

GLMM, Concepts, & R

Bill Last Updated:

22 March, 2020

Contents

Preface: Motivation	1
1 LM and GLM	3
1.1 LM	3
1.2 GLM-Definition	3
1.3 GLM-log link example	4
1.4 GLM-Reciprocal link:	4
1.5 GLM-exponential family:	4
1.6 Canonical exponential family	5
1.7 Canonical exponential family - Expected value and variance . . .	6
1.8 Expected value and variance - Poisson Example	8
1.9 Canonical link	9
1.10 Canonical link - Bernoulli	10
1.11 NR - Bernoulli	11
1.12 Iteratively Re-weighted Least Squares	12
2 Linear Mixed Models	13
2.1 LMM	13
2.2 Calculate mean	13
2.3 Test the treatment effect	15
2.4 Another example	16
2.5 Full LMM model	18
2.6 Serial correlations in time and space	20

3	Generalized Linear Mixed Models	23
3.1	Basics of GLMM	23
3.2	Some References	24

Preface: Motivation

All the notes I have done here are the preparation for my stat master project, which will be about Generalized Linear Mixed Models. While I have tried my best, probably there are still some typos and errors. Please feel free to let me know in case you find one. Thank you!

Chapter 1

LM and GLM

Before moving to LMM, I would like to review LM and GLM first.

1.1 LM

$$Y|X \sim N(\mu(X), \sigma^2 I)$$

$$E(Y|X) = \mu(X) = X^T \beta$$

where,

$\mu(X)$: *random component*

$X^T \beta$: *covariates*

1.2 GLM-Definition

Ref: https://ocw.mit.edu/courses/mathematics/18-650-statistics-for-applications-fall-2016/lecture-slides/MIT18_650F16_GLM.pdf

$$Y \sim \text{exponential family}$$

Link function

$$g(\mu(X)) = X^T \beta$$

1.3 GLM-log link example

$$\mu_i = \gamma e^{\delta t_i}$$

Link function is log link, and it becomes:

$$\log(\mu_i) = \log(\gamma) + \log(\delta t_i) = \beta_0 + \beta_1 t_i$$

(This is somehow similar to Poisson distribution.)

1.4 GLM-Reciprocal link:

$$\mu_i = \frac{\alpha x_i}{h + x_i}$$

Reciprocal link:

$$g(\mu_i) = \frac{1}{\mu_i} = \frac{1}{\alpha} + \frac{h}{\alpha} \frac{1}{x_i} = \beta_0 + \beta_1 \frac{1}{x_i}$$

1.5 GLM-exponential family:

In a more general sense, for exponential family:

$$\begin{aligned} P_\theta(X) = P(X, \theta) &= e^{\sum \eta_i(\theta) T_i(X)} C(\theta) h(x) \\ &= e^{\sum \eta_i(\theta) T_i(X)} e^{-\log(\frac{1}{c(\theta)})} h(x) \\ &= e^{\sum \eta_i(\theta) T_i(X) - \log(\frac{1}{c(\theta)})} h(x) \\ &= e^{\sum \eta_i(\theta) T_i(X) - B(\theta)} h(x) \end{aligned}$$

Normal distribution

For normal distributions, it belongs to exponential family.

$$\begin{aligned} P_\theta(X) &= \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} \\ &= e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} e^{\log(\frac{1}{\sigma \sqrt{2\pi}})} \\ &= e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2} - \log(\sigma \sqrt{2\pi})} \\ &= e^{-\frac{1}{2\sigma^2} x^2 - \frac{1}{2\sigma^2} \mu^2 + \frac{x\mu}{\sigma^2} - \log(\sqrt{2\pi}\sigma)} \\ &= e^{-\frac{1}{2\sigma^2} x^2 + \frac{x\mu}{\sigma^2} - (\frac{1}{2\sigma^2} \mu^2 + \log(\sqrt{2\pi}\sigma))} \end{aligned}$$

