BIOSTAT 571 Homework 6

Wenxiao Gu

February 26, 2014

Contents

1 Marginal models induced by linear mixed effects models

1.1 Mixed effects

$$\begin{aligned} Y_{ij} &\sim N(\mu+\gamma_i+\beta_j,\sigma^2) \text{ (independent conditional on } \gamma_i) \\ \gamma_i &\underset{iid}{\sim} N(0,\sigma_\gamma^2) \text{ (random effect always with mean zero.)} \end{aligned}$$

We know the operation of Kronecker product is as below,

$$A_{m\times n}\otimes B_{p\times q} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{1n}b_{1q} \\ a_{11}b_{21} & & \\ & & \\ a_{m1}b_{p1} & & a_{mn}b_{pq} \end{bmatrix}_{mn\times pq}$$

For a short try, $Y_{11}\sim N(\mu+\gamma_1+\beta_1,\sigma^2)$ and $Y_{12}\sim N(\mu+\gamma_1+\beta_2,\sigma^2)$.

$$Y_{2\times 1}|\gamma_{1\times 1} \sim N\left(\mu 1_{2\times 1}+\gamma_{1\times 1}\otimes 1_{2\times 1}+(\beta_1,\beta_2)^T,\sigma^2 I_2\right) = N\left(\mu 1_{2\times 1}+\gamma_{1\times 1}\otimes 1_{2\times 1}+1_1\otimes\beta_{2\times 1},\sigma^2 I_2\right)$$

The matrix notation of the above model is

$$\begin{aligned} Y_{mn\times 1}|\gamma_{n\times 1} &\sim N\left(\mu 1_{mn\times 1}+\gamma_{n\times 1}\otimes 1_{m\times 1}+1_n\otimes\beta_{m\times 1},\sigma^2 I_{mn}\right) \\ \gamma_{n\times 1} &\sim N(0,\sigma_\gamma^2 I_n) \end{aligned}$$

where $\gamma=(\gamma_1,\ldots,\gamma_n)^T$ and $\beta=(\beta_1,\ldots,\beta_m)^T$.
For all Y_{ij}, the marginal expectation is

$$E[Y_{ij}] = E\left[E\left(Y_{ij}|\gamma_i\right)\right] = E\left[\mu+\gamma_i+\beta_j\right] = \mu+E\left(\gamma_i\right)+\beta_j = \mu+\beta_j$$

and the marginal variance is

$$\begin{aligned} Var[Y_{ij}] &= Var\left[E\left(Y_{ij}|\gamma_i\right)\right]+E\left[Var\left(Y_{ij}|\gamma_i\right)\right] \\ &= Var\left[\mu+\gamma_i+\beta_j\right]+E\left[\sigma^2\right] \\ &= \sigma_\gamma^2+\sigma^2 \end{aligned}$$

The marginal covariance has two cases,

- Within group: using conditional independence between Y_{ij} and $Y_{ij'}$

$$\begin{aligned} Cov[Y_{ij},Y_{ij'}] &= Cov\left[E\left(Y_{ij}|\gamma_i\right),E\left(Y_{ij'}|\gamma_i\right)\right]+E\left[Cov\left(Y_{ij},Y_{ij'}|\gamma_i\right)\right] \\ &= Cov\left[\mu+\gamma_i+\beta_j,\mu+\gamma_i+\beta_{j'}\right]+0 \\ &= \sigma_\gamma^2 \end{aligned}$$

- Between group: between Y_{ij} and $Y_{i'j}$

$$\begin{aligned} Cov[Y_{ij},Y_{i'j}] &= Cov\left[E\left(Y_{ij}|\gamma_i\right),E\left(Y_{i'j}|\gamma_{i'}\right)\right]+E\left[Cov\left(Y_{ij},Y_{i'j}|\gamma_i,\gamma_{i'}\right)\right] \\ &= Cov\left[\mu+\gamma_i+\beta_j,\mu+\gamma_{i'}+\beta_j\right]+E\left[Cov\left(\mu+\gamma_i+\beta_j+\epsilon_{ij},\mu+\gamma_{i'}+\beta_j+\epsilon_{i'j}|\gamma_i,\gamma_{i'}\right)\right] \\ &= Cov\left[\gamma_i,\gamma_{i'}\right]+E\left[Cov\left(\epsilon_{ij},\epsilon_{i'j}|\gamma_i,\gamma_{i'}\right)\right] \\ &= 0 \end{aligned}$$

The convolution of normal is normal. We have the mean and the covariance, and now plug them into the marginal model of Y. Let $\mu_i = (\mu + \beta_1, \ldots, \mu + \beta_m)^T$. Next, construct the vector of the mean Y,

$$\mu_{mn\times 1} = \left(\mu_1^T, ..., \mu_n^T\right)^T = E(Y_{mn\times 1})$$

and its covariance matrix,

$$\Sigma_{mn\times mn} = \sigma^2 I_{mn} + \sigma_\gamma^2 I_n \otimes \left(1_m 1_m^T\right).$$

