Week 8

# ~Week 8 : Multilevel Models~

## ① Admin:

- Week 8 final day. Well done for completing this term...etc
- Python notebooks → 1 more notebook on MLM. Worth doing if you plan on doing MLM but no requirement if not. You have better things to be doing
- Regression problem set → Opinions? Again, if you haven't done it then don't worry but hopefully it was interesting for those of you that did it

⊛ Reminder that the papers are on GitHub. Had really good feedback from people so far. They are a high level but that is what you can aim for. A few of them are 75+

→ If you are thinking about impressive models to use the MLM is a really good one to think about. eg. if you were doing a MLR and want to take it further.

## ② MLM Thoughts:

- Stress again that MLM is a really good way to elevate your paper.
- Some of my thoughts on how hard it would be to use

⊙ Conceptually it is hard: Idea and notation well... harder...
⊙ Mathematically it is relatively easy
⊙ Implementation via python is easy

⊛ But big gains

→ Impressiveness to difficulty ratio is high so a good one to choose ⚠

## ③ Resources:

- Fox textbook is your friend: p. 699 →
  710-712 (23.3.2) → Variance Component Model
  712-714 (23.3.3) → Random Intercept (+Random Slope) Model
  * This is quite meaty but well worth reading. I read through yesterday and it is really good if you want to use it
- Adam's LN
- Bristol Videos → Short lectures and well worth learning
- Online resources are hard to find and quite bad

MLM goes by so many other names
- → Random Effects Models (generic term)
- → Variance and covariance component models
- → Hierarchical Linear Models
- → Contextual Effects Models
- → Random Coefficient models
- → Repeated-measures models

} Listed in Fox

## ④ Background: Fixed vs Random Effects

→ Understand this and everything else will be easy.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i$$

, $X_{1i}$ is a dummy

$X_{1i}$ is a fixed effect
- The effect applies to everyone in the population
- Generalize from the sample
- Generally "fixed effects" refer to subgroup dummies, often in panel data (within-estimator)

Random effect
- Random! $\varepsilon_i \sim N(0, \sigma^2_\varepsilon)$
- Residuals assumed to be random
- Not a population wide thing.
- Models variability
- Do not generalize from the sample
- Random Variable → coefficients are mean and variance

# Split Screen

* Two ways of controlling for groups: Fixed effects and random effect

(FE:) Introduce dummies (Categorical Variables)

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \varepsilon_i$$

Use J and J−1

→ Control for the groups by adding p−1 dummies
→ Can estimate the coefficient for each group
→ Hopefully these would generalize to the population
⚠ (*) Generally "fixed effects" refers to subgroup dummies in panel data models → [illegible]

OR... (this is where the conceptual jump comes in) -- random effects

(RE:) $Y_i = \beta_0 + \varepsilon_i$ } Simple regression model

→ Error term now: $\varepsilon_{ij} = u_j + e_{ij}$

$u_j$: Group random variation
$e_{ij}$: individual random error (as before)

$$Y_i = \beta_0 + u_j + e_{ij}$$

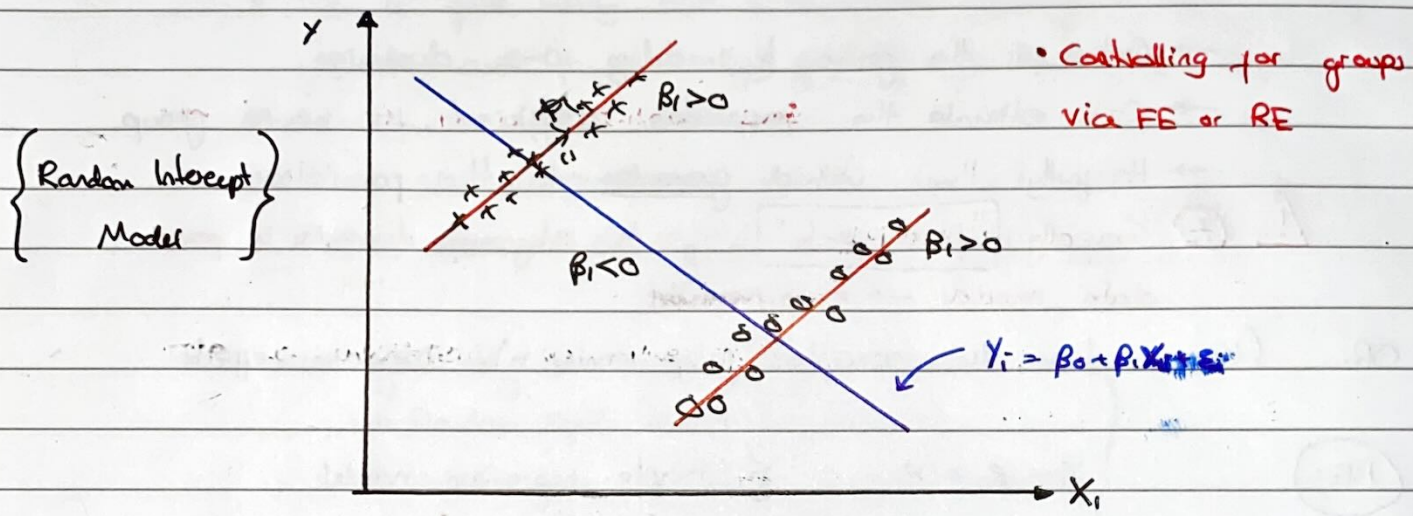
→ Assuming that differences in groups is because of random variation (error) rather than fixed difference.
→ If we were to sample again we might get diffent random group variation. ASSUMPTION

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

→ This will control for the group differences

(5) Reminder of the motivation for this + when you might use each model.