Where,

$$\eta_1 = -\frac{1}{2\sigma^2} \text{ and } T_1(x) = x^2$$

$$\eta_2 = -\frac{\mu}{\sigma^2} \text{ and } T_2(x) = x$$

$$B(\theta) = \frac{1}{2\sigma^2}\mu^2 + \log(\sqrt{2\pi}\sigma)$$

$$h(x) = 1$$

In the case above, $\theta = (\mu, \sigma^2)$. If σ^2 is known, $\theta = \mu$. In this case, we can rewrite the normal pdf as follows.

$$\begin{aligned} P_\theta(X) &= e^{-\frac{1}{2\sigma^2}x^2 - \frac{1}{2\sigma^2}\mu^2 + \frac{x\mu}{\sigma^2} - \log(\sqrt{2\pi}\sigma)} \\ &= e^{\frac{x\mu}{\sigma^2} - \frac{1}{2\sigma^2}\mu^2} e^{-\frac{1}{2\sigma^2}x^2 - \log(\sqrt{2\pi}\sigma)} \end{aligned}$$

Where,

$$\eta_1 = -\frac{\mu}{\sigma^2} \text{ and } T_1(x) = x$$

$$B(\theta) = \frac{1}{2\sigma^2}\mu^2$$

$$h(x) = e^{-\frac{1}{2\sigma^2}x^2 - \log(\sqrt{2\pi}\sigma)}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{x^2}{\sigma^2}}$$

Thus, we can see that $h(x)$ is a normal pdf $\sim N(0, \sigma^2)$.

Bernoulli

Another example, x is discrete. For example, Bernoulli:

$$\begin{aligned} &= p^x(1-p)^{1-x} \\ &= e^{\log(p^x(1-p)^{1-x})} \\ &= e^{x\log(p) + (1-x)\log(1-p)} \\ &= e^{x\log(p) - x\log(1-p) + \log(1-p)} \\ &= e^{x\log(\frac{p}{1-p}) + \log(1-p)} \end{aligned}$$

Where,

$$\eta_1 = \log\left(\frac{p}{1-p}\right) \text{ and } T_1(x) = x$$

$$B(\theta) = \log\left(\frac{1}{1-p}\right)$$

$$h(x) = 1$$

1.6 Canonical exponential family

Canonical exponential family:

$$f_{\theta}(x) = e^{\frac{x\theta - b(\theta)}{\phi} + c(x, \phi)}$$

where, $b(\cdot)$ and $c(\cdot, \cdot)$ are known.

Normal distribution

Again, use the normal pdf:

$$\begin{aligned} P_{\theta}(X) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} \\ &= e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} e^{\log(\frac{1}{\sigma\sqrt{2\pi}})} \\ &= e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2} - \log(\sigma\sqrt{2\pi})} \\ &= e^{-\frac{1}{2\sigma^2} x^2 - \frac{1}{2\sigma^2} \mu^2 + \frac{x\mu}{\sigma^2} - \log(\sqrt{2\pi}\sigma)} \\ &= e^{\frac{x\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2} + (-\frac{1}{2\sigma^2} x^2 - \log(\sqrt{2\pi}\sigma))} \\ &= e^{\frac{x\mu - \frac{1}{2}\mu^2}{\sigma^2} + (-\frac{1}{2\sigma^2} x^2 - \log(\sqrt{2\pi}\sigma))} \end{aligned}$$

Where (we assume σ^2 is known.),

$$\theta = \mu$$

$$\phi = \sigma^2$$

$$b(\theta) = \frac{1}{2}\theta^2$$

$$\begin{aligned} c(x, \phi) &= -\frac{1}{2\sigma^2} x^2 - \log(\sqrt{2\pi}\sigma) \\ &= -\frac{1}{2\sigma^2} x^2 - \frac{1}{2} \log(2\pi\sigma^2) \\ &= -\frac{1}{2} \left(\frac{x^2}{\sigma^2} + \log(2\pi\sigma^2) \right) \\ &= -\frac{1}{2} \left(\frac{x^2}{\phi} + \log(2\pi\phi) \right) \end{aligned}$$

