Stt864 Lab4

Nan Cao

April 27, 2016

Data preparing

```
setwd("C://Users//nan66//Google Drive//stt864//LAB4")
load(file="Alldata.Rdata")
set.seed(52871775)
library(nlme)
library(lme4)
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'lme4'
```

```
## The following object is masked from 'package:nlme':
##
##     lmList
```

```
library(MASS)
library(glmmML)
```

Q1

```
Goterms<-Alldata$originaldata[,128:130]
mc<-dim(Goterms)[1]
```

```
GOTermlist<-NULL
for (j in 1:mc)
{
    list<-NULL
    for (k in 1:3)
         {
        GOTerm22283<-as.character(Goterms[j,k])
        getGoTerms<-unlist(strsplit(GOTerm22283,"///"))
        repj<-rep(j,length(getGoTerms))
        newlist<-cbind(repj,getGoTerms)
        list<-rbind(list,newlist)
         }
    GOTermlist<-rbind(GOTermlist,list)
}
```

```
GO0006955<-which(GOTermlist[,2]=="GO:0006955")
rownums<-as.numeric(GOTermlist[GO0006955,1])
subsetGenes<-Alldata$originaldata[rownums,c(3:110)]
write.table(subsetGenes,file="subsetGenes.txt")
```

Q2

```
samIDs<-names(subsetGenes)
Beverages<-((samIDs%in%Alldata$trt1)*1
           +(samIDs%in%Alldata$trt2)*2
           +(samIDs%in%Alldata$trt3)*3
           +(samIDs%in%Alldata$trt4)*4)
Subject<-((samIDs%in%Alldata$ind1)*1
         +(samIDs%in%Alldata$ind2)*2
         +(samIDs%in%Alldata$ind3)*3
         +(samIDs%in%Alldata$ind4)*4
         +(samIDs%in%Alldata$ind5)*5
         +(samIDs%in%Alldata$ind6)*6)
hours<-((samIDs%in%Alldata$time_h0)*0
       +(samIDs%in%Alldata$time_h1)*1
       +(samIDs%in%Alldata$time_h2)*2
       +(samIDs%in%Alldata$time_h4)*4
       +(samIDs%in%Alldata$time_h12)*12)
```

Q3

```
BeverFac<-as.factor(Beverages)
hourFac<-as.factor(hours)
resp<-as.numeric(subsetGenes[1,])
lmmd2<-lmer(resp~BeverFac+(1|Subject)+(1|hours))
summary(lmmd2)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: resp ~ BeverFac + (1 | Subject) + (1 | hours)
##
## REML criterion at convergence: 61.1
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.95114 -0.67473 -0.03279  0.69286  2.74099
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 0.089614 0.29936
##  hours    (Intercept) 0.002754 0.05248
##  Residual             0.078447 0.28008
## Number of obs: 108, groups:  Subject, 6; hours, 5
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  5.97213    0.13670   43.69
## BeverFac2    0.12216    0.07891    1.55
## BeverFac3    0.06217    0.07717    0.81
## BeverFac4    0.10770    0.07699    1.40
##
## Correlation of Fixed Effects:
##           (Intr) BvrFc2 BvrFc3
## BeverFac2 -0.296
```

```
## BeverFac3 -0.300  0.519
## BeverFac4 -0.305  0.528  0.533
```

