

# Case Studies for Marginal Models

Keunbaik Lee  
Sungkyunkwan University

*keunbaik@skku.edu*

November 16, 2021

- Just from the table, it appears that the two different doses of the analgesic treatment (2 and 3) both did better than the placebo (1). The high dose may be slightly better than the low dose.
- Is there any carry-over (period-treatment interaction) effect? It looks that the low dose is better than the high dose in period 1 but worse in the last two periods.
- We fit two models, assuming working independence or unstructured correlation.

```
setwd('d:/course/SKKU/Longitudinal_Data_Analysis/2016Fall/R-codes')
```

```
# Crossover trial
xover3 <- read.table("xover3.data",col.names=c("id","class","relief",
  "intercept","tx2","tx3","p2","p3","ptx1","ptx2","ptx3"))
xover3$period <- ifelse(xover3$p2==1,2, ifelse (xover3$p3==1,3,1))
xover3$treatment <- ifelse(xover3$tx2==1,2, ifelse (xover3$tx3==1,3,1))
with(xover3,ftable(period,relief,treatment))
```

```
xover3$ptx <- ifelse(xover3$ptx1==1,1,
  ifelse (xover3$ptx2==1,2,3))
xover3$ptx[xover3$period==1]<-0
with(xover3,ftable(ptx,relief,treatment))
```

```
library(gee)
xover.gee <- gee(relief~p2+p3+tx2+tx3+ptx2+ptx3,
  data=xover3,scale.fix=TRUE,id = id,family = binomial)
summary(xover.gee)
```

```
# Code for Conditional Logisitic Regression
library(survival)
xover3.cl <- clogit(relief~tx2+tx3+p2+p3+ptx2+ptx3+strata(id),
  data=xover3)
```

```
summary(xover3.cl)
```

```
> xover3 <- read.table("xover3.data",col.names=c("id","class","relief",
+ "intercept","tx2","tx3","p2","p3","ptx1","ptx2","ptx3"))
> xover3$period <- ifelse(xover3$p2==1,2, ifelse (xover3$p3==1,3,1))
> xover3$treatment <- ifelse(xover3$tx2==1,2, ifelse (xover3$tx3==1,3,1))
> with(xover3,ftable(period,relief,treatment))
```

```
      treatment  1  2  3
period relief
1         0      24  4 10
         1       7 19 22
2         0      20 11  5
```

```

      1          9 18 23
3      0          20 7 5
      1          6 23 25
>
> xover3$ptx <- ifelse(xover3$ptx1==1,1,
+                     ifelse(xover3$ptx2==1,2,3))
> xover3$ptx[xover3$period==1]<-0
> with(xover3,ftable(ptx,relief,treatment))
      treatment  1  2  3
ptx relief
0      0          24  4 10
      1          7 19 22
1      0          0  6  5
      1          0 23 26
2      0          20  0  5
      1          9  0 22
3      0          20 12  0
      1          6 18  0

> library(gee)
> xover.gee <- gee(relief~p2+p3+tx2+tx3+ptx2+ptx3,
+                 data=xover3,scale.fix=TRUE,id = id,family = binomial)
Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
running glm to get initial regression estimate
(Intercept)          p2          p3          tx2          tx3          ptx2
-1.0864970    0.4123752    0.5866249    2.1055755    2.0683345    -0.1284182
          ptx3
-0.9285768
> summary(xover.gee)

GEE:  GENERALIZED LINEAR MODELS FOR DEPENDENT DATA
gee S-function, version 4.13 modified 98/01/27 (1998)

```

```

Model:
Link:                               Logit

```

Variance to Mean Relation: Binomial  
Correlation Structure: Independent

Call:

```
gee(formula = relief ~ p2 + p3 + tx2 + tx3 + ptx2 + ptx3, id = id,  
    data = xover3, family = binomial, scale.fix = TRUE)
```

Summary of Residuals:

	Min	1Q	Median	3Q	Max
	-0.8328140	-0.2522785	0.1819871	0.2652069	0.8323952

Coefficients:

	Estimate	Naive S.E.	Naive z	Robust S.E.	Robust z
(Intercept)	-1.0864970	0.3282391	-3.3100778	0.3180752	-3.4158489
p2	0.4123752	0.4608082	0.8948955	0.4118732	1.0012187
p3	0.5866249	0.4750698	1.2348183	0.4560502	1.2863166
tx2	2.1055755	0.4017464	5.2410563	0.4191703	5.0231983
tx3	2.0683345	0.3830991	5.3989548	0.4185131	4.9421017
ptx2	-0.1284182	0.5016688	-0.2559821	0.5022616	-0.2556799
ptx3	-0.9285768	0.4872822	-1.9056242	0.4457959	-2.0829638

Estimated Scale Parameter: 1

Number of Iterations: 1

Working Correlation

	[,1]	[,2]	[,3]
[1,]	1	0	0
[2,]	0	1	0
[3,]	0	0	1

```
>  
> # Code for Conditional Logistic Regression  
> library(survival)  
> xover3.cl <- clogit(relief~tx2+tx3+p2+p3+ptx2+ptx3+strata(id),data=xover3)  
>
```

```
> summary(xover3.cl)
Call:
coxph(formula = Surv(rep(1, 258L), relief) ~ tx2 + tx3 + p2 +
      p3 + ptx2 + ptx3 + strata(id), data = xover3, method = "exact")
```

```
n= 258, number of events= 152
```

	coef	exp(coef)	se(coef)	z	Pr(> z )
tx2	1.9792	7.2369	0.4533	4.366	1.26e-05 ***
tx3	1.7082	5.5189	0.4066	4.201	2.66e-05 ***
p2	0.6931	1.9999	0.5582	1.242	0.2144
p3	0.8511	2.3423	0.5809	1.465	0.1429
ptx2	-0.1379	0.8712	0.5979	-0.231	0.8176
ptx3	-1.2433	0.2884	0.6510	-1.910	0.0562 .

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

	exp(coef)	exp(-coef)	lower .95	upper .95
tx2	7.2369	0.1382	2.97644	17.596
tx3	5.5189	0.1812	2.48746	12.245
p2	1.9999	0.5000	0.66967	5.973
p3	2.3423	0.4269	0.75017	7.313
ptx2	0.8712	1.1479	0.26986	2.812
ptx3	0.2884	3.4669	0.08053	1.033

```
Concordance= 0.84 (se = 0.047 )
Likelihood ratio test= 57.58 on 6 df, p=1e-10
Wald test               = 37.37 on 6 df, p=1e-06
Score (logrank) test = 55.8 on 6 df, p=3e-10
```

# ICHS Data

## Fit a GEE model

```
# ICHS Data

setwd('d:/course/SKKU/Longitudinal_Data_Analysis/R-codes')

ICHS <- read.table("ICHS.dat", header = TRUE)
ICHS
library (gee)
ig1 <- gee (RESPONSE~VITA+AGE+I (AGE^2)+GENDER+TIME+I (TIME^2),
            scale.fix=TRUE,cor="exchangeable",
            id=ID,data=ICHS,family="binomial")
summary(ig1)

summary(glm (RESPONSE~VITA+AGE+I (AGE^2)+GENDER+TIME+I (TIME^2),
             data=ICHS,family="binomial"))

library (geepack)
ig2 <- geese (RESPONSE~VITA+AGE+I (AGE^2)+GENDER+TIME+I (TIME^2),
             id=ID,corstr="exchangeable",data=ICHS,family="binomial")
summary (ig2)

> library (gee)
> ig1 <- gee (RESPONSE~VITA+AGE+I (AGE^2)+GENDER+TIME+I (TIME^2),
+           scale.fix=TRUE,cor="exchangeable",
+           id=ID,data=ICHS,family="binomial")
Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
running glm to get initial regression estimate
      (Intercept)           VITA           AGE           I (AGE^2)           GENDER
-1.1071681059   0.2360480102   0.3374780043  -0.0548290461  -0.5808244758
              TIME           I (TIME^2)
    0.0249863323  -0.0005285361
> summary(ig1)
```

