# Complex Data lab 4

Anna Zaleska

29 maja 2018

Spaghetti plot of the heart rate over time in 2 groups

```
hrunbalanced <- read.csv(file = "../data/hrunbalanced.csv", header = TRUE)
hrunbalanced$id <- as.factor(hrunbalanced$id)

library(ggplot2)
p <- ggplot(hrunbalanced, aes(x=time, y=hr, color = id)) + geom_line()
p <- p + facet_grid(.~drug, shrink = TRUE, as.table = FALSE)
p + stat_summary(aes(group=drug), fun.y=mean, geom="line", colour="black", size = 1)
```
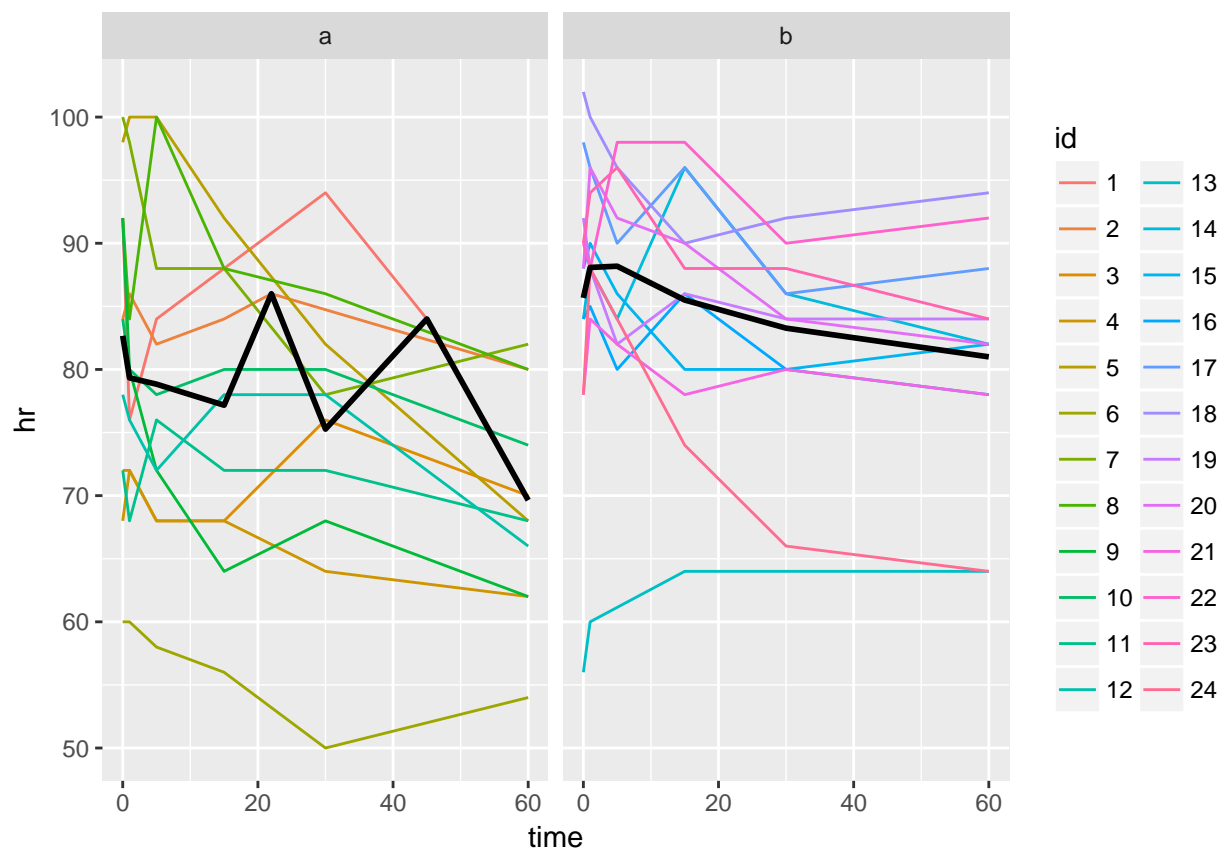


```
## Random intercept
fit.RandInt <- lme(hr ~ time*factor(drug),
random=~1|id,
data=hrunbalanced)
##Random intercept and slope
fit.RandSlp <- lme(hr ~ time*factor(drug),
random=~1+time|id,
data=hrunbalanced)
```

```
anova(fit.RandInt, fit.RandSlp)
```

```
##              Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## fit.RandInt     1  6 945.3558 962.9193 -466.6779
## fit.RandSlp     2  8 937.4996 960.9176 -460.7498 1 vs 2 11.85621  0.0027
```

```
varCorr <- VarCorr(fit.RandSlp)
covar <- getVarCov(fit.RandSlp)
```