REML estimates of variances:

```
lmreg<-lm(resp~BeverFac+Subject+hours)
anova(lmreg)
```

```
## Analysis of Variance Table
##
## Response: resp
##            Df  Sum Sq Mean Sq F value  Pr(>F)
## BeverFac    3  0.2402 0.08007  0.5192 0.67004
## Subject     1  0.7414 0.74141  4.8068 0.03062 *
## hours       1  0.0836 0.08360  0.5420 0.46330
## Residuals 102 15.7324 0.15424
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(lmreg)
```

```
##
## Call:
## lm(formula = resp ~ BeverFac + Subject + hours)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.75196 -0.31482  0.00507  0.28115  0.90104
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 5.768875   0.113141  50.988   <2e-16 ***
## BeverFac2   0.122258   0.110015   1.111   0.2691
## BeverFac3   0.063637   0.108077   0.589   0.5573
## BeverFac4   0.109990   0.107192   1.026   0.3073
## Subject     0.047577   0.021665   2.196   0.0304 *
## hours       0.006423   0.008724   0.736   0.4633
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3927 on 102 degrees of freedom
## Multiple R-squared:  0.06342,    Adjusted R-squared:  0.0175
## F-statistic: 1.381 on 5 and 102 DF,  p-value: 0.2375
```

```
sigmahat<-sqrt(0.07058)
sigmahat
```

```
## [1] 0.265669
```

```
sigmahat_Subject<-(0.15870-0.07058)/1
sigmahat_Subject
```

```
## [1] 0.08812
```

```
sigmahat_hours<-(0.34036-0.07058)/1
sigmahat_hours
```

```
## [1] 0.26978
```

H_0 : the means are the same; H_1 : the means are different. Statistical model:

$$Z = \frac{\hat{\beta}_1 - \hat{\beta}_2 - 1}{\sqrt{Var(\hat{\beta}_1) + Var(\hat{\beta}_2) - 2Cov(\hat{\beta}_1, \hat{\beta}_2)}} \sim Z$$

For test the significance of difference between alcohol and water group:

```
summary(lmmd2)$coefficients[4,3]
```

```
## [1] 1.39896
```

```
2*(1-pnorm(abs(summary(lmmd2)$coefficients[4,3])))
```

```
## [1] 0.161825
```

Reject the null hypothesis, the means are significantly different.

For test the significance of difference between alcohol and water group:

```
Zstat2<-(summary(lmmd2)$coefficients[3,1]
         -summary(lmmd2)$coefficients[4,1])/sqrt(vcov(lmmd2)[3,3]
         +vcov(lmmd2)[4,4]-2*vcov(lmmd2)[4,3])
Zstat2
```

```
## [1] -0.6115279
```

```
pnorm(abs(Zstat2))
```

```
## [1] 0.7295749
```

We can't reject the null hypothesis.

Q4

```
numgenes<-dim(subsetGenes)[1]
```

```
pvalset<-rep(0,numgenes)
pvalset2<-rep(0,numgenes)
for (i in 1:numgenes)
{
 resp<-as.numeric(subsetGenes[i,])
 lmmd2<-lmer(resp~BeverFac+(1|Subject)+(1|hours))
```

```
 Zstat<-summary(lmmd2)$coefficients[4,3]
 Zstat2<-(summary(lmmd2)$coefficients[3,1]
          -summary(lmmd2)$coefficients[4,1])/sqrt(vcov(lmmd2)[3,3]
          +vcov(lmmd2)[4,4]-2*vcov(lmmd2)[4,3])
 pvalset[i]<-2*(1-pnorm(abs(Zstat)))
 pvalset2[i]<-2*(1-pnorm(abs(Zstat)))
}
```

