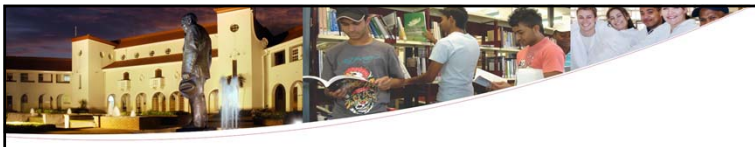


Multilevel models

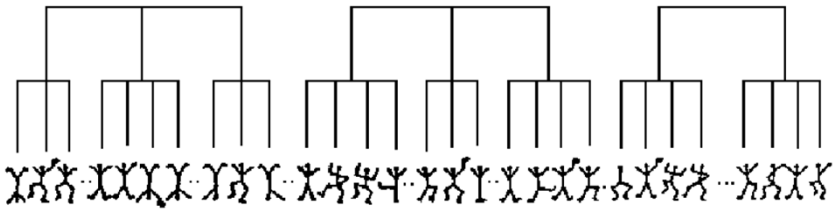
Marike Cockeran
MURIA symposium
25 July 2016

It all starts here

NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT
POTCHEFSTROOM CAMPUS

Objectives

- Recognise when there is a multilevel problem.
- Basic understanding of multilevel analysis.

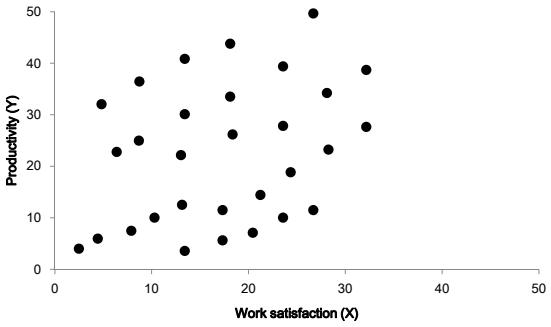


Background: Single-level model

- Simple linear regression is a technique that is used to explore the nature of the relationship between two continuous random variables.
- Regression analysis enables us to investigate the change in one variable, called the response (dependent variable), which corresponds to a given change in the other, known as the explanatory variable (independent variable).
- The ultimate objective of regression analysis is to predict or estimate the value of the response that is associated with a fixed value of the explanatory variable.

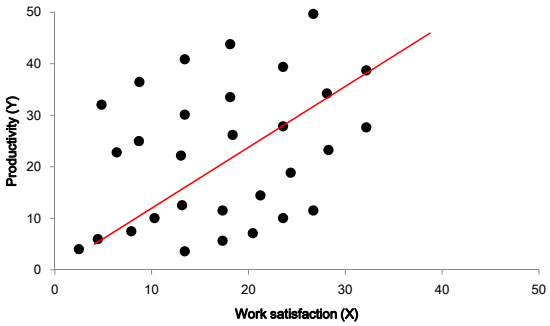
Example: Single-level model

- Is there a relationship between work satisfaction and productivity?



Example: Single-level model

- Is there a relationship between work satisfaction and productivity?



$$\text{Productivity}_i = \beta_0 + \beta_1 \text{Work satisfaction}_i + e_i$$

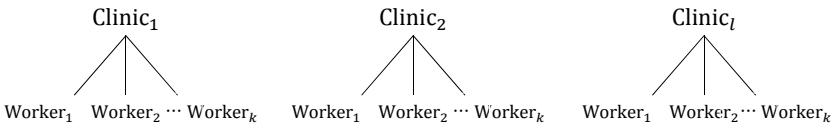
Example: Single-level model

$$\text{Productivity}_i = \beta_0 + \beta_1 \text{Work satisfaction}_i + e_i$$

- β_0 - Intercept.
  - The predicted (mean) productivity when work satisfaction is zero.
- β_1 - Slope.
  - The value by which predicted productivity changes for each unit of work satisfaction.
- e_i - Residual.
  - The part of productivity not accounted for by the model.
- β_0 and β_1 are fixed parameters. One value for the intercept and one value for the slope are calculated. All participants receive the same intercept and slope.

Example: Multilevel model

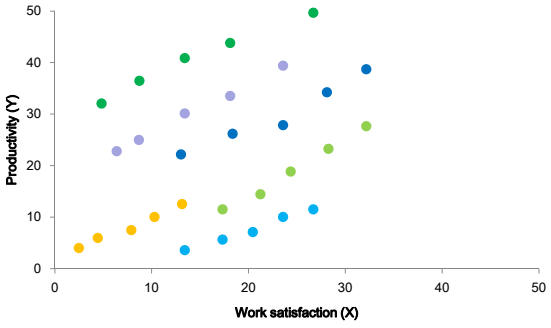
- Now suppose groups of participants work at the same clinics.
- Participants are nested within clinics.
- Two-level sample:
  - Level 1: The individual participants.
  - Level 2: The clinics.



This structure needs to be built into the model.

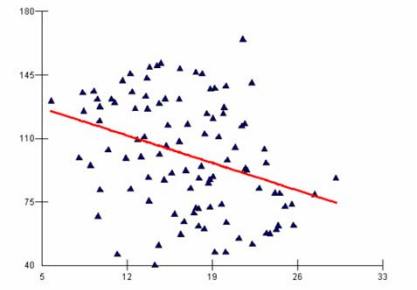
Example: Multilevel model

- Is it possible to determine how much of the variance in productivity might be attributed to the clinic and how much might be attributed to the individual participant?