Task 1

```
print(xtable(hrunbalanced[hrunbalanced$id %in% c(1,2,13),], caption = "Observations taken into account"))
```

id	drug	hr	time	t
1	a	92	0	1.00
1	a	76	1	2.00
1	a	84	5	3.00
1	a	88	15	4.00
1	a	94	30	5.00
1	a	84	45	5.50
2	a	84	0	1.00
2	a	86	1	2.00
2	a	82	5	3.00
2	a	84	15	4.00
2	a	86	22	4.50
2	a	80	60	6.00
13	b	56	0	1.00
13	b	60	1	2.00
13	b	64	15	4.00
13	b	64	60	6.00

Table 1: Observations taken into account

```
coefs <- fit.RandSlp$coefficients$fixed
#intercept, drug, time, time:drug
coefs <- matrix(c(coefs[1], coefs[3], coefs[2], coefs[4]), ncol = 1)

ones <- matrix(rep(1,6), nrow=6)
zeros <- matrix(rep(0,6), nrow = 6)
tN <- matrix(c(0,1,5,15,30,60), nrow = 6)
t1 <- matrix(hrunbalanced$time[1:6], nrow = 6)
t2 <- matrix(hrunbalanced$time[7:12], nrow = 6)
t13 <- matrix(hrunbalanced$time[73:76], nrow = 4)
x1 <- cbind(ones,zeros,t1,zeros)
x2 <- cbind(ones,zeros, t2,zeros)
x13 <- cbind(ones[1:4,],ones[1:4,],t13,t13)
xNA <- cbind(ones, zeros, tN, zeros)
xNB <- cbind(ones, ones, tN, tN)
ZN <- cbind(ones,tN)
Z1 <- cbind(ones,t1)
Z2 <- cbind(ones, t2)
Z13 <- cbind(ones[1:4], t13)
```

Calculated means:

```
means <- data.frame(drugA = as.numeric(xNA%*%coefs),
                    drugB = as.numeric(xNB%*%coefs),
                    Y1 = as.numeric(x1%*%coefs),
                    Y2 = as.numeric(x2%*%coefs),
                    Y13 = c(as.numeric(x13%*%coefs),NA,NA))
print(xtable(means, caption = "Conditional means vectors for considered observations"),  sanitize.text.
```

drugA	drugB	Y1	Y2	Y13
80.51	86.59	80.51	80.51	86.59
80.35	86.48	80.35	80.35	86.48
79.69	86.06	79.69	79.69	85.00
78.05	85.00	78.05	78.05	80.22
75.58	83.40	75.58	76.90	
70.65	80.22	73.11	70.65	

Table 2: Conditional means vectors for considered observations

Calculated covariances:

```
getCov <- function(Z){
  Z%*%covar%*%t(Z) + as.numeric(varCorr[3,1])*diag(1,nrow = dim(Z)[1], ncol = dim(Z)[1])
}

covZN <- as.data.frame(getCov(ZN))
names(covZN) <- NULL
print(xtable(covZN, caption = " $Cov(Y_i)$ "),  sanitize.text.function=function(x){x}, include.rownames=
```

132.15	112.51	110.00	103.72	94.30	75.46
112.51	130.91	109.43	103.27	94.02	75.54
110.00	109.43	126.17	101.47	92.93	75.87
103.72	103.27	101.47	115.98	90.21	76.69
94.30	94.02	92.93	90.21	105.13	77.93
75.46	75.54	75.87	76.69	77.93	99.42

Table 3: $Cov(Y_i)$

```
covZ1 <- as.data.frame(getCov(Z1))
names(covZ1) <- NULL
print(xtable(covZ1, caption = " $Cov(Y_1)$ "),  sanitize.text.function=function(x){x}, include.rownames=
```