1.7 Canonical exponential family - Expected value and variance

First derivative

Canonical exponential family:

$$f_{\theta}(x) = e^{\frac{x\theta - b(\theta)}{\phi} + c(x, \phi)}$$

1.7. CANONICAL EXPONENTIAL FAMILY - EXPECTED VALUE AND VARIANCE 7

log likelihood (only one observation)

$$\log f_{\theta}(x)$$

$$\begin{aligned} E\left[\frac{\partial(\log f_{\theta}(X))}{\partial\theta}\right] &= E\left[\frac{\frac{\partial f_{\theta}(X)}{\partial\theta}}{f_{\theta}(X)}\right] \\ &= \int \frac{\frac{\partial f_{\theta}(X)}{\partial\theta}}{f_{\theta}(X)} f_{\theta}(X) dx \\ &= \int \frac{\partial f_{\theta}(X)}{\partial\theta} dx \\ &= \frac{\partial}{\partial\theta} \int f_{\theta}(X) dx \\ &= \frac{\partial 1}{\partial\theta} \\ &= 0 \end{aligned}$$

Second derivative

Second derivative

$$\begin{aligned} E\left[\frac{\partial^2(\log f_{\theta}(X))}{\partial\theta^2}\right] &= E\left[\frac{\partial}{\partial\theta}\left(\frac{\frac{\partial f_{\theta}(X)}{\partial\theta}}{f_{\theta}(X)}\right)\right] \\ &= E\left[\frac{\frac{\partial^2 f_{\theta}(X)}{\partial\theta^2} f_{\theta}(X) - \left(\frac{\partial f_{\theta}(X)}{\partial\theta}\right)^2}{f_{\theta}^2(X)}\right] \\ &= \int \frac{\frac{\partial^2 f_{\theta}(X)}{\partial\theta^2} f_{\theta}(X) - \left(\frac{\partial f_{\theta}(X)}{\partial\theta}\right)^2}{f_{\theta}^2(X)} dx \\ &= \int \left(\frac{\partial^2 f_{\theta}(X)}{\partial\theta^2} - \frac{\left(\frac{\partial f_{\theta}(X)}{\partial\theta}\right)^2}{f_{\theta}(X)}\right) dx \\ &= \int \frac{\partial^2 f_{\theta}(X)}{\partial\theta^2} dx - \int \frac{\left(\frac{\partial f_{\theta}(X)}{\partial\theta}\right)^2}{f_{\theta}(X)} dx \\ &= \frac{\partial^2}{\partial\theta^2} \int f_{\theta}(X) dx - \int \frac{\left(\frac{\partial f_{\theta}(X)}{\partial\theta}\right)^2}{f_{\theta}(X)} dx \\ &= 0 - \int \frac{\left(\frac{\partial f_{\theta}(X)}{\partial\theta}\right)^2}{f_{\theta}(X)} dx \\ &= 0 - \int \frac{\left(\frac{\partial f_{\theta}(X)}{\partial\theta}\right)^2}{(f_{\theta}(X))^2} f_{\theta}(x) dx \\ &= -E\left[\left(\frac{\frac{\partial f_{\theta}(X)}{\partial\theta}}{f_{\theta}(X)}\right)^2\right] \\ &= -E\left[\left(\frac{\partial(\log f_{\theta}(X))}{\partial\theta}\right)^2\right] \end{aligned}$$

Based on the first derivative, we can get:

$$\begin{aligned} \log(f_\theta(X)) &= \frac{X\theta - b(\theta)}{\phi} + c(X, \phi) \\ E\left[\frac{\partial(\log(f_\theta(X)))}{\partial\theta}\right] &= E\left[\frac{X - b'(\theta)}{\phi}\right] = \frac{E(X) - b'(\theta)}{\phi} = 0 \end{aligned}$$

