Homework 7

Tim Vigers

25 April 2019

1. Linear mixed model

a. Provide an interpretation for each parameter in the fixed effects model.

$$\beta_0 = \text{Average gall bladder volume for a dog in the control group at time 0}$$
$$\beta_1 = \text{The average effect of a "1 unit increase" in time in the control group, holding everything else constant}$$
$$\beta_2 = \text{The average effect of treatment at time 0, holding everything else constant}$$
$$\beta_3 = \text{The interaction effect of time and treatment}$$

b. Is time being treated as continuous or categorical in this model?

Time is being treated as continuous, because there is one term for time, rather than a term for each timepoint.

c. Random intercept model

$$\mathbf{R}_i = \begin{pmatrix} \sigma_e^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_e^2 \end{pmatrix}_{r_i \times r_i}$$

d. Model without random intercept

$$\mathbf{R}_i = \begin{pmatrix} \sigma_I^2 + \sigma_e^2 & \dots & \sigma_I^2 \\ \vdots & \ddots & \vdots \\ \sigma_I^2 & \dots & \sigma_I^2 + \sigma_e^2 \end{pmatrix}_{r_i \times r_i}$$

2. Dog data

```
# Read in data
dogs <- read.csv("/Users/timvigers/Documents/School/Biostatistical Methods 2/Homeworks/H
omework 7/dogdata_long.csv")
# Format columns
dogs$minutes <- as.factor(dogs$minutes)
# Models
mod <- lme(GBV ~ trt * minutes, random = ~1|id,data = dogs)
mean_mod <- lme(GBV ~ trt:minutes-1, random = ~1|id,data = dogs)
kable(summary(mean_mod)$tTable)
```

	Value	Std.Error	DF	t-value	p-value
trtCholechystokynin:minutes0	19.70500	2.808738	58	7.015606	0.00e+00
trtClanobutin:minutes0	15.82000	2.808738	58	5.632423	5.00e-07
trtControl:minutes0	16.67167	2.808738	58	5.935643	2.00e-07
trtCholechystokynin:minutes30	13.64667	2.808738	58	4.858647	9.30e-06
trtClanobutin:minutes30	12.66000	2.808738	58	4.507363	3.24e-05
trtControl:minutes30	16.30333	2.808738	58	5.804505	3.00e-07
trtCholechystokynin:minutes60	15.24167	2.808738	58	5.426518	1.20e-06
trtClanobutin:minutes60	13.89000	2.808738	58	4.945282	6.80e-06
trtControl:minutes60	16.71167	2.808738	58	5.949885	2.00e-07
trtCholechystokynin:minutes90	17.56333	2.808738	58	6.253105	1.00e-07
trtClanobutin:minutes90	14.87000	2.808738	58	5.294193	1.90e-06
trtControl:minutes90	16.51500	2.808738	58	5.879865	2.00e-07
trtCholechystokynin:minutes120	19.14833	2.808738	58	6.817415	0.00e+00
trtClanobutin:minutes120	15.12833	2.808738	58	5.386168	1.40e-06
trtControl:minutes120	16.74833	2.808738	58	5.962939	2.00e-07

a. Why might adding a random intercept for dogs (relative to the same model but without the random term) help the model?

Each dog is being measured multiple times, so the random intercept for dogs will help account for correlation between these measurements.

b. Is there a mean difference between the two drug groups (Cholechystokynin and Clanobutin) for at least one time point?

$$H_0 : (\mu_{1|0} = \mu_{2|0}, \mu_{1|30} = \mu_{2|30}, \mu_{1|60} = \mu_{2|60}, \mu_{1|90} = \mu_{2|90}, \mu_{1|120} = \mu_{2|120})$$

```
# Contrast
cmat <- rbind(c(-1,1,0,0,0,0,0,0,0,0,0,0,0,0,0),
              c(0,0,0,-1,1,0,0,0,0,0,0,0,0,0,0),
              c(0,0,0,0,0,0,-1,1,0,0,0,0,0,0,0),
              c(0,0,0,0,0,0,0,0,0,-1,1,0,0,0,0),
              c(0,0,0,0,0,0,0,0,0,0,0,0,-1,1,0))
numerator <- cmat %*% summary(mean_mod)$tTable[,1]
f1 <- t(numerator) %*% solve(cmat %*% vcov(mean_mod) %*% t(cmat)) %*% numerator / 5 # Di
vide by # rows in contrast matrix
f1
```

```
##          [,1]
## [1,] 6.902941
```

```
pf(f1,5,((nrow(dogs)-1)-(nrow(summary(mean_mod)$tTable)-1)-(length(unique(dogs$id))-1)),
lower.tail=FALSE)
```

```
##              [,1]
## [1,] 4.076914e-05
```