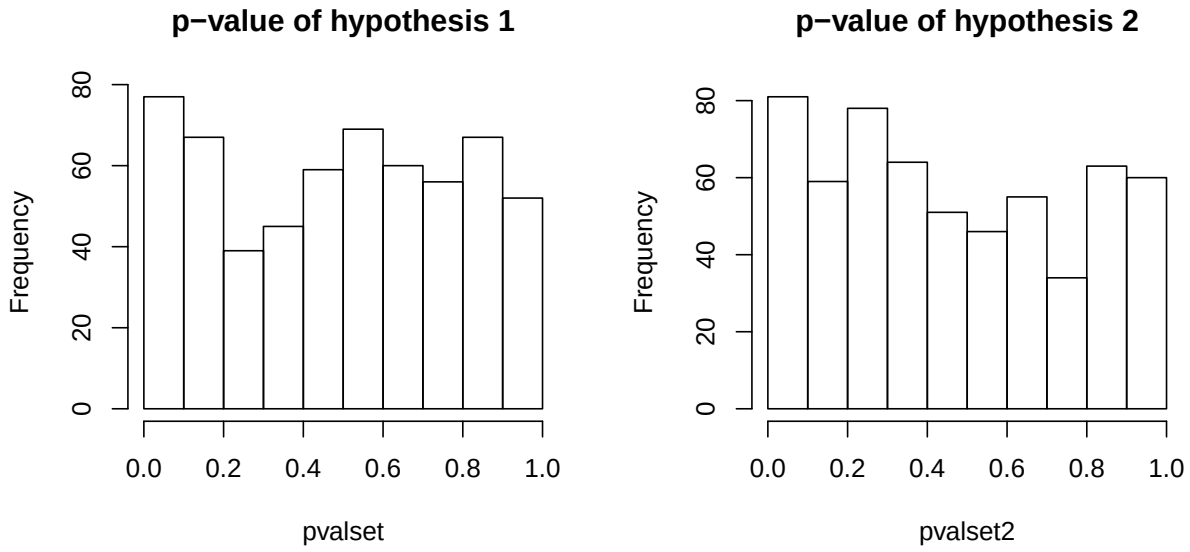
```
smallest<-which.min(pvalset)
smallest
```

```
## [1] 416
```

```
smallest2<-which.min(pvalset2)
smallest2
```

```
## [1] 119
```

```
oldpar <- par(mfrow = c(1, 2))
hist(pvalset,main="p-value of hypothesis 1")
hist(pvalset2,main="p-value of hypothesis 2")
```




```
par(oldpar)
0.05/numgenes
```

```
## [1] 8.460237e-05
```

```
min(pvalset)
```

```
## [1] 1.696569e-05
```

```r
min(pvalset2)
```

```
## [1] 0.0001308026
```

Since there's no gene significant for the second hypothesis, there's no gene significant for both hypothesises.

```r
newresp<-as.numeric(subsetGenes[smallest,])
lmmd3<-lmer(newresp~BeverFac+(1|Subject)+(1|hours))
summary(lmmd3)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: newresp ~ BeverFac + (1 | Subject) + (1 | hours)
##
## REML criterion at convergence: 12.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.4003 -0.6617 -0.0362  0.6129  3.3348
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 0.026456 0.16265
##  hours    (Intercept) 0.003286 0.05732
##  Residual             0.050135 0.22391
## Number of obs: 108, groups:  Subject, 6; hours, 5
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  9.02833    0.08432  107.07
## BeverFac2   -0.28806    0.06310   -4.56
## BeverFac3   -0.19812    0.06170   -3.21
## BeverFac4   -0.26476    0.06155   -4.30
##
## Correlation of Fixed Effects:
##           (Intr) BvrFc2 BvrFc3
## BeverFac2 -0.384
## BeverFac3 -0.389  0.519
## BeverFac4 -0.395  0.528  0.533
```