GEE: GENERALIZED LINEAR MODELS FOR DEPENDENT DATA  
 gee S-function, version 4.13 modified 98/01/27 (1998)

Model:

Link: Logit  
 Variance to Mean Relation: Binomial  
 Correlation Structure: Exchangeable

Call:

```
gee(formula = RESPONSE ~ VITA + AGE + I(AGE^2) + GENDER + TIME +
    I(TIME^2), id = ID, data = ICHS, family = "binomial", corstr = "exchangeable",
    scale.fix = TRUE)
```

Summary of Residuals:

	Min	1Q	Median	3Q	Max
	-0.4751476	-0.3092661	-0.2373356	0.5952528	0.8596809

Coefficients:

	Estimate	Naive S.E.	Naive z	Robust S.E.	Robust z
(Intercept)	-1.1077203785	0.435203082	-2.5452953	0.448021579	-2.4724710
VITA	0.2211277915	0.222612016	0.9933327	0.223274883	0.9903836
AGE	0.3471462874	0.263699430	1.3164469	0.252813294	1.3731330
I(AGE^2)	-0.0566719852	0.034669589	-1.6346310	0.031893068	-1.7769374
GENDER	-0.5791466168	0.216144184	-2.6794458	0.216864474	-2.6705463
TIME	0.0249671660	0.028572074	0.8738311	0.029395526	0.8493526
I(TIME^2)	-0.0005279881	0.001813595	-0.2911279	0.001858689	-0.2840647

Estimated Scale Parameter: 1

Number of Iterations: 2

Working Correlation

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	1.0000000	0.5002811	0.5002811	0.5002811	0.5002811	0.5002811
[2,]	0.5002811	1.0000000	0.5002811	0.5002811	0.5002811	0.5002811



```
[3,] 0.5002811 0.5002811 1.0000000 0.5002811 0.5002811 0.5002811
[4,] 0.5002811 0.5002811 0.5002811 1.0000000 0.5002811 0.5002811
[5,] 0.5002811 0.5002811 0.5002811 0.5002811 1.0000000 0.5002811
[6,] 0.5002811 0.5002811 0.5002811 0.5002811 0.5002811 1.0000000
```

```
>
```

```
> summary(glm(RESPONSE~VITA+AGE+I (AGE^2)+GENDER+TIME+I (TIME^2),
+             data=ICHS,family="binomial"))
```

Call:

```
glm(formula = RESPONSE ~ VITA + AGE + I(AGE^2) + GENDER + TIME +
     I(TIME^2), family = "binomial", data = ICHS)
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-1.1368	-0.8614	-0.7356	1.3494	1.9660

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	-1.1071681	0.2566737	-4.314	1.61e-05	***
VITA	0.2360480	0.1189761	1.984	0.04726	*
AGE	0.3374780	0.1409397	2.394	0.01664	*
I(AGE^2)	-0.0548290	0.0185135	-2.962	0.00306	**
GENDER	-0.5808245	0.1155841	-5.025	5.03e-07	***
TIME	0.0249863	0.0403973	0.619	0.53624	
I(TIME^2)	-0.0005285	0.0025646	-0.206	0.83672	

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 1822.3  on 1499  degrees of freedom
Residual deviance: 1773.5  on 1493  degrees of freedom
AIC: 1787.5
```