Hence, the distribution of Y is as below,

$$Y_{mn\times 1} \sim MVN\left(\mu_{mn\times 1}, \Sigma_{mn\times mn}\right).$$

1.2 Random effects

$$\begin{aligned}
Y_{ij} &\sim N(\mu + \gamma_i + \beta_j, \sigma^2) \text{ (independent conditional on } \alpha_j \text{ and } \gamma_i) \\
\gamma_i &\underset{iid}{\sim} N(0, \sigma_\gamma^2) \text{ (random effect always with mean zero.)} \\
\alpha_j &\underset{iid}{\sim} N(0, \sigma_\alpha^2) \text{ (random effect always with mean zero.)} \\
\alpha_j &\perp \gamma_i
\end{aligned}$$

From part a, we know how to contrct a random column effect. Now we construct a random row effect. Similarly, the matrix notation of the above model is

$$\begin{aligned}
Y_{mn\times 1}|\gamma_{n\times 1} &\sim N\left(\mu 1_{mn\times 1} + \gamma_{n\times 1} \otimes 1_{m\times 1} + 1_n \otimes \alpha_{m\times 1}, \sigma^2 I_{nm}\right) \\
\gamma_{n\times 1} &\sim N(0, \sigma_\gamma^2 I_n) \\
\alpha_{m\times 1} &\sim N(0, \sigma_\alpha^2 I_m) \\
\alpha_j &\perp \gamma_i,
\end{aligned}$$

where $\gamma = (\gamma_1, \ldots, \gamma_n)^T$ and $\alpha = (\alpha_1, \ldots, \alpha_m)^T$.
For all Y_{ij}, the marginal expectation is

$$E[Y_{ij}] = E\left[E\left(Y_{ij}|\gamma_i, \alpha_j\right)\right] = E\left[\mu + \gamma_i + \alpha_j\right] = \mu + E\left(\gamma_i\right) + E\left(\alpha_j\right) = \mu$$

and the marginal variance is

$$\begin{aligned}
Var[Y_{ij}] &= Var\left[E\left(Y_{ij}|\gamma_i, \alpha_j\right)\right] + E\left[Var\left(Y_{ij}|\gamma_i\right)\right] \\
&= Var\left[\mu + \gamma_i + \alpha_j\right] + E\left[\sigma^2\right] \\
&= \sigma_\gamma^2 + \sigma_\alpha^2 + \sigma^2
\end{aligned}$$

The marginal covariance has three cases,

- Within "row": using conditional independence between Y_{ij} and $Y_{ij'}$

$$\begin{aligned}
Cov[Y_{ij}, Y_{ij'}] &= Cov\left[E\left(Y_{ij}|\gamma_i, \alpha_j\right), E\left(Y_{ij'}|\gamma_i, \alpha_{j'}\right)\right] + E\left[Cov\left(Y_{ij}, Y_{ij'}|\gamma_i, \alpha_j, \alpha_{j'}\right)\right] \\
&= Cov\left[\mu + \gamma_i + \alpha_j, \mu + \gamma_i + \alpha_{j'}\right] + 0 \\
&= Cov\left(\gamma_i, \gamma_i\right) + Cov\left(\gamma_i, \alpha_{j'}\right) + Cov\left(\alpha_j, \gamma_i\right) + Cov\left(\alpha_j, \alpha_{j'}\right) \\
&= \sigma_\gamma^2 + 0 + 0 + 0 \\
&= \sigma_\gamma^2
\end{aligned}$$

- Within "column": using conditional independence between Y_{ij} and $Y_{i'j}$

$$\begin{aligned} Cov[Y_{ij}, Y_{i'j}] &= Cov\left[E\left(Y_{ij}|\gamma_i, \alpha_j\right), E\left(Y_{i'j}|\gamma_{i'}, \alpha_j\right)\right] + E\left[Cov\left(Y_{ij}, Y_{i'j}|\gamma_i, \gamma_{i'}, \alpha_j\right)\right] \\ &= Cov\left[\mu + \gamma_i + \alpha_j, \mu + \gamma_{i'} + \alpha_j\right] + 0 \\ &= Cov\left(\gamma_i, \gamma_{i'}\right) + Cov\left(\gamma_i, \alpha_j\right) + Cov\left(\alpha_j, \gamma_{i'}\right) + Cov\left(\alpha_j, \alpha_j\right) \\ &= 0 + 0 + 0 + \sigma_\alpha^2 \\ &= \sigma_\alpha^2 \end{aligned}$$

- Between Y_{ij} and $Y_{i'j'}$

$$\begin{aligned} Cov[Y_{ij}, Y_{i'j'}] &= Cov\left[E\left(Y_{ij}|\gamma_i, \alpha_j\right), E\left(Y_{i'j'}|\gamma_{i'}, \alpha_{j'}\right)\right] + E\left[Cov\left(Y_{ij}, Y_{i'j}|\gamma_i, \gamma_{i'}, \alpha_j, \alpha_{j'}\right)\right] \\ &= Cov\left[\mu + \gamma_i + \alpha_j, \mu + \gamma_{i'} + \alpha_{j'}\right] + 0 \\ &= Cov\left(\gamma_i, \gamma_{i'}\right) + Cov\left(\gamma_i, \alpha_{j'}\right) + Cov\left(\alpha_j, \gamma_{i'}\right) + Cov\left(\alpha_j, \alpha_{j'}\right) \\ &= 0 + 0 + 0 + 0 \\ &= 0 \end{aligned}$$