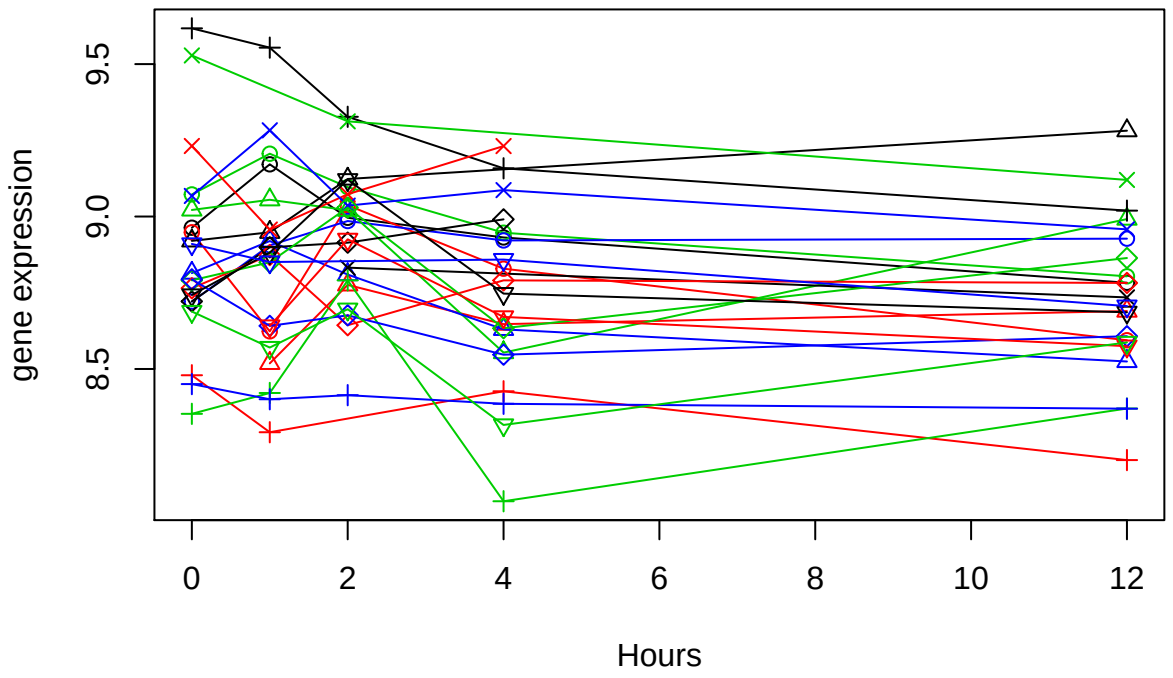
```r
newresp2<-as.numeric(subsetGenes[smallest2,])
lmmd32<-lmer(newresp2~BeverFac+(1|Subject)+(1|hours))
summary(lmmd32)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: newresp2 ~ BeverFac + (1 | Subject) + (1 | hours)
##
## REML criterion at convergence: -91.5
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.14696 -0.69569 -0.00443  0.55191  3.06336
```

```
##
## Random effects:
##  Groups   Name        Variance  Std.Dev.
##  Subject  (Intercept) 0.0015740 0.03967
##  hours    (Intercept) 0.0003597 0.01897
##  Residual             0.0202629 0.14235
## Number of obs: 108, groups:  Subject, 6; hours, 5
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  5.65034    0.03397  166.35
## BeverFac2   -0.03207    0.04001   -0.80
## BeverFac3    0.05852    0.03920    1.49
## BeverFac4   -0.08602    0.03901   -2.21
##
## Correlation of Fixed Effects:
##           (Intr) BvrFc2 BvrFc3
## BeverFac2 -0.603
## BeverFac3 -0.612  0.519
## BeverFac4 -0.619  0.526  0.533
```

Q5

```
plot(hours,newresp,type="n",xlab="Hours", ylab="gene expression")
for (sub in 1:6)
 for (bever in 1:4)
{
points(
      hours[(Subject==sub)&(Beverages==bever)],
      newresp[(Subject==sub)&(Beverages==bever)],
      col=bever,pch=sub
      )
lines(
      hours[(Subject==sub)&(Beverages==bever)],
      newresp[(Subject==sub)&(Beverages==bever)],
      col=bever
      )
}
```



There're obvious changes of gene expression in the first 2 hours, but the change from the 2nd hour to the 12th hour.