132.15	112.51	110.00	103.72	94.30	84.88
112.51	130.91	109.43	103.27	94.02	84.78
110.00	109.43	126.17	101.47	92.93	84.40
103.72	103.27	101.47	115.98	90.21	83.45
94.30	94.02	92.93	90.21	105.13	82.02
84.88	84.78	84.40	83.45	82.02	99.61

Table 4: $Cov(Y_1)$

```r
covZ2 <- as.data.frame(getCov(Z2))
names(covZ2) <- NULL
print(xtable(covZ2, caption = " $Cov(Y_2)$ "),  sanitize.text.function=function(x){x}, include.rownames=
```

132.15	112.51	110.00	103.72	99.32	75.46
112.51	130.91	109.43	103.27	98.95	75.54
110.00	109.43	126.17	101.47	97.48	75.87
103.72	103.27	101.47	115.98	93.81	76.69
99.32	98.95	97.48	93.81	110.25	77.27
75.46	75.54	75.87	76.69	77.27	99.42

Table 5: $Cov(Y_2)$

```r
covZ13 <- as.data.frame(getCov(Z13))
names(covZ13) <- NULL
print(xtable(covZ13, caption = " $Cov(Y_13)$ "),  sanitize.text.function=function(x){x}, include.rowname
```

132.15	112.51	103.72	75.46
112.51	130.91	103.27	75.54
103.72	103.27	115.98	76.69
75.46	75.54	76.69	99.42

Table 6: $Cov(Y_13)$

Task 2

```r
hrunbalanced <- read.csv(file = "../data/hrunbalanced.csv", header = TRUE)
#step one
coef.hr <- matrix(NA,ncol = 4, nrow = 0)
for (i in unique(hrunbalanced$id)) {
  model <- lm(hr~time, data = hrunbalanced, subset = id == i)
  model.drug <- unique(hrunbalanced[hrunbalanced$id == i, "drug"])
  coef.hr <- rbind(coef.hr, c(i, model.drug, coef(model)))
}
colnames(coef.hr) <- c("id","drug","intercept","slope")
coef.hr <- as.data.frame(coef.hr)
```

```r
# step two
hr.int <- lm(intercept~drug, data = coef.hr)
hr.slp <- lm(slope~drug, data=coef.hr)
print(summary(hr.int))
```

```
##
## Call:
## lm(formula = intercept ~ drug, data = coef.hr)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -27.3357  -4.6227   0.8607   6.2569  19.6306
##
```

4

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   74.353      7.055  10.539 4.61e-10 ***
## drug           6.098      4.462   1.367    0.186
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.93 on 22 degrees of freedom
## Multiple R-squared:  0.07825,    Adjusted R-squared:  0.03635
## F-statistic: 1.868 on 1 and 22 DF,  p-value: 0.1856
```

```
print(summary(hr.slp))
```

```
##
## Call:
## lm(formula = slope ~ drug, data = coef.hr)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.38592 -0.04489  0.02057  0.06470  0.23708
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.21235    0.09128  -2.326   0.0296 *
## drug         0.05332    0.05773   0.924   0.3657
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1414 on 22 degrees of freedom
## Multiple R-squared:  0.03732,    Adjusted R-squared:  -0.006434
## F-statistic: 0.853 on 1 and 22 DF,  p-value: 0.3657
```

We can see from models' summaries that p-values for drug effect in intercept and slope are 0.186 and 0.366 respectively. Thus we conclude that there is no significant difference neither in mean intercepts nor mean slopes in different drugs groups.

```
info.int <- round(summary(hr.int)$coefficients[,c(1,4)], 4)
info.slp <- round(summary(hr.slp)$coefficients[,c(1,4)], 4)
info.mixed <- round(summary(fit.RandSlp)$tTable[,c(1,5)], 4)
```

Looking at Table 7 we can observe some differences in coefficients estimates between the two models. However, the differences do not change the significance of p-values. In both models drug and time*drug effects are not significant.

Table 7: Estimates of coefficients and corresponding p-values in two-stage model and mixed-effects model

	Two-Stage	Mixed Effects
Intercept	74.353(0)	80.514(0)
Time	$-0.2123(0.0296)$	$-0.1645(10^{-4})$
Drug	6.098(0.1856)	6.0748(0.186)
Time*Drug	0.0533(0.3657)	0.0582(0.3041)