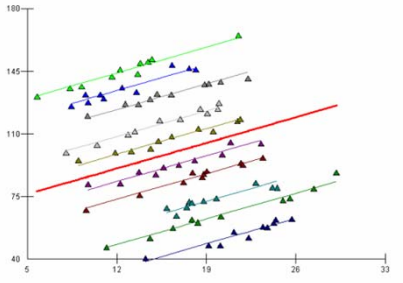


Example: Multilevel modelling

- Single-level models can be misleading:



Treating data as single-level model



Treating data as multilevel model

Consequence of clustering

- Measurements on units within a cluster are more similar than measurements on units in different clusters.
- Examples:
  - Two doctors selected at random from the same clinic are expected to respond more similarly than two doctors randomly selected from different clinics.
  - Repeated measurements on same participant
- Multilevel models are also known as:
  - Random-effects models
  - Hierarchical models
  - Variance-components models
  - Random-coefficient models
  - Mixed models

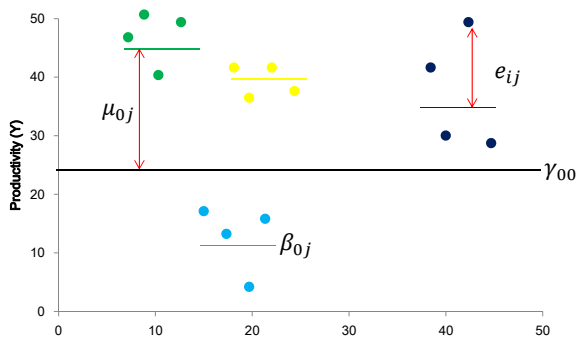
Steps: Multilevel models

- Step 1: Intercept only model
  - To what extent does productivity vary between clinics?
- Step 2: Add fixed effects
  - Add level 1 predictors
    - Do work satisfaction predict productivity?
  - Add level 2 predictors
    - Does clinic size predict productivity?
- Step 3: Evaluate random slope variance
  - Does the relationship between work satisfaction and productivity vary across clinics?

Step 1: Intercept only model

- To what extent does productivity vary between clinics?

$$Y_{ij} = \gamma_{00} + \mu_{0j} + e_{ij} \qquad Y_{ij} = \beta_{0j} + e_{ij}$$



Step 1: Intercept only model

- Calculate the intra-class / cluster-correlation coefficient

$$\rho = \frac{{\sigma^2}_{between}}{{\sigma^2}_{between} + {\sigma^2}_{within}}$$

Estimates of Covariance Parameters[a]

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		1.221793	0.039641	30.821	.000	1.146517	1.302012
Intercept [subject = Clinic]	Variance	0.702105	0.108620	6.464	.001	0.518461	0.950797

a. Dependent Variable: Productivity.

Step 2: Add predictors

- Add level 1 predictors (independent variables)
  - Work satisfaction
- Add level 2 predictors (independent variables)
  - Size of clinic
- Suppose work satisfaction is added as a level 1 predictor:

$$\text{Productivity}_{ij} = \beta_{0j} + \beta_{1j}\text{Work satisfaction}_{ij} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

Step 3: Random slope variance

- Does the relationship between work satisfaction and productivity vary across clinics?

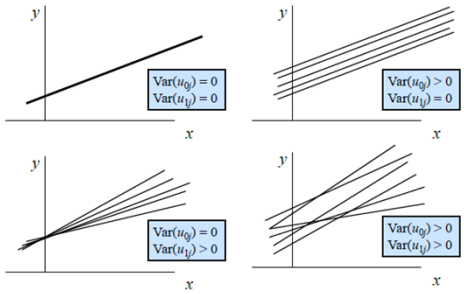
$$\text{Productivity}_{ij} = \beta_{0j} + \beta_{1j}\text{Work satisfaction}_{ij} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \mu_{0j}$$

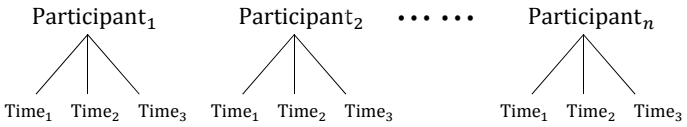
Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	258.608696	16.108847	22.929	16.054	.000	225.279258	291.938133
Work satisfaction	25.898522	3.575625	109.000	7.243	.000	18.811750	32.985293
Clinic size	18.014522	3.575625	109.000	5.038	.000	10.927750	25.101293

Multilevel models

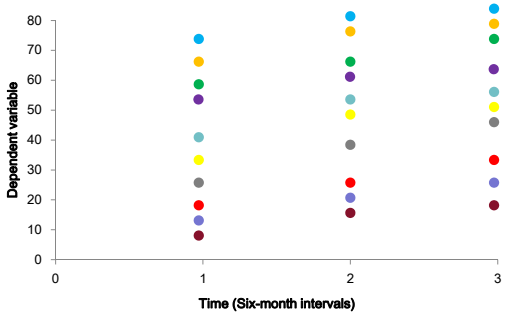


Longitudinal studies

- Multilevel data structures also arise in longitudinal studies where an individual's responses over time are correlated with each other.
- The clusters are composed of the repeated measurements obtained from a single individual at different occasions.
- In longitudinal studies the level 1 units are the repeated occasions of measurement and the level 2 units are the subjects.



Longitudinal studies



- At the first time point, before participants enter into the study there is already a lot of variability in the values of the dependent variable. This suggests that intercepts should differ across participants.
- There is also a lot of variables at all subsequent time points. This suggests that the slope of the relationship between time and the dependent variable might vary across participants also.

Format of dataset

- **Wide format**

Subject	SBP_1	SBP_2	SBP_3
1	115	118	117
2	120	125	115
3	130	132	128
⋮			

- **Long format**

Subject	Time	BP
1	1	115
1	2	118
1	3	117
2	1	120
2	2	125
2	3	115
⋮		