Thus, we can get,

$$E(X) = b'(\theta)$$

For second derivative, from the calculation above, we know that,

$$\begin{aligned} E\left[\frac{\partial^2(\log f_\theta(X))}{\partial\theta^2}\right] &= -E\left[\left(\frac{\partial(\log f_\theta(X))}{\partial\theta}\right)^2\right] \\ &= -E\left[\left(\frac{X - b'(\theta)}{\phi}\right)^2\right] \\ &= -E\left[\left(\frac{X - E(X)}{\phi}\right)^2\right] \\ &= -\frac{\text{Var}(X)}{\phi^2} \end{aligned}$$

At the same time,

$$\begin{aligned} E\left[\frac{\partial^2(\log f_\theta(X))}{\partial\theta^2}\right] &= E\left[\frac{\partial\left(\frac{X - b'(\theta)}{\phi}\right)}{\partial\theta}\right] \\ &= E\left[-\frac{b''(\theta)}{\phi}\right] \\ &= -\frac{b''(\theta)}{\phi} \end{aligned}$$

Thus,

$$\text{Var}(X) = b''(\theta)\phi$$

1.8 Expected value and variance - Poisson Example

Example of poisson distribution

$$P(\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

If we put k as y , and λ as μ , we can get:

$$P(\mu) = \frac{\mu^y e^{-\mu}}{y!}$$

Compare to,

$$f_{\theta}(y) = e^{\frac{y\theta - b(\theta)}{\phi} + c(y, \phi)}$$

We can write it as the exponential format:

$$\begin{aligned} P(\mu) &= \frac{\mu^y e^{-\mu}}{y!} \\ &= e^{\log(\mu^y) + \log(e^{-\mu}) - \log(y!)} \\ &= e^{y \log(\mu) - \mu - \log(y!)} \end{aligned}$$

We thus know that $\theta = \log(\mu)$. We can continue to write the equation above as follows.

$$= e^{y\theta - e^{\theta} - \log(y!)}$$

Thus, we can get:

$$\begin{aligned} E(X) &= \frac{\partial(e^{\theta})}{\partial\theta} = e^{\theta} = \mu \\ Var(X) &= \frac{\partial''(e^{\theta})}{\partial\theta^2} \phi = \frac{\partial''(e^{\theta})}{\partial\theta^2} = \mu \end{aligned}$$

1.9 Canonical link

A link function can link $X^T\beta$ to the mean μ .

That is,

$$g(\mu) = X^T\beta \rightarrow \mu = g^{-1}(X^T\beta)$$

We know that

$$\mu = b'(\theta)$$

Thus,

$$b'(\theta) = g^{-1}(X^T \beta)$$

Thus,

$$g = b'^{-1}(\theta)$$

1.10 Canonical link - Bernoulli

PMF of Bernoulli:

$$\begin{aligned} &= p^y (1-p)^{1-y} \\ &= e^{\log(p^y (1-p)^{1-y})} \\ &= e^{y \log(p) + (1-y) \log(1-p)} \\ &= e^{y \log(p) - y \log(1-p) + \log(1-p)} \\ &= e^{y \log(\frac{p}{1-p}) + \log(1-p)} \end{aligned}$$

Copared to the following:

$$f_\theta(y) = e^{\frac{y\theta - b(\theta)}{\phi} + c(y, \phi)}$$

We need to change the format of Bernoulli:

$$\theta = \log \frac{p}{1-p}$$

Thus,

$$e^\theta = \frac{p}{1-p} \rightarrow p = \frac{e^\theta}{1+e^\theta}$$

After that, we can contintue the Bernoulli:

$$\begin{aligned} &= e^{y\theta + \log(1 - \frac{e^\theta}{1+e^\theta})} \\ &= e^{y\theta + \log(\frac{1}{1+e^\theta})} \\ &= e^{y\theta - \log(1+e^\theta)} \end{aligned}$$

Where,

$$b(\theta) = \log(1 + e^\theta)$$

We can then try to calculate the derivative:

$$b'(\theta) = \frac{\partial(\log(1 + e^\theta))}{\partial\theta} = \frac{e^\theta}{1 + e^\theta}$$