There is a mean difference between the two drugs for at least one timepoint (F = 6.9, p <0.0001).

c. Does the difference between the two drug groups change over time?

$$H_0 : (\mu_{1|0} - \mu_{2|0} = \mu_{1|30} - \mu_{2|30} = \mu_{1|60} - \mu_{2|60} = \mu_{1|90} - \mu_{2|90} = \mu_{1|120} - \mu_{2|120})$$

```
# Contrast
cmat <- rbind(c(-1,1,0,1,-1,0,0,0,0,0,0,0,0,0,0),
              c(0,0,0,-1,1,0,1,-1,0,0,0,0,0,0,0),
              c(0,0,0,0,0,0,-1,1,0,1,-1,0,0,0,0),
              c(0,0,0,0,0,0,0,0,0,-1,1,0,1,-1,0))
numerator <- cmat %*% summary(mean_mod)$tTable[,1]
f1 <- t(numerator) %*% solve(cmat %*% vcov(mean_mod) %*% t(cmat)) %*% numerator / 4 # Di
vide by # rows in contrast matrix
f1
```

```
##          [,1]
## [1,] 8.521355
```

```
pf(f1,4,((nrow(dogs)-1)-(nrow(summary(mean_mod)$tTable)-1)-(length(unique(dogs$id))-1)),
lower.tail=FALSE)
```

```
##              [,1]
## [1,] 1.767402e-05
```

The mean difference between the two drugs changes over time (F = 8.52, p <0.0001).

d. Estimate and 95% confidence interval for the mean change in GBV from baseline to 60 minutes after for the Cholechystokynin group.

```
# Contrast
cmat <- c(-1,0,0,0,0,0,1,0,0,0,0,0,0,0,0)
numerator <- cmat %*% summary(mean_mod)$tTable[,1]
numerator
```

```
##           [,1]
## [1,] -4.463333
```

```
se <- sqrt(diag(cmat %*% vcov(mean_mod) %*% as.matrix(cmat)))
ci <- c(numerator - 1.96*se,numerator + 1.96*se)
ci
```

```
## [1] -5.403140 -3.523526
```

```
f1 <-
  numerator %*% solve(cmat %*% vcov(mean_mod) %*% as.matrix(cmat)) %*% t(numerator)
f1
```

```
##          [,1]
## [1,] 86.64695
```

```
pf(f1,1,((nrow(dogs)-1)-(nrow(summary(mean_mod)$tTable)-1)-(length(unique(dogs$id))-1)),
lower.tail=FALSE)
```

```
##              [,1]
## [1,] 4.124327e-13
```

```
t1 <- sqrt(f1)
t1
```

```
##          [,1]
## [1,] 9.308435
```

