

GLMM vs GEE

Alexander McLain

Contents

The data are from a randomized, double-blind, parallel-group, multicenter study comparing two oral treatments (denoted A and B) for toe-nail infection. Patients were evaluated for the degree of onycholysis (the degree of separation of the nail plate from the nail-bed) at baseline (week 0) and at weeks 4, 8, 12, 24, 36, and 48 thereafter. The onycholysis outcome variable is binary (none or mild versus moderate or severe). The binary outcome was evaluated on 294 patients comprising a total of 1908 measurements.

```
library(tidyverse)
toenail <- read.csv("toenail.csv")
head(toenail)
```

ID	Response	Treatment	Month	Visit
1	1	1	0.0000000	1
1	1	1	0.8571429	2
1	1	1	3.5357143	3
1	0	1	4.5357143	4
1	0	1	7.5357143	5
1	0	1	10.0357143	6

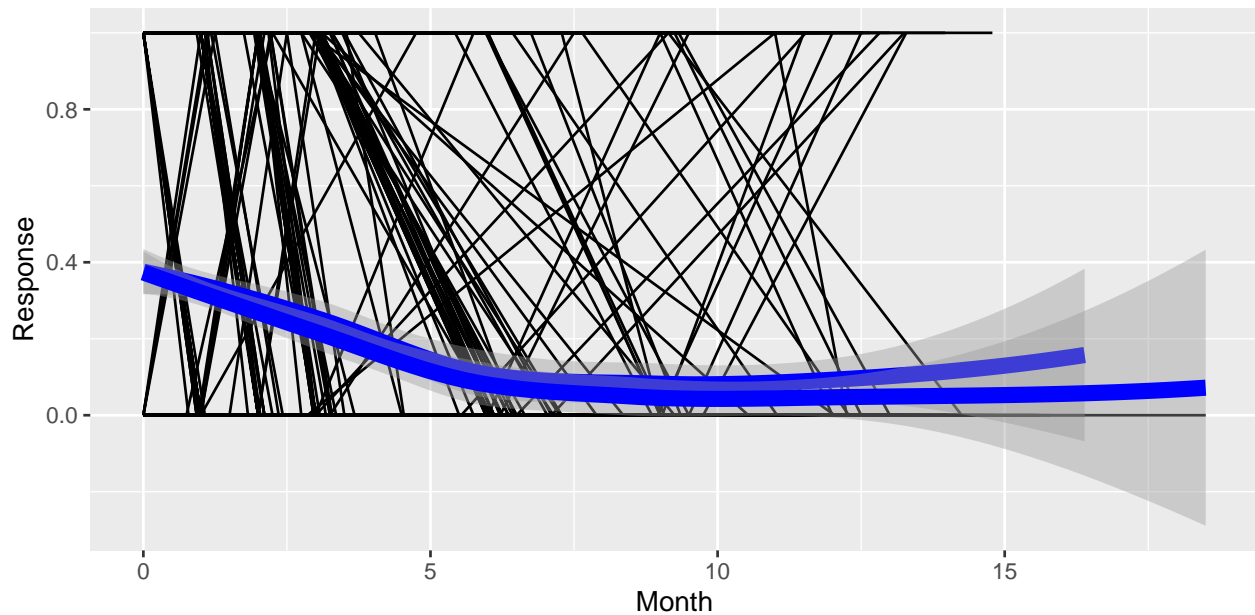
```
str(toenail)

## 'data.frame': 1908 obs. of 5 variables:
## $ ID : int 1 1 1 1 1 1 1 2 2 2 ...
## $ Response : int 1 1 1 0 0 0 0 0 0 1 ...
## $ Treatment: int 1 1 1 1 1 1 1 0 0 0 ...
## $ Month : num 0 0.857 3.536 4.536 7.536 ...
## $ Visit : int 1 2 3 4 5 6 7 1 2 3 ...
```

```
table( toenail$Treatment )
```

0	1
937	971

```
p <- ggplot(toenail, aes(x = Month, y = Response, group = ID) )
p + geom_line() + geom_smooth(aes(group = Treatment), method = "loess",
                              color = "blue", size = 3)
```



```
library(lme4)
library(lmerTest)
```

```
fm <- Response ~ Treatment*(Month + log(Month+1)) + (1|ID)
```

```
system.time(glmm_int_50 <- glmer(formula = fm, data = toenail, nAGQ = 50,
                                family = "binomial"))
```

```
##    user  system elapsed
##   6.718   0.017   6.754
```

```
system.time(glmm_int_20 <- glmer(formula = fm, data = toenail, nAGQ = 20,
                                family = "binomial"))
```

```
##    user  system elapsed
##   4.652   0.003   4.661
```

```
system.time(glmm_int_10 <- glmer(formula = fm, data = toenail, nAGQ = 10,
                                family = "binomial"))
```

```
##    user  system elapsed
##   2.673   0.003   2.676
```

```
system.time(glmm_int_5 <- glmer(formula = fm, data = toenail, nAGQ = 5,
                                family = "binomial"))
```

```
##    user  system elapsed
##   3.075   0.004   3.080
```

```
system.time(glmm_int_1 <- glmer(formula = fm, data = toenail, nAGQ = 1,
                                family = "binomial"))
```

```
##    user  system elapsed
##   3.467   0.004   3.472
```

```
system.time(glmm_int_0 <- glmer(formula = fm, data = toenail, nAGQ = 0,
                                family = "binomial"))
```

```
## user system elapsed
## 0.173 0.000 0.172
```

```
rbind( logLik(glmm_int_50), logLik(glmm_int_20), logLik(glmm_int_10),
       logLik(glmm_int_5) , logLik(glmm_int_1) , logLik(glmm_int_0))
```

```

-621.3481
-621.3342
-621.4979
-625.7075
-624.2390
-635.8513
```

```
VarCorr(glmm_int_50)
```

```
## Groups Name      Std.Dev.
## ID      (Intercept) 4.089
```

```
VarCorr(glmm_int_5)
```

```
## Groups Name      Std.Dev.
## ID      (Intercept) 3.7821
```

```
VarCorr(glmm_int_1)
```

```
## Groups Name      Std.Dev.
## ID      (Intercept) 4.6598
```

```
VarCorr(glmm_int_0)
```

```
## Groups Name      Std.Dev.
## ID      (Intercept) 3.6036
```

```
round( coef( summary( glmm_int_50)), 3)
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.296	0.472	-2.742	0.006
Treatment	-0.150	0.644	-0.232	0.816
Month	-0.209	0.100	-2.091	0.037
log(Month + 1)	-0.882	0.445	-1.983	0.047
Treatment:Month	-0.098	0.154	-0.633	0.527
Treatment:log(Month + 1)	-0.115	0.660	-0.173	0.862