We know that

$$b'(\theta) = \mu$$

Thus, we can get

$$\mu = \frac{e^\theta}{1 + e^\theta}$$

We can then calculate the inverse function:

$$\theta = \log\left(\frac{\mu}{1 - \mu}\right)$$

Thus,

$$g(\mu) = \log\left(\frac{\mu}{1 - \mu}\right) = X^T \beta$$

1.11 NR - Bernoulli

We know that the PMF for Bernoulli:

$$\begin{aligned} &= p^y(1 - p)^{1-y} \\ &= e^{y\theta - \log(1 + e^\theta)} \\ &= e^{yx^T \beta - \log(1 + e^{x^T \beta})} \end{aligned}$$

Thus,

$$\ell(\beta|Y, X) = \sum_{i=1}^n (Y_i X_i^T \beta - \log(1 + e^{X_i^T \beta}))$$

Thus, the gradient is:

$$\nabla_\ell(\beta) = \sum_{i=1}^n (Y_i X_i - \frac{e^{X_i^T \beta}}{1 + e^{X_i^T \beta}})$$

The Hessian is:

$$H_\ell(\beta) = - \sum_{i=1}^n \frac{e^{X_i^T \beta}}{(1 + e^{X_i^T \beta})^2} X_i X_i^T$$

Thus,

$$\beta^{k+1} = \beta^k - (H_\ell(\beta^k))^{-1} \nabla_\ell(\beta^k)$$

1.12 Iteratively Re-weighted Least Squares

Chapter 2

Linear Mixed Models

2.1 LMM

The following is a shortened version of Jonathan Rosenblatt's LMM tutorial.
<http://www.john-ros.com/Rcourse/lme.html>.

In addition, another reference is from Douglas Bates's R package document.
https://cran.r-project.org/web/packages/lme4/vignettes/lmer.pdf?fbclid=IwAR1nmmRP9A0BrhKdgBibNjM5acR_spTpXV8QlQGdmTWyQz3ZtV3LYn6kCbQ

Assume that y is a function of x and u , where x is the fixed effect and u is the random effect. Thus, we can get,

$$y|x, u = x'\beta + z'u + \epsilon$$

For random effect, one example can be that you want to test the treatment effect, and sample 8 observations from 4 groups. You measure before and after the treatment. In this case, x represents the treatment effect, whereas z represents the group effect (i.e., random effect). Note that, in this case, it reminds the paired t-test. Remember in SPSS, why do we do paired t-test? Typically, it is the case when we measure a subject (or, participant) twice. In this case, we can consider each participant as an unit of random effect (rather than as group in the last example.)

2.2 Calculate mean

The following code generates 4 numbers ($N(0, 10)$) for 4 groups. Then, replicate it within each group. That is, in the end, there are 8 observations.

Note that, in the following code, there are no “independent variables”. Both the linear model and mixed model are actually just trying to calculate the mean. Note that `lmer(y~1+1|groups)` and `lmer(y~1|groups)` will generate the same results.

```
set.seed(123)
n.groups <- 4 # number of groups
n.repeats <- 2 # samples per group
#Generating index for observations belong to the same group
groups <- as.factor(rep(1:n.groups, each=n.repeats))
n <- length(groups)
#Generating 4 random numbers, assuming normal distribution
z0 <- rnorm(n.groups, 0, 10)
z <- z0[as.numeric(groups)] # generate and inspect random group effects
z
```

```
## [1] -5.6047565 -5.6047565 -2.3017749 -2.3017749 15.5870831 15.5870831 0.7050839
## [8] 0.7050839
```

```
epsilon <- rnorm(n,0,1) # generate measurement error
beta0 <- 2 # this is the actual parameter of interest! The global mean.
y <- beta0 + z + epsilon # sample from an LMM

# fit a linear model assuming independence
# i.e., assume that there is no "group things".
lm.5 <- lm(y~1)

# fit a mixed-model that deals with the group dependence
#install.packages("lme4")
library(lme4)
lme.5.a <- lmer(y~1+1|groups)
lme.5.b <- lmer(y~1|groups)
lm.5
```

```
##
## Call:
## lm(formula = y ~ 1)
##
## Coefficients:
## (Intercept)
##          4.283
```

```
lme.5.a
```

```
## Linear mixed model fit by REML ['lmerMod']
```

```
## Formula: y ~ 1 + 1 | groups
## REML criterion at convergence: 36.1666
## Random effects:
##   Groups   Name          Std.Dev.
##   groups   (Intercept) 8.8521
##   Residual                0.8873
## Number of obs: 8, groups:  groups, 4
## Fixed Effects:
## (Intercept)
##          4.283
```

```
lme.5.b
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: y ~ 1 | groups
## REML criterion at convergence: 36.1666
## Random effects:
##   Groups   Name          Std.Dev.
##   groups   (Intercept) 8.8521
##   Residual                0.8873
## Number of obs: 8, groups:  groups, 4
## Fixed Effects:
## (Intercept)
##          4.283
```

