

Comparing GLMM and GEE

Demetris Athienitis

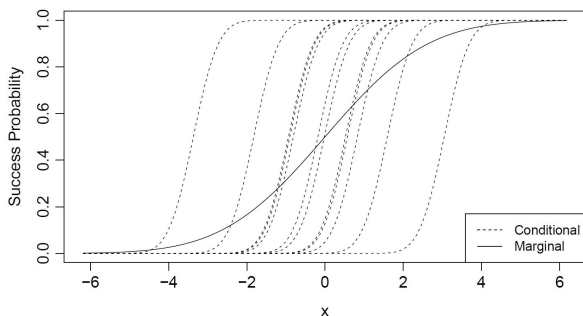


Recall that the two methods covered that can accomodate correlated responses are:

- GLMM is a subject specific model, modeling how the predictors affect an individual subject's response
- GEE is a marginal model, modeling how the predictors affect the average of the subject's response

Comparison

- When $\hat{\sigma}_u = 0$ in GLMM or $\hat{\rho} = 0$ in GEE, estimates and standard errors same as treating repeated observations as independent
- When $\hat{\sigma}_u$ is large, estimated β 's from random effects logit model usually larger than from marginal GEE model. They are estimating different things



Example (Teratology Overdispersions)

Female rats on iron-deficient diets assigned to four groups:

- 1 placebo
- 2 iron injections on days 7 and 10
- 3 iron injections on days 0 and 7
- 4 iron injections weekly

Then they are made pregnant and sacrificed after 3 weeks. The response is whether fetus is dead or alive and the *cluster* is the litter.

Notation:

- GRP = group,
- LS = litter size,
- ND = number dead in litter

Example (continued)

$$\text{logit}[P(\text{fetus } t \text{ in litter } i \text{ dead})] = \alpha + \beta_2 z_{i2} + \beta_3 z_{i3} + \beta_4 z_{i4}$$

$$z_{ij} = \begin{cases} 1 & \text{if litter } i \text{ in group } j \\ 0 & \text{otherwise} \end{cases}$$

We will fit regular binomial GLM, i.e. assuming independence.

```
> terat.binom=glm(cbind(ND,N-ND)~GRP, family=binomial)
> summary(terat.binom)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.1440	0.1292	8.855	< 2e-16 ***
GRP2	-3.3225	0.3308	-10.043	< 2e-16 ***
GRP3	-4.4762	0.7311	-6.122	9.22e-10 ***
GRP4	-4.1297	0.4762	-8.672	< 2e-16 ***

Null deviance:	509.43	on 57	degrees of freedom
Residual deviance:	173.45	on 54	degrees of freedom

Example (continued)

```
> X2=sum(resid(terat.binom,type="pearson")^2);X2
[1] 154.707
> 1-pchisq(X2,df.residual(terat.binom)) # GoF Pearson
[1] 1.187217e-11

> X2/df.residual(terat.binom) # Evidence of overdispersion
[1] 2.864945
```

- Poor fit (GoF)
- Overdispersion, and standard errors invalid (too small) due to overdispersion
- Inter-litter variability that cannot be accounted for. Fetuses are more alike within litters than across litters, even within the same treatment group

Example (continued)

Possible solutions:

- GEE: models marginal (population averaged) effect of treatment
- GLMM: models litter-specific effect
- At least two other approaches not discussed (thoroughly) in this class:
 - Quasi-binomial: simplified version of GEE
 - Beta-binomial: parametric mixture model, analogous to negative-binomial for count data. Motivation similar to GLMM

Example (continued)

```
> terat.gee <- gee((Resp == "Dead") ~ GRP, id = Litter,  
+ data=teratbnry, family=binomial, corstr="exchangeable")  
> summary(terat.gee)
```

Coefficients:

	Estimate	Robust S.E.	Robust z
(Intercept)	-0.5889477	0.2966943	-1.985032
GRP2	1.2429690	0.5612748	2.214546
GRP3	1.6997950	0.8877114	1.914806
GRP4	1.9028396	0.7226377	2.633186

```
> # Big cor matrix (17 x 17), but all off-diag entries  
> terat.gee$working.correlation[1,2]  
[1] 0.8051211
```


Example (continued)

```
> terat.glmm <- glmer((Resp == "Dead") ~ GRP + (1|Litter),  
+ data = teratbnry, family = binomial)
```

```
> summary(terat.glmm)
```

Random effects:

Groups Name	Variance	Std.Dev.
-------------	----------	----------

Litter (Intercept)	2.284	1.511
--------------------	-------	-------

Number of obs: 607, groups: Litter, 58

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.8094	0.3616	-5.004	5.62e-07	***
GRP2	4.5396	0.7345	6.181	6.39e-10	***
GRP3	5.8833	1.1754	5.005	5.58e-07	***
GRP4	5.6062	0.9076	6.177	6.54e-10	***

Example (continued)

	Binomial ML	GEE	GLMM
(Intercept)	1.14 (0.13)	1.21 (0.27)	1.81 (0.33)
GRP2	-3.32 (0.33)	-3.37 (0.43)	-4.54 (0.68)
GRP3	-4.48 (0.73)	-4.58 (0.62)	-5.88 (1.18)
GRP4	-4.13 (0.48)	-4.25 (0.6)	-5.61 (0.86)

- SEs for binomial ML fit invalid (because of lack of fit)
- GEE estimates are similar to binomial but with larger SEs. Estimate marginal (population averaged) effects
- GLMM estimates are larger in magnitude. Estimate conditional (within litter) effects

Example (continued)

It seems that there are differences between groups 2,3,4 with the base group 1. As an exercise compare groups 2 and 3 for the GLMM model.

Hint: Create pairwise CI for $\beta_i - \beta_j$ for which you will need the estimated covariances $\text{Cov}(\hat{\beta}_i, \hat{\beta}_j)$ and remember to use Bonferonni adjustment.

We learned

- GLMM is a subject specific model, modeling how the predictors affect an individual subject's response
- GEE is a marginal model, modeling how the predictors affect the average of the subject's response