Number of Fisher Scoring iterations: 4

```

> library (geepack)
> ig2 <- geese(RESPONSE~VITA+AGE+I(AGE^2)+GENDER+TIME+I(TIME^2),
+             id=ID, corstr="exchangeable", data=ICHS, family="binomial")
> summary (ig2)

Call:
geese(formula = RESPONSE ~ VITA + AGE + I(AGE^2) + GENDER + TIME +
      I(TIME^2), id = ID, data = ICHS, family = "binomial", corstr = "exchangeable")

Mean Model:
Mean Link:                logit
Variance to Mean Relation: binomial

Coefficients:
              estimate      san.se      wald      p
(Intercept) -1.1077246216 0.448030120 6.11292679 0.013419673
VITA         0.2210371826 0.223275827 0.98004780 0.322187006
AGE          0.3472061304 0.252820689 1.88603414 0.169649911
I(AGE^2)     -0.0566833780 0.031893921 3.15860716 0.075527598
GENDER       -0.5791397231 0.216866328 7.13152576 0.007574025
TIME         0.0249671534 0.029395657 0.72139267 0.395687466
I(TIME^2)    -0.0005279876 0.001858698 0.08069188 0.776361979

Scale Model:
Scale Link:                identity

Estimated Scale Parameters:
              estimate      san.se      wald p
(Intercept) 0.9945625 0.06322268 247.4676 0

Correlation Model:
Correlation Structure:      exchangeable
Correlation Link:          identity

```

```
Estimated Correlation Parameters:
      estimate      san.se      wald p
alpha 0.5016817 0.05550145 81.70484 0
```

```
Returned Error Value: 0
```

```
Number of clusters: 250 Maximum cluster size: 6
```

# Seizure Data

```
# Seizure Data
```

```
setwd('d:/course/SKKU/Longitudinal_Data_Analysis/2016Fall/R-codes')
```

```
library(gee)
```

```
library(geepack)
```

```
seize<-read.table("seize.data",  
col.names=c("id","seizure","week","progabide","baseline8","age"))  
seize.base <- data.frame(id=seize$id,seizure=seize$baseline8,week=seize
```

```
$week,progabide=seize$prog,age=seize$age)  
seize.base <- seize.base[seize.base$week==1,]  
seize.base$week<-0  
seize.full<-rbind(seize[,c(1:4,6)],seize.base[,])  
seize.full <- seize.full[order(seize.full$id,seize.full$week),]  
seize.full$time <- ifelse(seize.full$week==0,8,2)  
seize.full$post <- seize.full$week!= 0  
seize.full[1:10,]
```

```
sg2<-gee(seizure~progabide+post+post:progabide+offset(log(time)),  
data=seize.full,id=id,family="poisson",cor="exchangeable")  
summary(sg2)
```

```
sg2 <- geese(seizure ~ progabide+post+post:progabide+offset(log(time)),  
sformula= ~ progabide,data=seize.full,id=id,family="poisson",  
corstr="exchangeable")  
summary(sg2)
```

```
> sg2<-gee(seizure~progabide+post+post:progabide+offset(log(time)),  
+ data=seize.full,id=id,family="poisson",cor="exchangeable")  
Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27  
running glm to get initial regression estimate
```

(Intercept)	progabide	postTRUE	progabide:postTRUE
1.34760922	0.02651461	0.11079814	0.10368067

```
> summary(sg2)
```

```
GEE:  GENERALIZED LINEAR MODELS FOR DEPENDENT DATA  
gee S-function, version 4.13 modified 98/01/27 (1998)
```

```
Model:
```

```
Link:                      Logarithm  
Variance to Mean Relation: Poisson  
Correlation Structure:     Exchangeable
```

```
Call:
```

```
gee(formula = seizure ~ progabide + post + post:progabide + offset(log(time)),  
    id = id, data = seize.full, family = "poisson", corstr = "exchangeable")
```

```
Summary of Residuals:
```

Min	1Q	Median	3Q	Max
-4.299107	-1.299107	2.020161	10.374640	147.048387

```
Coefficients:
```