2.3 Test the treatment effect

As we can see that, LLM and paired t-test generate the same t-value.

```
times<-rep(c(1,2),4) # first time and second time
times
```

```
## [1] 1 2 1 2 1 2 1 2
```

```
data_combined<-cbind(y,groups,times)
data_combined
```

```
##           y groups times
## [1,] -3.4754687     1     1
## [2,] -1.8896915     1     2
## [3,]  0.1591413     2     1
## [4,] -1.5668361     2     2
```

```
## [5,] 16.9002303      3      1
## [6,] 17.1414212      3      2
## [7,]  3.9291657      4      1
## [8,]  3.0648977      4      2

lme_diff_times<- lmer(y~times+(1|groups))

t_results<-t.test(y~times, paired=TRUE)

lme_diff_times

## Linear mixed model fit by REML ['lmerMod']
## Formula: y ~ times + (1 | groups)
## REML criterion at convergence: 35.0539
## Random effects:
## Groups      Name          Std.Dev.
## groups      (Intercept) 8.845
## Residual                1.013
## Number of obs: 8, groups: groups, 4
## Fixed Effects:
## (Intercept)          times
##      4.5691      -0.1908

print("The following results are from paired t-test")

## [1] "The following results are from paired t-test"

t_results$statistic

##           t
## 0.2664793
```

2.4 Another example

```
data(Dyestuff, package='lme4')
attach(Dyestuff)

## The following objects are masked from Dyestuff (pos = 5):
##
##      Batch, Yield
```

```
Dyestuff
```

```
##      Batch Yield
## 1      A  1545
## 2      A  1440
## 3      A  1440
## 4      A  1520
## 5      A  1580
## 6      B  1540
## 7      B  1555
## 8      B  1490
## 9      B  1560
## 10     B  1495
## 11     C  1595
## 12     C  1550
## 13     C  1605
## 14     C  1510
## 15     C  1560
## 16     D  1445
## 17     D  1440
## 18     D  1595
## 19     D  1465
## 20     D  1545
## 21     E  1595
## 22     E  1630
## 23     E  1515
## 24     E  1635
## 25     E  1625
## 26     F  1520
## 27     F  1455
## 28     F  1450
## 29     F  1480
## 30     F  1445
```

```
lme_batch<- lmer( Yield ~ 1 + (1|Batch) , Dyestuff )
summary(lme_batch)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Yield ~ 1 + (1 | Batch)
##      Data: Dyestuff
##
## REML criterion at convergence: 319.7
##
## Scaled residuals:
```

```
##      Min      1Q  Median      3Q      Max
## -1.4117 -0.7634  0.1418  0.7792  1.8296
##
## Random effects:
## Groups   Name                Variance Std.Dev.
## Batch    (Intercept) 1764      42.00
## Residual                2451      49.51
## Number of obs: 30, groups:  Batch, 6
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 1527.50      19.38      78.8
```

2.5 Full LMM model

In the following, I used the data from the package of `lme4`. For `Days + (1 | Subject)`, it only has random intercept; in contrast, `Days + (Days| Subject)` has both random intercept and random slope for `Days`. Note that, random effects do not generate specific slopes for each level of `Days`, but rather just a variance of all the slopes.