```
round( coef( summary( glmm_int_1)), 3)
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-2.257	0.838	-2.693	0.007
Treatment	-0.320	0.750	-0.426	0.670
Month	-0.226	0.101	-2.246	0.025
log(Month + 1)	-0.834	0.443	-1.885	0.059
Treatment:Month	-0.106	0.155	-0.687	0.492
Treatment:log(Month + 1)	-0.078	0.655	-0.119	0.905

```
round( coef( summary( glmm_int_0)), 3)
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.529	0.390	-1.357	0.175
Treatment	-0.046	0.553	-0.084	0.933
Month	-0.182	0.092	-1.981	0.048
log(Month + 1)	-0.763	0.410	-1.859	0.063
Treatment:Month	-0.077	0.142	-0.541	0.589
Treatment:log(Month + 1)	-0.109	0.609	-0.178	0.858

Now let's fit the same model with a gee

```
library(geepack)
fm_gee <- Response ~ Treatment*(Month + log(Month+1))

gee_int <- geeglm(formula = fm_gee, data = toenail, family = binomial,
                  id = ID, corstr = "exchangeable")
round( coef(summary(gee_int)) , 3)
```

	Estimate	Std.err	Wald	Pr(> W)
(Intercept)	-0.435	0.178	5.985	0.014
Treatment	-0.049	0.248	0.038	0.845
Month	-0.108	0.054	4.044	0.044
log(Month + 1)	-0.290	0.189	2.347	0.126
Treatment:Month	-0.096	0.103	0.872	0.350
Treatment:log(Month + 1)	0.108	0.311	0.120	0.729

```
round( coef( summary( glmm_int_50)), 3)
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.296	0.472	-2.742	0.006
Treatment	-0.150	0.644	-0.232	0.816
Month	-0.209	0.100	-2.091	0.037
log(Month + 1)	-0.882	0.445	-1.983	0.047
Treatment:Month	-0.098	0.154	-0.633	0.527
Treatment:log(Month + 1)	-0.115	0.660	-0.173	0.862

Hence, the parameter vector in the GEE model needs to be interpreted completely different from the parameter vector in the GLMM: - GEE: marginal interpretation - GLMM: conditional interpretation, conditionally upon level of random effects

In general, the model for the marginal average is not of the same parametric form as the conditional average in the GLMM.

For logistic mixed models, with normally distributed random intercepts, it can be shown that the marginal model can be well approximated by again a logistic model, but with parameters approximately satisfying

$$\frac{\beta_{GLMM}}{\beta_{GEE}} = \sqrt{c^2\sigma^2 + 1} > 1$$

where σ is the standard deviation of the random intercepts and $c = 0.5881$. In the toenail example we had random intercept with standard deviation:

```
VarCorr(glmm_int_50)
```

```
## Groups Name      Std.Dev.
## ID      (Intercept) 4.089
```

```
( (0.5881^2) * (4.089^2) + 1)^(1/2)
```

```
## [1] 2.604377
```

the actual ratios were

```
res <- cbind(fixef(glmm_int_50), fixef(glmm_int_50)/2.604, gee_int$coefficients)
colnames(res) <- c("GLMM", "GLMM/2.604", "GEE")
res
```

	GLMM	GLMM/2.604	GEE
(Intercept)	-1.2955792	-0.4975343	-0.4349642
Treatment	-0.1496465	-0.0574679	-0.0485553
Month	-0.2085953	-0.0801057	-0.1079477
log(Month + 1)	-0.8821684	-0.3387744	-0.2896097
Treatment:Month	-0.0976014	-0.0374813	-0.0958057
Treatment:log(Month + 1)	-0.1145764	-0.0440002	0.1078359

```
rand_ef <- rnorm(100, 0 , 1)
Xi <- seq(-5, 5, 0.1)
p_mat <- NULL
for(i in 1:100 ){
  pi = exp( Xi + rand_ef[i])/( 1 + exp( Xi + rand_ef[i]) )
  p_mat <- rbind( p_mat, pi)
}
p_avg <- apply( p_mat, 2, mean)

matplot(Xi,t(p_mat), type = "l", xlab = expression(X[i]),
        ylab = expression(p[i]))
lines(Xi,p_avg,lwd=5)
```