Ecological Fallacy: (Confounding Effects)




{Random Intercept Model}

• Controlling for groups via FE or RE

→ Clearly important to control for the diffent groups but when do we assume group effects are FE vs RE?

**FE**

- Small number of groups
- Care about estimating the group effects
- Groups generalise to population in a "meaningful" way
- Distribution of group effects are not ≈ normal

→ eg: Ethnicity, gender, education level (if categorical)

**RE**

- Lots of groups (using FE is not feasible as lots of coefficient need estimating) → Degrees of freedom problems
- Don't care about estimating the effects of specific groups
- Don't really expect specific groups to have an impact (we think group is important but the importance is kind of random)
- Categories might change if we sample again

→ eg. Schools if sampling 100 schools, hospital ID, – etc
Neighbourhood for house price, ...

⚠ Both controlling for group effects via FE and RE will overcome the ecological fallacy. BUT used in diffent cases ⚠

→ Estimation of coefficient in model ($\hat{\beta}_1$) will be diffent so important that you get the choice right

— DON'T GO PAST HERE UNLESS THEY GET IT —

↳ May just finish here if doing other things...

⑤ Variance Component Model:

* Very very simple
* Motivation is not the ecological fallacy but estimating proportion of variance
* Is multilevel modeling ever necessary?

→ Empty regression:

$$Y_i = \beta_0 + \varepsilon_i \quad \} \quad \hat{\beta}_0 = \text{mean } Y_i$$

But Group Effects! $Y_i = \beta_0 + u_j + e_{ij}$

→ Estimate $\hat{\beta}_0$, $\sigma^2_u$, $\sigma^2_e$ ($\hat{\beta}_0$ the same as in just model)

Reason we might do the VC model...

---

$\rho$ = Variance Partition Coefficient = $\dfrac{\sigma^2_u}{\sigma^2_u + \sigma^2_e}$ ⚠

---

→ Proportion of variance in $Y_i$ due to group effects

→ Kind of interesting. Diffent type of research question being asked

→ But not that interesting...

## (7) Random Intercept Model:

MLR: $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{1i} + \ldots + \beta_k X_{ki} + \varepsilon_i$

- But group effects! Confounding effects and $\beta$s potentially biased (think about the ecological fallacy)

- Model the groups as random effects

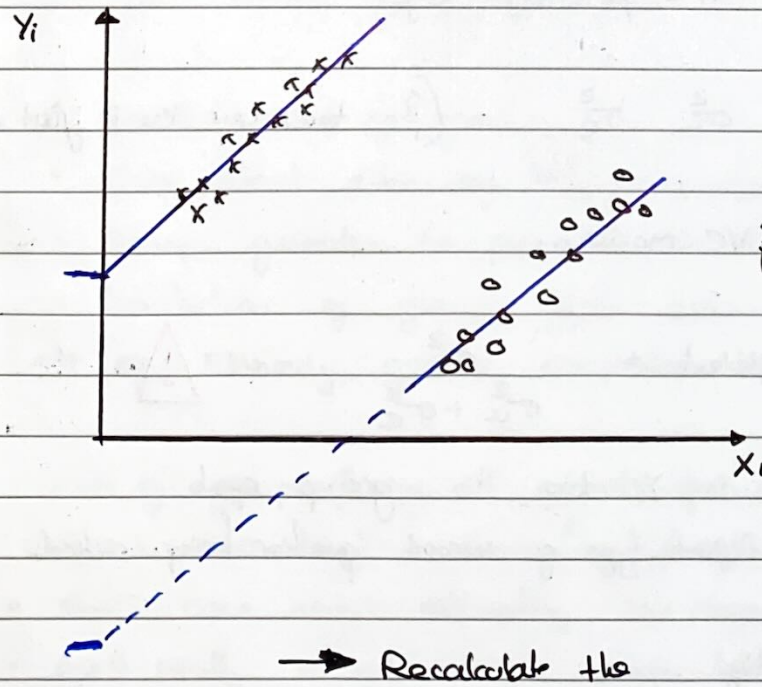
$$Y_i = \beta_0 + \beta_1 X_{1i} + \ldots + \beta_k X_{ki} + u_j + e_{ij}$$

→ Basically lets the intercept shift for each group (intercept is random)

→ $\hat{\beta}_0$ is the average intercept.

→ $\hat{\beta}_1, \ldots, \hat{\beta}_k$ are the average within-group coefficients

Nice. Overcomes the ecological fallacy.




$\hat{\beta}_1$ a much better estimator

→ Recalculate the variance partition coefficient, $\rho$

$$\frac{\sigma^2_u}{\sigma^2_u + \sigma^2_e}$$

↳ will change as you control for things!

## (8) The Point Of MLM

→ "Nested", "Hierarchical" data (groups)

and...

(1) Assumption mean FE is not appropriate. This could well be the case

or

(2) Interested in estimating the variance partitioning coefficient, ρ

## (9) Extensions:

→ Random Slope + Random Intercept (See Fox)

→ Hypothesis testing for whether you should account for [redacted] regressors (Likelihood ratio test)

Intuition:
- 2 models (with and without additional regressors)
- Calculate log-likelihood (likelihood of seeing the data given the model)
- Compare difference in likelihood → if sufficiently large jump then significant

→ [redacted]

→ Pseudo $R^2$ point

$$2(\ell(1) - \ell(0))$$

$\ell(1)$: with addition; $\ell(0)$: without

QUESTIONS