The convolution of normal is normal. We have the mean and the covariance, and now plug them into the marginal model of Y. Let $\mu_i = \mu 1_n$. Next, construct the vector of the mean Y,

$$\mu_{mn\times 1} = \left(\mu_1^T, ..., \mu_n^T\right)^T = E(Y_{mn\times 1})$$

and its covariance matrix,

$$\Sigma_{mn\times mn} = \sigma^2 I_{mn} + \sigma_\gamma^2 I_n \otimes \left(1_m 1_m^T\right) + \sigma_\alpha^2 \left(1_n 1_n^T\right) \otimes I_m.$$

Hence, the distribution of Y is as below,

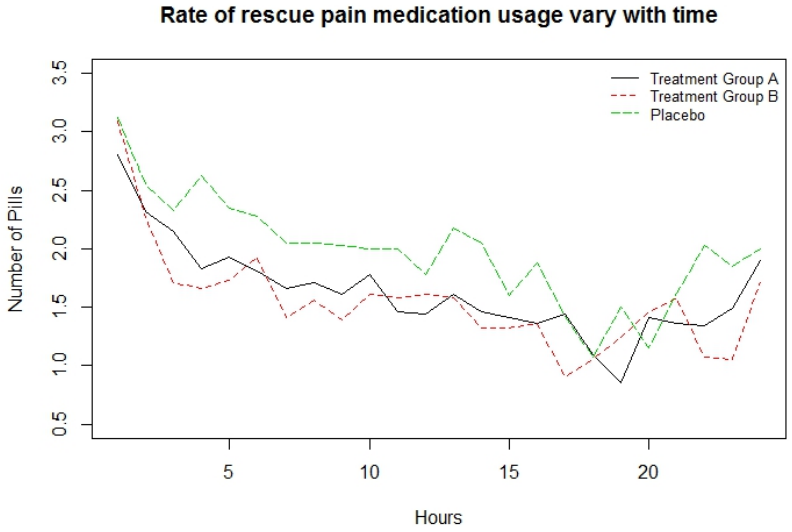
$$Y_{mn\times 1} \sim MVN\left(\mu_{mn\times 1}, \Sigma_{mn\times mn}\right).$$

2 Pain data

2.1 Descriptive analysis

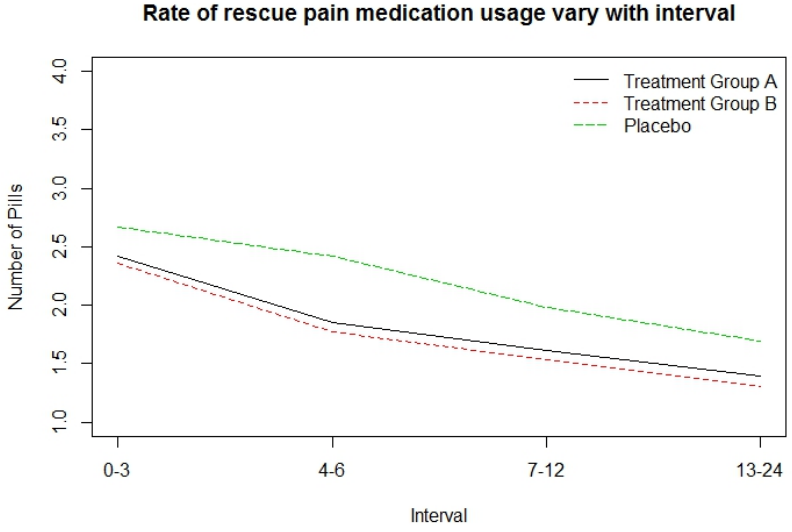
2.1.1 a)

1) The number of pills used decreases during the inverval between 0 and 19 hours.
2) We also see an increase in the number of pills used between 19 and 24 hours.
3) The number of pills used in the placebo group are generally larger than the other two treatment groups



2.1.2 b)

Now draw a plot in which the x-axis is interval instead of hours. We see that the relationship is linear for each of the three groups. The number of pills generally decreases over time. However, the number of pills used in the placebo group are generally larger than the other two treatment groups.