	Estimate	Naive S.E.	Naive z	Robust S.E.	Robust z
(Intercept)	1.34760922	0.1511851	8.9136359	0.1573571	8.5640166
progabide	0.02651461	0.2072721	0.1279217	0.2218539	0.1195138
postTRUE	0.11079814	0.1547038	0.7161956	0.1160997	0.9543358
progabide:postTRUE	-0.10368067	0.2199500	-0.4713830	0.2136100	-0.4853736

```
Estimated Scale Parameter: 19.70269
```

```
Number of Iterations: 1
```

```
Working Correlation
```

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	1.000000	0.771588	0.771588	0.771588	0.771588
[2,]	0.771588	1.000000	0.771588	0.771588	0.771588
[3,]	0.771588	0.771588	1.000000	0.771588	0.771588
[4,]	0.771588	0.771588	0.771588	1.000000	0.771588

```
[5,] 0.771588 0.771588 0.771588 0.771588 1.000000
>
> sg2 <- geese(seizure ~ progabide+post+post:progabide+offset(log(time)),
+             sformula= ~ progabide,data=seize.full,id=id,family="poisson",
+             corstr="exchangeable")
> summary(sg2)
```

Call:

```
geese(formula = seizure ~ progabide + post + post:progabide +
      offset(log(time)), sformula = ~progabide, id = id, data = seize.full,
      family = "poisson", corstr = "exchangeable")
```

Mean Model:

```
Mean Link:          log
Variance to Mean Relation: poisson
```

Coefficients:

	estimate	san.se	wald	p
(Intercept)	1.34760922	0.1573571	73.34238067	0.0000000
progabide	0.02651461	0.2218539	0.01428355	0.9048683
postTRUE	0.11079814	0.1160997	0.91075687	0.3399137
progabide:postTRUE	-0.10368067	0.2136100	0.23558757	0.6274113

Scale Model:

```
Scale Link:          identity
```

Estimated Scale Parameters:

	estimate	san.se	wald	p
(Intercept)	14.228486	4.450311	10.22202	0.001387738
progabide	9.910183	16.618842	0.35560	0.550960076

Correlation Model:

```
Correlation Structure:  exchangeable
Correlation Link:       identity
```

```
Estimated Correlation Parameters:  
      estimate      san.se      wald p  
alpha 0.7451173 0.08108394 84.44611 0
```

```
Returned Error Value:      0
```

```
Number of clusters:      59      Maximum cluster size: 5
```

## Seizure Data: GEE1.5

For illustration, we assume an exchangeable correlation structure where the correlation may depend on age.

$$\begin{aligned} \text{cor}(Y_{ij}, Y_{ik}) &= \rho_i, \\ \log\left(\frac{1 + \rho_i}{1 - \rho_i}\right) &= \alpha_1 + \alpha_2 \text{Age}_i. \end{aligned}$$

This model can be fitted using `geese`. The design matrix for the correlation model has to be constructed by hand. The matrix  $Z$  has the same number of rows as the number of clusters (the covariate should be invariant within a cluster).

```
# Seizure Data: GEE1.5
setwd('d:/course/SKKU/Longitudinal_Data_Analysis/2016Fall/R-codes')
library(gee)
library(geepack)

seize<-read.table("seize.data",col.names=c("id","seizure","week","progabide","baseline8","age"))
seize.base <- data.frame(id=seize$id,seizure=seize$baseline8,week=seize$week,progabide=seize$progabide)
seize.base <- seize.base[seize.base$week==1,]
seize.base$week<-0
seize.full<-rbind(seize[,c(1:4,6)],seize.base[,])
seize.full <- seize.full[order(seize.full$id,seize.full$week),]
seize.full$time <- ifelse(seize.full$week==0,8,2)
seize.full$post <- seize.full$week!= 0
```



```

z <- cbind(1, seize.full$age[seize.full$week==0])
sg2 <-geese(seizure~progabide+post+post:progabide+offset(log(time)),
            sformula=~progabide, data=seize.full, id=id, family="poisson",
            cor.link="fisherz", zcor=z, corstr="exchangeable")
summary(sg2)

sg2 <-geese(seizure~progabide+post+post:progabide+offset(log(time)),
            data=seize.full, id=id, family="poisson",
            corstr="exchangeable", jack=TRUE, jls=TRUE, fij=TRUE)
summary(sg2)

> sg2 <-geese(seizure~progabide+post+post:progabide+offset(log(time)),
+             sformula=~progabide, data=seize.full, id=id, family="poisson",
+             cor.link="fisherz", zcor=z, corstr="exchangeable")
> summary(sg2)

Call:
geese(formula = seizure ~ progabide + post + post:progabide +
      offset(log(time)), sformula = ~progabide, id = id, data = seize.full,
      zcor = z, family = "poisson", cor.link = "fisherz", corstr = "exchangeable")