Therefore, we can see that “`Days + (Days| Subject)`” and “`Days + (1+Days| Subject)`” generate the same results. For more discussion, you can refer to the following link: <https://www.jaredknowles.com/journal/2013/11/25/getting-started-with-mixed-effect-models-in-r>

```
data(sleepstudy, package='lme4')
attach(sleepstudy)
```

```
## The following objects are masked from sleepstudy (pos = 5):
##
##      Days, Reaction, Subject
```

```
fm1 <- lmer(Reaction ~ Days + (1 | Subject), sleepstudy)
summary(fm1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Reaction ~ Days + (1 | Subject)
##      Data: sleepstudy
##
## REML criterion at convergence: 1786.5
##
## Scaled residuals:
```

```
##      Min      1Q  Median      3Q      Max
## -3.2257 -0.5529  0.0109  0.5188  4.2506
##
## Random effects:
## Groups   Name                Variance Std.Dev.
## Subject  (Intercept) 1378.2    37.12
## Residual                    960.5    30.99
## Number of obs: 180, groups: Subject, 18
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 251.4051    9.7467    25.79
## Days        10.4673     0.8042    13.02
##
## Correlation of Fixed Effects:
##      (Intr)
## Days -0.371
```

```
fm2<-lmer ( Reaction ~ Days + ( Days| Subject ) , data= sleepstudy )
summary(fm2)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Reaction ~ Days + (Days | Subject)
## Data: sleepstudy
##
## REML criterion at convergence: 1743.6
##
## Scaled residuals:
##      Min      1Q  Median      3Q      Max
## -3.9536 -0.4634  0.0231  0.4633  5.1793
##
## Random effects:
## Groups   Name                Variance Std.Dev. Corr
## Subject  (Intercept) 611.90    24.737
##          Days        35.08    5.923   0.07
## Residual                    654.94    25.592
## Number of obs: 180, groups: Subject, 18
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 251.405    6.824    36.843
## Days        10.467     1.546     6.771
##
## Correlation of Fixed Effects:
##      (Intr)
```

```
## Days -0.138
```

```
fm3<-lmer ( Reaction ~ Days + (1+Days| Subject ) , data= sleepstudy )
summary(fm3)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Reaction ~ Days + (1 + Days | Subject)
## Data: sleepstudy
##
## REML criterion at convergence: 1743.6
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.9536 -0.4634  0.0231  0.4633  5.1793
##
## Random effects:
## Groups Name Variance Std.Dev. Corr
## Subject (Intercept) 611.90 24.737
## Days 35.08 5.923 0.07
## Residual 654.94 25.592
## Number of obs: 180, groups: Subject, 18
##
## Fixed effects:
## Estimate Std. Error t value
## (Intercept) 251.405 6.824 36.843
## Days 10.467 1.546 6.771
##
## Correlation of Fixed Effects:
## (Intr)
## Days -0.138
```

2.6 Serial correlations in time and space

The hierarchical model of $y|x, u = x'\beta + z'u + \epsilon$ can work well for correlations within blocks, but not for correlations in time as the correlations decay in time. The following uses nlme package to calculate time serial data.

```
library(nlme)
head(nlme::Ovary,n=50)
```

```
## Grouped Data: follicles ~ Time | Mare
## Mare Time follicles
## 1 1 -0.13636360 20
```