2.1.3 c)

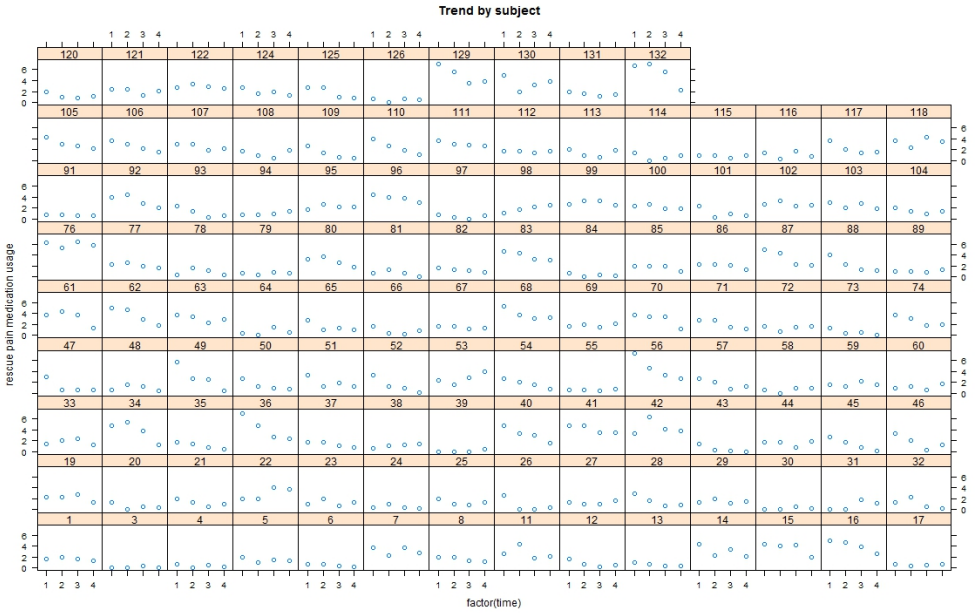
According to the second plot, the temporal trend is negative.

2.1.4 d)

We see the temporal trends are different between the treatment groups and placebo group, but are not very different between treatments A and B.

2.1.5 e)

In each group, I made the plot of number of pills over time for each patient. They show very different linear relationship.



2.2 Candidates

2.2.1 (i) no random effects with unstructured homoscedastic residuals

It's not plausible, because in our data the variance of Y across time is not constant.

2.2.2 (ii) random intercepts with independent homoscedastic residuals

It's not plausible, because in our data the variance of Y across time is not constant.

2.2.3 (iii) independent random intercepts and slopes, with independent homoscedastic residuals

It's not plausible, because in our data the variance of Y across time is decreasing. However the independent $Var(Y_{ij}) = G_{11} + G_{22}t_{ij}^2 + \sigma^2$ increases over time.

2.2.4 (iv) correlated random intercepts and slopes, with independent homoscedastic residuals

It may be plausible, because in our data the variance of Y across time is decreasing. However the independent $Var(Y_{ij}) = G_{11} + 2G_{12}t_{ij} + G_{22}t_{ij}^2 + \sigma^2$ decreases over time if $G_{12} < -4G_{22}$.

2.2.5 (v) no random effects with homoscedastic AR-1 residuals

It's not plausible, because in our data the variance of Y across time is not constant.

2.2.6 (vi) no random effects with unstructured heteroscedastic residuals

It's plausible, because in our data the variance of Y across time is heteroscedastic and the unstructured structure can refect the correlation in the data.

2.2.7 (vii) random intercepts with independent heteroscedastic residuals**

It's plausible, because in our data the variance of Y across time is heteroscedastic and the model takes into account the correlation by assuming that it comes from the random intercept.

2.2.8 (viii) independent random intercepts and slopes with independent heteroscedastic residuals

From problem 1, we know that the correlation between measurements can be built by the varaince of the random effects. So it's plausible.

2.2.9 (ix) correlated random intercepts and slopes with independent heteroscedastic residuals

It may be plausible, because in our data the variance of Y across time is decreasing. However the independent $Var(Y_{ij}) = G_{11} + 2G_{12}t_{ij} + G_{22}t_{ij}^2 + \sigma^2$ decreases over time if $G_{12} < -4G_{22}$.

2.2.10 (x) no random effects with heteroscedastic AR-1 residuals.

It's plausible, because in our data the variance of Y across time is heteroscedastic and the AR-1 structure can refect the correlation in the data.

2.3 Estimate unstructured within-subject covariance and correlation matrices for the purpose of exploratory analysis

2.3.1 The within-participant covariance and correlation matrices

$$\hat{R} = \begin{bmatrix} 1 & .812 & .728 & .603 \\ & 1 & .818 & .664 \\ & & 1 & .782 \\ & & & 1 \end{bmatrix}$$

$$\hat{\Sigma} = \begin{bmatrix} 2.618 & 1.988 & 1.490 & 1.002 \\ & 2.289 & 1.565 & 1.032 \\ & & 1.599 & 1.016 \\ & & & 1.055 \end{bmatrix}$$

From the outputs, the model (vii) is excluded because it induces an exchangeable covariance structure which contradicts our matrix. Also, the model (viii) is excluded because it's an special case of (ix), so we do not need to fit a model that is not as good as another candidate.

2.4 Summary of scientific findings

The models (vi), (ix) and (x) remains in the list. Their AICs are 1309, 1314, and 1306 respectively. We go with the model (x) with the smallest AIC. The output of (x) indicated that the overall temporal trend is negative.