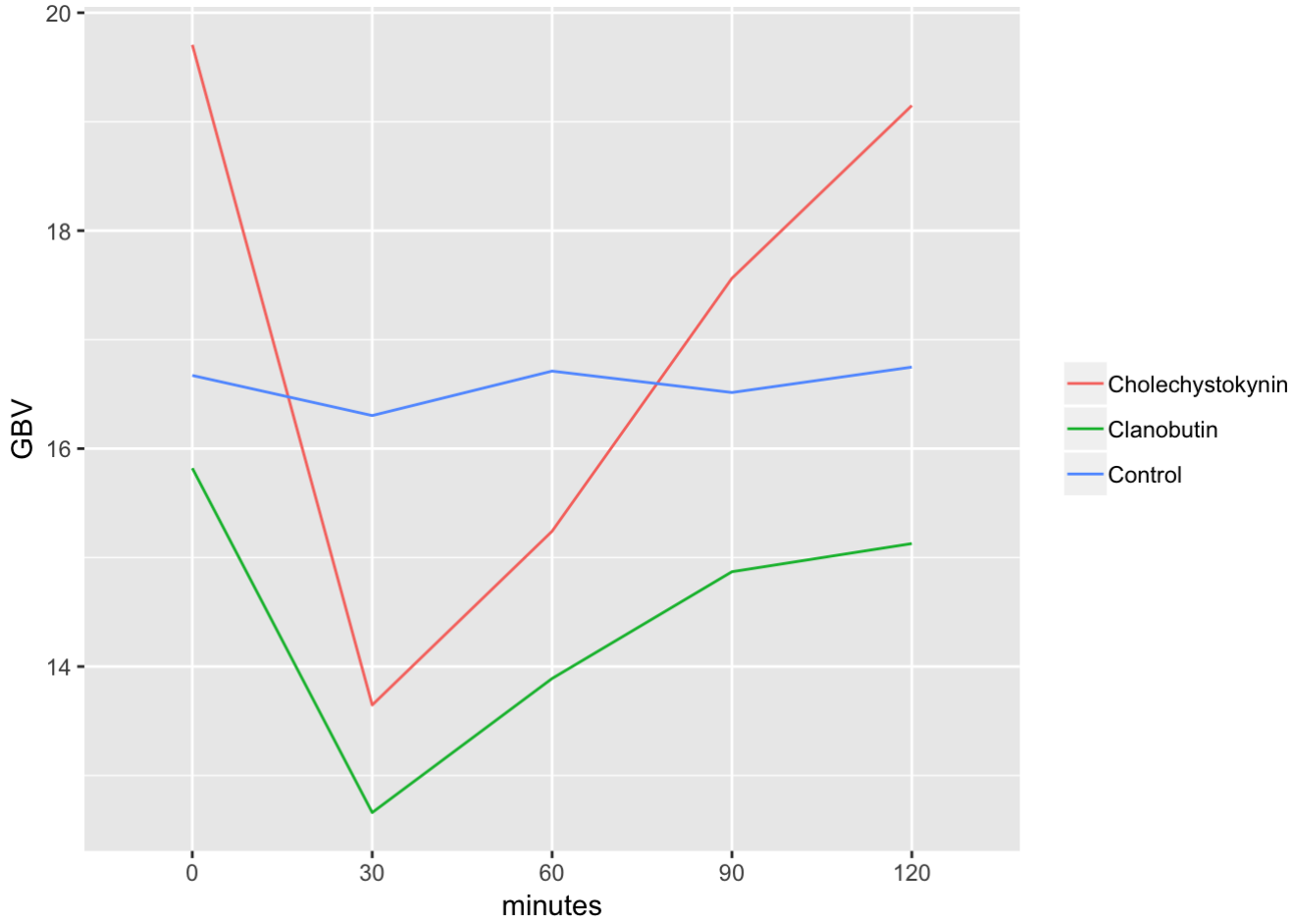
```
pt(t1,((nrow(dogs)-1)-(nrow(summary(mean_mod)$tTable)-1)-(length(unique(dogs$id))-1)),lo
wer.tail=FALSE)
```

```
##              [,1]
## [1,] 2.062164e-13
```

On average GBV decreased from baseline to the 60 minute timepoint for the Cholechystokynin group by 4.463 (95% CI: -5.403, -3.523). This difference was statistically significant ($p < 0.0001$).

e. Summary

```
# Plot
plot <- ggplot(dogs,aes(x=minutes,y=GBV,group=as.factor(trt)))+
  stat_summary(fun.y=mean, geom="line",aes(color = as.factor(trt)))+
  theme(legend.title=element_blank())
plot
```



The aim of this study was to determine the effect of Cholechystokynin and Clanobutin on gall bladder volume over time. We found that there was a significant difference ($p < 0.0001$) between Cholechystokynin and Clanobutin for at least one of the timepoints measured (0, 30, 60, 90, and 120 minutes), and this difference changed over time ($p< 0.0001$). We also tested the average difference in GBV between time 0 and time 60 for those on Cholechystokynin. On average GBV decreased by 4.463 (95% CI: -5.403, -3.523) and this change was statistically significant ($p < 0.0001$).

I would include all of this information and the coefficients of the means model in my report to an investigator, along with the graph above.

f. Time as continuous

When treating time as a continuous variable, there could be an issue with correlation between GBV and time and you would need to account for the collinearity (perhaps by “centering” the time variable with time 60 as the reference time instead of time 0). This issue would increase when modeling higher order time variables, like time^2 or time^3.

Also, if there’s no data at the timepoint the investigator is interested in, then predictions about that timepoint will most likely not be accurate.