```

```

Mean Model:
Mean Link:          log
Variance to Mean Relation: poisson

```

```

Coefficients:
              estimate      san.se      wald      p
(Intercept)    1.413723297 0.1637308 7.455354e+01 0.0000000
progabide       0.003521332 0.2181660 2.605192e-04 0.9871222
postTRUE       0.119463923 0.1377950 7.516344e-01 0.3859593
progabide:postTRUE -0.321152168 0.2845887 1.273464e+00 0.2591179

```

```

Scale Model:
Scale Link:          identity

```

```

Estimated Scale Parameters:
      estimate      san.se      wald      p
(Intercept) 13.30317  4.055167 10.762000 0.001036054
progabide   14.30612 19.011502  0.5662533 0.451751620

```

```

Correlation Model:
Correlation Structure:  exchangeable
Correlation Link:       fisherz

```

```

Estimated Correlation Parameters:
      estimate      san.se      wald      p
alpha:1  5.8652470 2.78349010 4.440105 0.03510409
alpha:2 -0.1370331 0.08058039 2.891961 0.08902250

```

```

Returned Error Value: 0
Number of clusters: 59 Maximum cluster size: 5

```

```

>
> sg2 <-geese(seizure~progabide+post+post:progabide+offset(log(time)),
+           data=seize.full,id=id,family="poisson",
+           corstr="exchangeable",jack=TRUE,jls=TRUE,fij=TRUE)
> summary(sg2)

```

```

Call:
geese(formula = seizure ~ probabide + post + post:progabide +
      offset(log(time)), id = id, data = seize.full, family = "poisson",
      corstr = "exchangeable", jack = TRUE, jls = TRUE, fij = TRUE)

```

```

Mean Model:
Mean Link:      log
Variance to Mean Relation: poisson

```

```

Coefficients:
      estimate      san.se      ajs.se      jls.se      fij.se
(Intercept) 1.34760922 0.1573571 0.1546652 0.1546652 0.1583893

```

progabide	0.02651461	0.2218539	0.2176720	0.2176720	0.2250338
postTRUE	0.11079814	0.1160997	0.1141136	0.1141136	0.1202731
progabide:postTRUE	-0.10368067	0.2136100	0.2094277	0.2094277	0.2419675
		wald		p	
(Intercept)	73.34238067	0.0000000			
progabide	0.01428355	0.9048683			
postTRUE	0.91075687	0.3399137			
progabide:postTRUE	0.23558757	0.6274113			

Scale Model:

Scale Link: identity

Estimated Scale Parameters:

	estimate	san.se	ajs.se	jls.se	fij.se	wald	p
(Intercept)	19.43553	8.697971	9.356896	8.956847	8.842138	4.992946	0.02545084

Correlation Model:

Correlation Structure: exchangeable

Correlation Link: identity

Estimated Correlation Parameters:

	estimate	san.se	ajs.se	jls.se	fij.se	wald	p
alpha	0.776891	0.07534604	0.3140797	0.1753292	0.175155	106.3162	0

Returned Error Value: 0

Number of clusters: 59 Maximum cluster size: 5