Name of coefficient	$\hat{\beta}$	Std. Err.	95% CI	
Intercept	2.98	0.311	2.37	3.59
time	-0.323	0.072	-0.465	-0.180
groupC (control)	ref			
groupA	-0.281	0.437	-1.14	0.577
groupB	0.338	0.437	-1.20	0.521
time×groupC (control)	ref			
time×groupA	-0.00159	0.102	-0.198	0.201
time×groupB	-0.00709	0.102	-0.207	0.193

- In the pain medication usage study, there was evidence of a linear trend in usage of pain medication over time. We estimated that one unit long time (interval) was associated with 0.323 unit less intake of medication on average in the placebo group (95% CI: -0.465, -0.180). Since the CI does not contain zero, there is evidence that the trend in placebo group over time.
- On average, the difference comparing the trend in medication usage in treatment group A to placebo group is -0.00159 (95% CI: -0.198, 0.201). Then the estimated trend in group A is -0.3246 (95% CI: -0.523, -0.124). Since the CI of the difference contains zero, there is not enough evidence that the trend in treatment group A is any different than the on in control group.
- On average, the difference comparing the trend in medication usage in treatment group B to placebo group is -0.709 (95% CI: -0.207, 0.193). The estimated trend in group B is thus -0.330 (95% CI: -0.537, -0.137). Since the CI of the difference contains zero, there is not enough evidence that the trend in treatment group B is any different than the on in control group.

2.5 AIC criterion

Model	AIC
1	1342
2	1390
3	1392
4	1318
5	1339
6	1309
7	1364
8	1366
9	1314
10	1306

The 3 models screen by us in previous parts are now with the smallest AICs. They are models (xi), (ix) and (x), as expected.

3 Sandwich and bootstrap standard error estimates

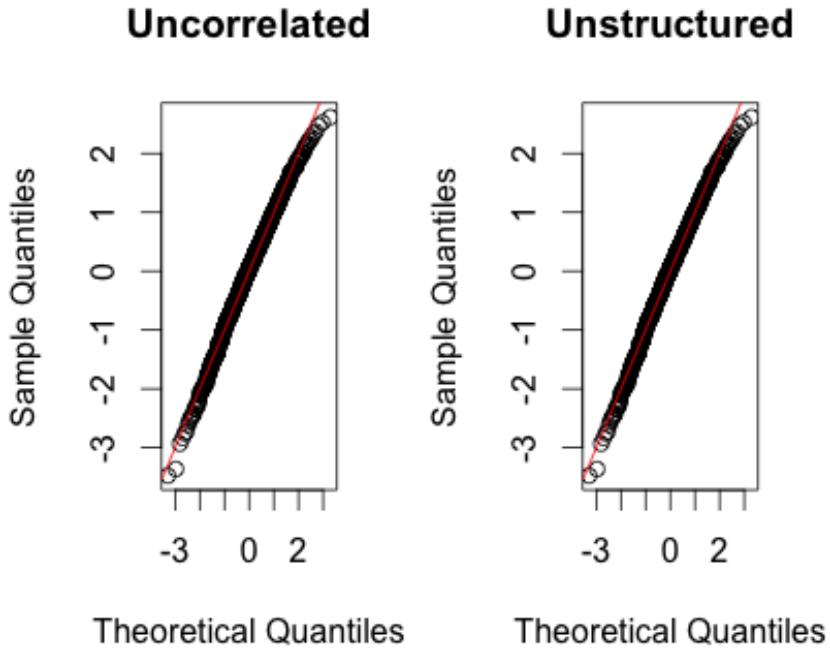
3.1 (a)

The sandwich-based standard error estimate from model (i) unstructured is 0.117; the sandwich-based standard error estimate from model (ii) exchangable is 0.113.

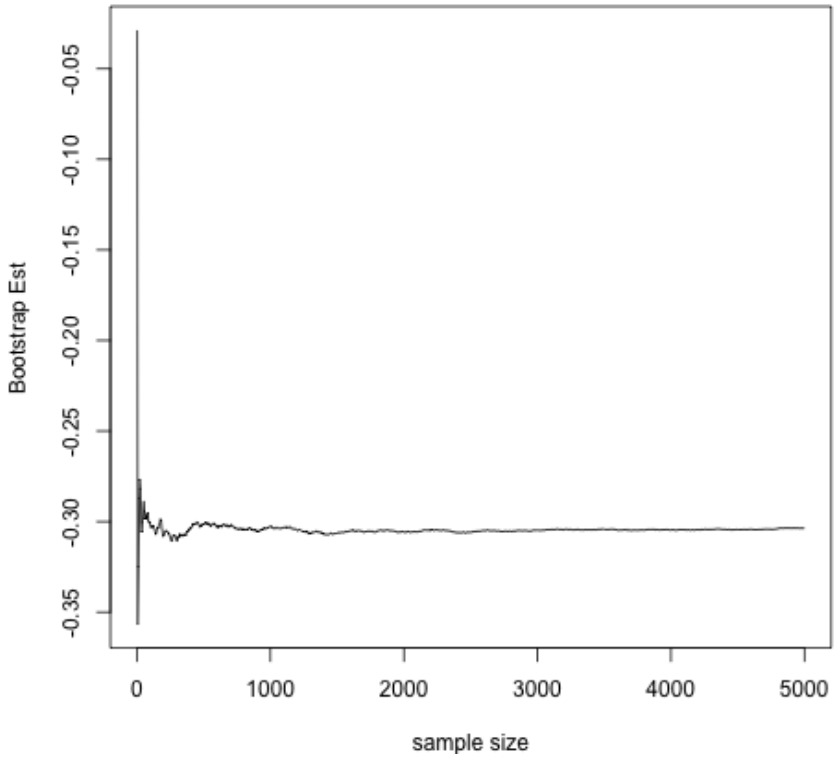
3.2 (b)

