution of groups
- The **random intercept** model allows us to include explanatory variables at the individual and group level to explain variation in our dependent variable

National Centre for Research Methods

For more information visit
www.ncrm.ac.uk