## 2	1	-0.09090910	15
## 3	1	-0.04545455	19
## 4	1	0.00000000	16
## 5	1	0.04545455	13
## 6	1	0.09090910	10
## 7	1	0.13636360	12
## 8	1	0.18181820	14
## 9	1	0.22727270	13
## 10	1	0.27272730	20
## 11	1	0.31818180	22
## 12	1	0.36363640	15
## 13	1	0.40909090	18
## 14	1	0.45454550	17
## 15	1	0.50000000	14
## 16	1	0.54545450	18
## 17	1	0.59090910	14
## 18	1	0.63636360	16
## 19	1	0.68181820	17
## 20	1	0.72727270	18
## 21	1	0.77272730	18
## 22	1	0.81818180	17
## 23	1	0.86363640	14
## 24	1	0.90909090	12
## 25	1	0.95454550	12
## 26	1	1.00000000	14
## 27	1	1.04545500	10
## 28	1	1.09090900	11
## 29	1	1.13636400	16
## 30	2	-0.15000000	6
## 31	2	-0.10000000	6
## 32	2	-0.05000000	8
## 33	2	0.00000000	7
## 34	2	0.05000000	16
## 35	2	0.10000000	10
## 36	2	0.15000000	13
## 37	2	0.20000000	9
## 38	2	0.25000000	7
## 39	2	0.30000000	6
## 40	2	0.35000000	8
## 41	2	0.40000000	8
## 42	2	0.45000000	6
## 43	2	0.50000000	8
## 44	2	0.55000000	7
## 45	2	0.60000000	9
## 46	2	0.65000000	6
## 47	2	0.70000000	4

```
## 48    2  0.75000000    5
## 49    2  0.80000000    8
## 50    2  0.85000000   11
```

```
fm10var.lme <- nlme::lme(fixed=follicles ~ sin(2*pi*Time) + cos(2*pi*Time),
  data = Ovary,
  random = pdDiag(~sin(2*pi*Time)),
  correlation=corAR1() )
summary(fm10var.lme)
```

```
## Linear mixed-effects model fit by REML
## Data: Ovary
##      AIC      BIC    logLik
## 1563.448 1589.49 -774.724
##
## Random effects:
## Formula: ~sin(2 * pi * Time) | Mare
## Structure: Diagonal
##      (Intercept) sin(2 * pi * Time) Residual
## StdDev:      2.858385          1.257977 3.507053
##
## Correlation Structure: AR(1)
## Formula: ~1 | Mare
## Parameter estimate(s):
##      Phi
## 0.5721866
## Fixed effects: follicles ~ sin(2 * pi * Time) + cos(2 * pi * Time)
##              Value Std.Error DF   t-value p-value
## (Intercept)  12.188089 0.9436602 295 12.915760 0.0000
## sin(2 * pi * Time) -2.985297 0.6055968 295 -4.929513 0.0000
## cos(2 * pi * Time) -0.877762 0.4777821 295 -1.837159 0.0672
## Correlation:
##              (Intr) s(*p*T
## sin(2 * pi * Time) 0.000
## cos(2 * pi * Time) -0.123 0.000
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.34910093 -0.58969626 -0.04577893 0.52931186 3.37167486
##
## Number of Observations: 308
## Number of Groups: 11
```

Chapter 3

Generalized Linear Mixed Models

3.1 Basics of GLMM

Recall the formula in the probit model:

$$Y^* = X\beta + \epsilon, \epsilon \sim N(0, \sigma^2) = N(0, I)$$

Similar to LMM, binary model with random effect can be written as follows.

$$Y^* = X\beta + Zu + \epsilon$$

where,

$$\epsilon \sim N(0, I)$$

$$u \sim N(0, D)$$

We also assume ϵ and u are independent. Thus, we know that D represents the variances of the random effects. If we make $u = 1$, the model becomes the usual probit model. McCulloch (1994) states that there are a few advantages to use probit, rather than logit models. (Note that, however, probit is not canonical link function, but logit is!)

The following is the note from Charle E. McCulloch's "Maximum likelihood algorithms for Generalized Linear Mixed Models"

3.2 Some References

<http://www.biostat.umn.edu/~baolin/teaching/linmods/glmm.html>

http://www.biostat.umn.edu/~baolin/teaching/probmods/GLMM_mcmc.html

<https://bbolker.github.io/mixedmodels-misc/glmmFAQ.html>