We resampled 27 individuals (clusters), keeping observations intact, and replicated for 10000 times. The bootstrap standard error estimates from model (i) is 0.119, from model (ii) is 0.156.

From the qq plots of the simulated distributions, we can say that the normality is gauranteed. So bootstrap estimates should be reliable.



The sample size gets larger, the estimates converge.



3.3 (c)

- For bootstrap resampling both cluster and observation, we first randomly resampled 27 individuals, then resampled 4 observations within individual.
- For bootstrap ignoring clusters, we randomly resampled 27 times 4 observations.

Bootstrap type	Model	SD
Cluster and observation	model (i)	0.219
Ignore cluster	model (i)	0.196

3.4 (d)

The standard deviation in model (i) from different bootstraps are as bellow. We can see that the boostrap only resampling clusters work as good as the sandwich estimates. The bootstrap with resampling both cluster and within-cluster observations does not work very well because it gives the largest standard deviation.

Type	SD
Sandwich	0.117
Bootstrap: resampling only cluster	0.119
Bootstrap: cluster and observation	0.219
Bootstrap: ignore cluster	0.196

Appendix

```
## 571 HW4

## PR2
library("nlme")
data <- read.csv("pain.csv")
head(data)
attach(data)
# intervals (0-3 hours, 4-6 hours, 7-12 hours, 13-24 hours)
# intervals = 1:4

reshape <- matrix(nrow = nrow(data), ncol = 4)
reshape[, 1] <- rowMeans(data[, 3:5])
reshape[, 2] <- rowMeans(data[, 6:8])
reshape[, 3] <- rowMeans(data[, 9:14])
reshape[, 4] <- rowMeans(data[, 15:26])

time <- rep(1:4, each = nrow(data))
response <- c(reshape)
id.long <- rep(id, 4)
group.long <- rep(group, 4)
data.long = data.frame(id.long,time,group.long,response)

# (i) no random effects with unstructured homoscedastic residuals,
# (ii) random intercepts with independent homoscedastic residuals,
# (iii) independent random intercepts and slopes with independent homoscedastic residuals,
# (iv) correlated random intercepts and slopes with independent homoscedastic residuals,
# (v) no random effects with homoscedastic AR-1 residuals,
# (vi) no random effects with unstructured heteroscedastic residuals,
# (vii) random intercepts with independent heteroscedastic residuals,
# (viii) independent random intercepts and slopes with independent heteroscedastic residuals,
# (ix) correlated random intercepts and slopes with independent heteroscedastic residuals,
# (x) no random effects with heteroscedastic AR-1 residuals.

# 2a.a)
# hours
plot( colMeans(data[which(data$group == "A"), 3:26])~ c(1:24), col=1,type = "l" ,
      pch=19,
      lty = 1,xlab="Hours", ylab="Number of Pills",
     ylim=c(0.5, 3.5),main="Rate of rescue pain medication usage vary with time"
     )
lines(colMeans(data[which(data$group == "B"), 3:26]) ~ c(1:24), col=2, lty = 2)
lines(colMeans(data[which(data$group == "C"), 3:26]) ~ c(1:24), col=3, lty = 5)
legend("topright",
       legend=c("Treatment Group A", "Treatment Group B", "Placebo"),
       col=1:3,
       bty="n",
       lty=c(1,2,5),
       cex=0.8)

## comment:
# 1) The number of pills used decreases during the inverval between 0 and 19 hours.
# 2) We also see an increase in the number of pills used between 19 and 24 hours.
```

```
# 3) The number of pills used in the placebo group are generally larger than the other two treatment gro
# 2a.b)
# time (interval)
# # comment:
# Now draw a plot in which the x-axis is interval instead of hours.
# We see that the relationship is linear for each of the three groups. The number of pills generally dec
# However, the number of pills used in the placebo group are generally larger than the other two treatme
# 2a.c)
# According to the second plot, the temporal trend is negative
# 2a.d) We see the temporal trends are different between the treatment groups and placebo group, but are
# different between treatments A and B.

plot(x = 1:4, y = colMeans(reshape[group == "A", ] ), col = 1, lty = 1, xaxt="n",
     ylim = c(1,4),
     type = "l", xlab = "Interval", ylab = "Number of Pills",main="Rate of rescue pain medication usage
axis(1, at=c(1, 2, 3, 4), labels=c("0-3", "4-6", "7-12", "13-24"))
