Index of Lecture 10–L

— in addition to overheads from Davis (2002):
597–611,613–14,618,624,630,632,637–641

Practical Information

Today's lecture:

- review:
  - * random effects models (mostly by examples),
  - * GEE estimation, part I:
    - classical method (sometimes called "GEE1"),
    - some methodological issues postponed to next week,
- examples:
  - * cross-over trials with two and three periods (Diggle et al., 2002),
  - * seizure counts (Davis, 2002, and Diggle et al., 2002),
  - * clinical trial on spasmodic torticollis $\sim$ neurological disease (Davis, 2002).

Plan for rest of course:

- final of GEE + transitional models next week (Thursday!),
- summary of approaches + software + worked example,
- missing data,
- Mabrok's seminar.

Analysis of Random Effect Models

Many approaches for GLMM analysis:

- conditional (VER: "stratification"):
  - consider the random effects as nuisance parameters, and eliminate by "conditioning" (statistical technique) $\Rightarrow$ uses only within- cluster information,
  - binary outcome: Mantel-Haenszel statistic and conditinal logistic regression,
  - no assumptions about random effects,
  - limited scope: no between-group predictors,
- quasi-likelihood approximations:
  - numerically simple and fast,
  - biased (towards 0) estimates in certain situations,
  - no likelihood-based analysis,
- numerical integration:
  - numerically demanding and not always numerically robust,
  - `gllamm` macro in Stata = state-of-the-art software,
- Bayesian analysis (all above are "frequentist").

Examples: Cross-over Trials

2×2 trial:

- 2 treatments (A=active, B=placebo) and 2 periods ⇒ 2 treatment orders: AB or BA,
- outcome = electrocardiogram abnormal/normal (0/1),
- data:

	Responses				
Group	(1,1)	(0,1)	(1,0)	(0,0)	Total
AB	22	0	6	6	34
BA	18	4	2	9	33

3×3 trial:

- 3 treatments (A=placebo, B/C=low/high analgesic) and 3 periods ⇒ 6 treatment orders (below), corresponding to two Latin squares,
- outcome = relief (no/yes) of primary dysmenorrhoea,
- data:

	Responses								
Group	000	100	010	001	110	101	011	111	Total
ABC	0	0	2	2	1	0	9	1	15
ACB	2	1	0	0	0	0	9	4	16
BAC	0	1	1	1	0	8	3	1	15
BCA	0	1	1	1	8	0	0	1	12
CAB	3	0	0	0	1	7	2	1	14
CBA	1	5	0	0	4	3	1	0	14

Results for 2×2 Cross-over Trial

Random effects/GEE logistic regression estimates:

	Estimate (SE) on logit scale		
Parameter	Conditional	Num. Integration	GEE (classical)
intercept	–	1.52 (.86)	0.52 (.25)
β(tx)	1.61 (.77)	1.60 (.77)	0.56 (.24)
clustering	–	$\hat{\sigma} = 4.37$	$\tilde{\rho} = 0.62$
intercept	–	2.17 (1.10)	0.67 (.29)
β(tx)	10.9 (.)	1.84 (0.90)	0.57 (.23)
β(period)	-10.3 (.)	-1.02 (0.80)	-0.30 (.23)
clustering	–	$\hat{\sigma} = 4.84$	$\tilde{\rho} = 0.64$

- – numerical integration by `SAS,proc nlmixed`
- – $\tilde{\rho}$ = working correlation coefficient

Comments:

- agreement between two random effects analyses (first model),
- conditional estimation fails due to sparse data (12/67 subjects),
- very strong clustering,
- approximate correction from SS to PA by factor $\sqrt{1 + 0.346 \cdot 4.84^2} = 3.0 \Rightarrow$ fair agreement,
- no tx×period interaction (not shown).

Results for 3×3 Cross-over Trial

Random effects logistic regression estimates (logit scale):

Parameter		Estimate (SE) Conditional	Num. Integration	GEE (classical)
intercept		–	-1.09 (.33)	-1.08 (.31)
β(tx)	B	1.98 (.45)	2.11 (.40)	2.10 (.42)
β(tx)	C	1.71 (.41)	2.07 (.38)	2.07 (.42)
β(per)	2	0.69 (.56)	0.41 (.46)	0.42 (.41)
β(per)	3	0.85 (.58)	0.59 (.48)	0.59 (.46)
β(prev)	B	-0.14 (.60)	-0.13 (.50)	-0.14 (.51)
β(prev)	C	-1.24 (.65)	-0.93 (.49)	-0.93 (.44)
clustering		–	$\hat{\sigma} = 0.013$	$\tilde{\rho} = -0.03$

- numerical integration by (crude) `glmmML` library in R
- $\tilde{\rho}$ = working correlation coefficient

Comments:

- β(prev) $\sim$ carry-over effects, e.g. $\hat{\beta}_c = -0.93 \sim$ reduction in odds of relief in period following C (relative to A) by factor $e^{-0.93} = 0.40$,
- close agreement between numerical integration and GEE approaches, but some differences to conditional,
- no substantial clustering,
- no treatment by period interactions (not shown).

Example: Seizure Counts

Clinical trial involving 59 epileptics:

- two treatment groups (progabide medicament and placebo),
- counts of seizures in four 2-week intervals, plus 8-week baseline interval prior to treatment,
- additional information on age.

Data analysis preliminaries:

- outcome (counts) strongly non-normal $\Rightarrow$
  - summary statistic approach (Ch. 2 of Davis, 2002),
  - GLMM, possibly a Poisson regression:
    - length of weekly intervals $\sim$ "offset" (time at risk),
    - dichotomous predictor (time) to separate baseline and study periods:
      time=0 for baseline, and =1 for study period.
- patient no. 207 very unusual (very high counts), but cannot be excluded as an error,
- seizure rates (per 2 weeks):

Treatment	Time	Seizure rate
Progabide	0	7.90
	1	7.96
Placebo	0	7.70
	1	8.60

  — weak treatment effect (if any).

Results for Seizure Counts

Random effects/GEE Poisson regression estimates:

	Estimate (SE) on log scale			
Parameter	Conditional	Num. Integration		GEE (classical)
intercept	–	1.03 (.15)	1.07 (.14)	1.35 (.16)
β(tx)	–	-0.02 (.21)	0.05 (.19)	0.03 (.22)
β(time)	0.11 (.047)	0.11 (.047)	-0.002 (.11)	0.11 (.12)
β(tx∗time)	-0.10 (.065)	-0.10 (.065)	-0.31 (.15)	-0.10 (.21)
clustering	–	$\hat{\sigma} = 0.78$	$\hat{\sigma}_0 = 0.71$	$\tilde{\rho} = 0.78$
	–	–	$\hat{\sigma}_{\text{time}} = 0.48$	–
	–	–	$\hat{\rho} = 0.17$	–

- – conditional by explicit formula calculations
- – numerical integration by `SAS,proc nlmixed`
- – $\tilde{\rho}$ = working correlation coefficient

Comments:

- fair agreement between methods (except for random effects and GEE intercepts),
- strong evidence of random slopes:
  $G^2 = 2021.4 - 1849.3 = 172.1$ (df=2),
- log-scale variances of random slope model:

  $$\text{time} = 0: \ \sigma^2 = 0.71^2 = 0.50,$$
  $$\text{time} = 1: \ \sigma^2 = 0.71^2 + 0.48^2 + 2 \cdot 0.71 \cdot 0.48 \cdot 0.16 = 0.84,$$

- smaller correlation between baseline count and study week count than between two study week counts.