Mixed model analysis case studies

Rasmus Waagepetersen

April 24, 2023

Mixed model analysis of randomized studies:

- evaluation of new mathematics teaching method in primary school
- comparison of whole grain vs. refined grain diets

TRACK (Teaching Routines and Content Knowledge) project

123 Danish primary schools randomized into treatment and control group.

Treatment: new mathematics teaching method inspired by Singaporean practice.

Follow pupils over three years starting with 4th grade.

Data available now: mathematics test result at beginning of study (baseline).

My colleagues from Aarhus wanted to test adequate randomization by assessing treatment effect at baseline.

Is this a good idea ?

OLS

They used ordinary least squares (OLS) but with adjusted standard errors taking into account correlation between schools and classes.

$$\hat{\beta} = (X^\top X)^{-1} X^\top Y \quad \mathbb{V}\mathrm{ar}\hat{\beta} = (X^\top X)^{-1} X^\top V X (X^\top X)^{-1}$$

Plug-in empirical estimate of V.

Note: this is not BLUE !

They found slightly significant negative treatment effect !! (p slightly less than 5%)

Conclusion ??

Mixed models analysis

I used linear mixed models with school and class random effects.

In this case

$$\hat{\beta} = (X^\mathsf{T} V^{-1} X)^{-1} X^\mathsf{T} V^{-1} Y \quad \mathbb{V}\mathrm{ar}\hat{\beta} = (X^\mathsf{T} V^{-1} X)^{-1}$$

Is this BLUE ?

Possible disadvantage of mixed model analysis ?

Results regarding fraction (brøk-regning) test results

Estimates of treatment effects and p-values:

OLS	OLS p-value	permutation p-value
-2.39	0	0.014
Mixed (BLUE)	approx t p-value	permutation
-1.34	0.151	0.145

Permutation: randomly permute schools into treatment and control and assess treatment effect for each permuted data set. Under null-hypothesis, data should come from this permutation distribution.

Whole grain (WG) vs. refined grain (RG)

Subjects randomly allocated to two treatment arms:

Group 1:	baseline	WG	RG
Group 2:	baseline	RG	WG

Outcome: LDL cholesterol in blood

Note: possible cross over effect (treatment effect WG-RG may depend on order of treament (WG first or last)

For ith subject three measurements Y_{it}, $t = 1, 2, 3$

Standard approach: regression using baseline Y_{1t} as covariate:

$$Y_{it} = \mu_{it} + \alpha Y_{i1} + \epsilon_{it}, \quad t = 2, 3$$

μ_{it}: two sided ANOVA based on Group (1, 2) and Treatment (WG, RG)

Problem: we need to skip all observations for i if baseline is missing !

Alternative: mixed model with subject random effect

$$Y_{it} = \mu_{it} + U_{it} + \epsilon_{it}, \quad t = 1, 2, 3$$

Specification of μ_{it} more complicated since Treatment now has three levels WG, RG and baseline. Due to randomization, no group effect for baseline !

Results: no cross over effect, WG good for reducing LDL :)