lines(x = 1:4, y = colMeans(reshape[group == "B", ] ), lty = 2,col = 2)
lines(x = 1:4, y = colMeans(reshape[group == "C", ] ), lty = 5,col = 3)
legend("topright",
       legend=c("Treatment Group A", "Treatment Group B", "Placebo"),
       col=1:3,
       bty="n",
       lty=c(1,2,5),
       cex=1)
# 2a.e)
library(lme4)
a1=xyplot(response ~ factor(time) | factor(id.long), data=data.long,
          main="Trend by subject", ylab="rescue pain medication usage",
          panel = function(x,y) {
            j <- panel.number()
            panel.xyplot(x,y)
            })
print(a1)

## In each group, I made the plot of number of pills over time for each patient.
# They show very different linear relationship.

for (k in c("A","B","C")){
  ind.a=which(data[,1]==k)
  pat.a= unique(data[ind.a,2])
  n.a = length(pat.a)
  par(mfrow = c(5,8),oma=c(2,2,2,2),mar=c(2,2,2,2))
  pp=n.a

    for (i in 1:(pp-pp%%40)){
      plot(1:24,c(data[ind.a[i],-c(1,2)]),pch = 19,"l",
  #          xaxt="n",
  #          yaxt="n",
           cex = 0.5)
    }
  title(paste("Patients in treatment",k,"\nNumber of pills v.s. hours"),out = T)
}
```

```r
# 2C
res = lm(response~time*group.long, data = data.long)$residuals
uid = unique( data.long$id.long)
S = matrix(0,4,4)
nrow(data)
for(i in 1:nrow(data)){
  S = S + res[ data.long$id.long==uid[i]]%*%t(res[ data.long$id.long==uid[i]])
}
Sigma = S/nrow(data)
R = cov2cor(S)
Sigma
R

# 2D
data.long$group.long = factor( data.long$group.long, c("C", "A", "B")) # ref. level placebo
model10 = gls(response~time*group.long,
              corr = corExp(form = ~ 1+time|id.long),
              weight=varIdent(form= ~1|time),
              method = "ML",
              data = data.long)
intervals(model10)

# AIC
#(i) no random effects with unstructured homoscedastic residuals
model1 = gls(response~time*group.long,
             corr = corSymm(form = ~1|id.long),
             weight=NULL,
             method = "ML",data =  data.long)
#(ii) random intercepts with independent homoscedastic residuals
model2 = lme(response~time*group.long,
             random=reStruct(~1|id.long, pdClass="pdDiag"), #intercept random.  pdClass="pdDiag": ind
             weight=NULL,
             method="ML",
             data =  data.long)
#(iii) independent random intercepts and slopes with independent homoscedastic residuals
model3 = lme(response~time*group.long,
             random=reStruct(~1+time|id.long, pdClass="pdDiag"),
             weight=NULL,
             method="ML",
             data =  data.long)
#(iv) correlated random intercepts and slopes with independent homoscedastic residuals
model4 = lme(response~time*group.long,
             random=reStruct(~1+time|id.long, pdClass="pdSymm"), # unstructured random
             weight=NULL,
             method="ML",
             data =  data.long)
#(v) no random effects with homoscedastic AR-1 residuals
model5 = gls(response~time*group.long,
             corr = corExp(form = ~1+time|id.long),
             weight=NULL,
             method = "ML",
             data =  data.long)
```

```
#(vi) no random effects with unstructured heteroscedastic residuals
model6 = gls(response~time*group.long,
             corr = corSymm(form = ~1|id.long),
             weight=varIdent(form=~1|time),
             method = "ML",
             data =  data.long)
#(vii) random intercepts with independent heteroscedastic residuals
model7 = lme(response~time*group.long,
             random=reStruct(~1|id.long, pdClass="pdDiag"),
             weight=varIdent(form=~1|time),
             method="ML",
             data =  data.long)
#(viii) independent random intercepts and slopes with independent heteroscedastic residuals
model8 = lme(response~time*group.long,
             random=reStruct(~1+time|id.long, pdClass="pdDiag"),
             weight=varIdent(form=~1|time),
             method="ML",
             data =  data.long)
#(ix) correlated random intercepts and slopes with independent heteroscedastic residuals
model9 = lme(response~time*group.long,
             random=reStruct(~1+time|id.long, pdClass="pdSymm"),
             weight=varIdent(form=~1|time),
             method="ML",
             data =  data.long)
#(x) no random effects with heteroscedastic AR-1 residuals
model10 = gls(response~time*group.long,
              corr = corExp(form = ~1+time|id.long),
              weight=varIdent(form=~1|time),
              method = "ML",
              data =  data.long)
# AIC
for (i in 1:10)
  print(eval(parse(text=paste("AIC(model",i,")",sep=""))))



##pr3

data(Orthodont)
data = Orthodont
Y = distance
X = matrix(nrow = nrow(Orthodont),ncol = 4)
X[,1] = rep(1,nrow(Orthodont))
X[,2] = age - 8
X[,3] = as.numeric(Sex) - 1
X[,4] = (age - 8) * (as.numeric(Sex) - 1)
# uncorrelated homoscedastic covariance
fit.uncorrelated <- gls(distance ~ I(age-8)*Sex,method="REML",data=Orthodont)
beta.uncorrelated = coef(fit.uncorrelated)
sigma2.uncorrelated = (fit.uncorrelated$sigma)^2

# unstructured homoscedastic covariance
fit.unstructured <- gls(distance ~ I(age-8)*Sex,method="REML",
```

```
                        corr=corSymm(form = ~1 | Subject),
                        data=Orthodont)
est.beta = coef(fit.unstructured)
beta.unstructured = coef(fit.unstructured)
sigma2.unstructured = (fit.unstructured$sigma)^2

# Uncorrelated
par = beta.uncorrelated
res = Y - X %*% par[1:4]
ressq = res %*% t(res) # Cross residual
R = diag(rep((fit.uncorrelated$sigma)^2,4))
library(Matrix)
R = as.matrix( bdiag(replicate(n=n,R,simplify=FALSE)) )
temp = matrix(1,nrow = 4,ncol = 4)
ressq = ressq * kronecker(diag(nrow =  27),temp)
var1 = solve(t(X) %*% solve(R) %*% X) %*% (t(X) %*% solve(R) %*% ressq %*%
                                           solve(R) %*% X) %*% solve(t(X) %*% solve(R) %*% X)
se1 = sqrt(var1[4,4])

# Symmetric
par = beta.unstructured
res = Y - X %*% par[1:4]
ressq = res %*% t(res)
temp = matrix(1,nrow = 4,ncol = 4)
ressq = ressq * kronecker(diag(nrow =  27),temp)
R = matrix(c(getVarCov(fit.unstructured)),m,m)
R = as.matrix( bdiag(replicate(n=n,R,simplify=FALSE)) )
var2 = solve(t(X) %*% solve(R) %*% X) %*% (t(X) %*% solve(R) %*% ressq %*%
                                           solve(R) %*% X) %*% solve(t(X) %*% solve(R) %*% X)
se2 = sqrt(var2[4,4])

# Bootstrap only clusters
cluster.beta1 = c()
cluster.beta2 = c()
for(i in 1:10000){
  id = matrix(nrow = 27,ncol = 4)
  id[,1] = 4 * sample( 27,size = 27,replace = TRUE) - 3
  id[,2] = id[,1] + 1
  id[,3] = id[,1] + 2
  id[,4] = id[,1] + 3
  id = c(t(id))
  new.dist = Orthodont$distance[id]
  new.age = Orthodont$age[id]
  new.sex = Orthodont$Sex[id]
  new.Sub = rep(1:27,each = 4)
  result1 = gls(new.dist ~ I(new.age - 8) * new.sex,method = "REML")
  cluster.beta1[i] = result1$coefficients[4]
  result2 = gls(new.dist ~ I(new.age - 8) * new.sex,method = "REML",corr = corSymm(form = ~ 1 | new.Sub)
  cluster.beta2[i] = result2$coefficients[4]
  print(i)
}
par(mfrow = c(1,2))
hist(cluster.beta1,main = "Uncorrelated")
```

```
hist(cluster.beta2,main = "Unstructured")
par(mfrow = c(1,2))
qqnorm(scale(cluster.beta1), main="Uncorrelated")
abline(0,1,col="red")
qqnorm(scale(cluster.beta1), main="Unstructured")
abline(0,1,col="red")
par(mfrow = c(1,1))
est = c();est2 = c()
n.sample = seq(1,10000,length.out=5000)
for(i in n.sample){
  est=c(est, mean(cluster.beta1[1:i]))
  est2=c(est2, mean(cluster.beta2[1:i]))
}
plot(n.sample,est,xlab="Sample Size",ylab="Bootstrap Estimate",type="l")
lines(n.sample,est2,lty=3)


sd(cluster.beta1)
sd(cluster.beta2)
# Bootstrap that resamples observations
clusterobs.beta = c()
for(i in 1:10000){
  id = sample(108,size = 108,replace = TRUE)
  new.dist = Orthodont$distance[id]
  new.age = Orthodont$age[id]
  new.sex = Orthodont$Sex[id]
  new.Sub = rep(1:27,each = 4)
  result1 = gls(new.dist ~ I(new.age - 8) * new.sex,method = "REML")
  clusterobs.beta[i] = result1$coefficients[4]
  print(i)
}
sd(clusterobs.beta)
# Bootstrap that resamples clusters and observations
obs.beta = c()
for(i in 1:5000){
  id = matrix(nrow = 27,ncol = 4) for(j in 1:27){
    sub = sample( 27,size = 1,replace = TRUE)
    id[j,] = sample((4 * sub - 3):(4 * sub),size = 4,replace = TRUE) }
  id = c(t(id))
  new.dist = Orthodont$distance[id]
  new.age = Orthodont$age[id]
  new.sex = Orthodont$Sex[id]
  new.Sub = rep(1:27,each = 4)
  result1 = gls(new.dist ~ I(new.age - 8) * new.sex,method = "REML")
  obs.beta[i] = result1$coefficients[4]
  print(i)
}
sd(obs.beta)
```