Q6

```
resp<-newresp
hours2<-(hours^2)
hours3<-(hours^3)
combfacs<-Subject*10+Beverages
glscsh<-gls(
    resp~BeverFac+hours+hours2+hours3,
    correlation=corCompSymm(form=~1|combfacs),
    weights=varIdent(form=~1|combfacs),
    method="REML"
    )
summary(glscsh)
```

```
## Generalized least squares fit by REML
##   Model: resp ~ BeverFac + hours + hours2 + hours3
##   Data: NULL
##        AIC      BIC   logLik
##   25.99752 109.6814 19.00124
##
## Correlation Structure: Compound symmetry
##  Formula: ~1 | combfacs
##  Parameter estimate(s):
##       Rho
```

```
## 0.7332594
## Variance function:
##  Structure: Different standard deviations per stratum
##  Formula: ~1 | combfacs
##  Parameter estimates:
##        11        12        13        14        21        22        23
## 1.0000000 1.8295252 1.0600585 1.4341701 2.4060559 1.4688175 1.9886239
##        24        31        32        33        34        41        42
## 1.0054293 2.4793686 1.6586804 2.8338081 1.2161911 0.4534302 2.1202614
##        43        44        51        52        53        54        61
## 2.1389126 1.3764744 1.4416210 1.1997375 1.4298310 0.8834376 1.4674247
##        62        63        64
## 1.2811624 1.7399692 0.6516821
##
## Coefficients:
##                 Value  Std.Error   t-value p-value
## (Intercept)  8.894572 0.06421264 138.51746  0.0000
## BeverFac2   -0.142405 0.10958938  -1.29944  0.1968
## BeverFac3    0.044379 0.11510923   0.38554  0.7006
## BeverFac4   -0.092622 0.08386592  -1.10441  0.2720
## hours        0.038807 0.03466376   1.11951  0.2656
## hours2      -0.019956 0.01082251  -1.84393  0.0681
## hours3       0.001316 0.00068181   1.93009  0.0564
##
##  Correlation:
##           (Intr) BvrFc2 BvrFc3 BvrFc4 hours  hours2
## BeverFac2 -0.497
## BeverFac3 -0.479  0.271
## BeverFac4 -0.663  0.372  0.355
## hours     -0.367  0.039  0.048  0.082
## hours2     0.313 -0.041 -0.047 -0.081 -0.964
## hours3    -0.295  0.043  0.047  0.081  0.940 -0.997
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.04623400 -0.63629640  0.07814485  0.74801217  1.97615550
##
## Residual standard error: 0.1693538
## Degrees of freedom: 108 total; 101 residual
```

```
glsarh1<-gls(
    resp~BeverFac+hours+hours2+hours3,
    correlation=corAR1(form=~1|combfacs),
    weights=varIdent(form=~1|combfacs),
    method="REML"
    )
summary(glsarh1)
```

```
## Generalized least squares fit by REML
##   Model: resp ~ BeverFac + hours + hours2 + hours3
##   Data: NULL
##        AIC      BIC   logLik
##   31.69968 115.3835 16.15016
##
```

```
## Correlation Structure: AR(1)
##  Formula: ~1 | combfacs
##  Parameter estimate(s):
##       Phi
## 0.7529408
## Variance function:
##  Structure: Different standard deviations per stratum
##  Formula: ~1 | combfacs
##  Parameter estimates:
##        11        12        13        14        21        22        23
## 1.0000000 2.1763013 0.7912687 0.8613836 1.7510901 1.2505422 2.1894371
##        24        31        32        33        34        41        42
## 0.7170562 1.9960873 1.9230115 3.3853442 1.1988921 0.5083076 2.4074760
##        43        44        51        52        53        54        61
## 2.0696207 1.5857085 1.0039380 1.2833450 1.6469664 0.8885409 1.4083980
##        62        63        64
## 1.3768529 1.9994487 0.8304457
##
## Coefficients:
##                 Value  Std.Error   t-value p-value
## (Intercept)  8.880734 0.06509153 136.43457  0.0000
## BeverFac2   -0.164592 0.10660794  -1.54390  0.1257
## BeverFac3    0.076963 0.10073119   0.76404  0.4466
## BeverFac4   -0.110683 0.07706837  -1.43617  0.1540
## hours        0.069731 0.03537997   1.97091  0.0515
## hours2      -0.027355 0.00993549  -2.75324  0.0070
## hours3       0.001743 0.00061026   2.85580  0.0052
##
##  Correlation:
##           (Intr) BvrFc2 BvrFc3 BvrFc4 hours  hours2
## BeverFac2 -0.476
## BeverFac3 -0.524  0.301
## BeverFac4 -0.685  0.394  0.422
## hours     -0.386  0.024  0.069  0.093
## hours2     0.278 -0.018 -0.056 -0.077 -0.941
## hours3    -0.242  0.017  0.051  0.070  0.901 -0.994
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.02312422 -0.72618150  0.08860765  0.72481580  2.16428146
##
## Residual standard error: 0.1704837
## Degrees of freedom: 108 total; 101 residual
```

```
glscsM2<-gls(
    resp~BeverFac+BeverFac*hours+BeverFac*hours2+BeverFac*hours3,
    correlation=corCompSymm(form=~1|combfacs),
    method="REML"
    )
summary(glscsM2)
```

```
## Generalized least squares fit by REML
##   Model: resp ~ BeverFac + BeverFac * hours + BeverFac * hours2 + BeverFac *      hours3
##   Data: NULL
```

```
##        AIC      BIC    logLik
##   88.64523 134.0374 -26.32262
##
## Correlation Structure: Compound symmetry
##  Formula: ~1 | combfacs
##  Parameter estimate(s):
##       Rho
## 0.7659703
##
## Coefficients:
##                       Value  Std.Error  t-value p-value
## (Intercept)       8.964483 0.11656416 76.90600  0.0000
## BeverFac2        -0.174336 0.16710384 -1.04328  0.2996
## BeverFac3        -0.077384 0.16216840 -0.47718  0.6344
## BeverFac4        -0.166716 0.16213689 -1.02824  0.3065
## hours             0.125167 0.08320032  1.50440  0.1359
## hours2           -0.044369 0.02675679 -1.65825  0.1007
## hours3            0.002779 0.00170570  1.62898  0.1067
## BeverFac2:hours  -0.195780 0.12363494 -1.58353  0.1167
## BeverFac3:hours   0.043770 0.11385166  0.38445  0.7015
## BeverFac4:hours  -0.086370 0.11533952 -0.74883  0.4559
## BeverFac2:hours2  0.067626 0.03836483  1.76271  0.0813
## BeverFac3:hours2 -0.038707 0.03653247 -1.05953  0.2921
## BeverFac4:hours2  0.026033 0.03666247  0.71008  0.4795
## BeverFac2:hours3 -0.004316 0.00241363 -1.78807  0.0771
## BeverFac3:hours3  0.002915 0.00232392  1.25422  0.2129
## BeverFac4:hours3 -0.001587 0.00232247 -0.68324  0.4962
##
##  Correlation:
##                  (Intr) BvrFc2 BvrFc3 BvrFc4 hours  hours2 hours3 BvrF2:
## BeverFac2        -0.698
## BeverFac3        -0.719  0.501
## BeverFac4        -0.719  0.501  0.517
## hours            -0.386  0.269  0.277  0.277
## hours2            0.305 -0.213 -0.219 -0.219 -0.956
## hours3           -0.279  0.195  0.201  0.201  0.927 -0.996
## BeverFac2:hours   0.260 -0.415 -0.187 -0.187 -0.673  0.644 -0.624
## BeverFac3:hours   0.282 -0.197 -0.360 -0.203 -0.731  0.699 -0.678  0.492
## BeverFac4:hours   0.278 -0.194 -0.200 -0.362 -0.721  0.690 -0.669  0.485
## BeverFac2:hours2 -0.213  0.339  0.153  0.153  0.667 -0.697  0.695 -0.963
## BeverFac3:hours2 -0.224  0.156  0.283  0.161  0.701 -0.732  0.729 -0.471
## BeverFac4:hours2 -0.223  0.155  0.160  0.289  0.698 -0.730  0.727 -0.470
## BeverFac2:hours3  0.198 -0.313 -0.142 -0.142 -0.655  0.704 -0.707  0.937
## BeverFac3:hours3  0.205 -0.143 -0.259 -0.147 -0.681  0.731 -0.734  0.458
## BeverFac4:hours3  0.205 -0.143 -0.148 -0.265 -0.681  0.731 -0.734  0.458
##                  BvrF3: BvrF4: BvF2:2 BvF3:2 BvF4:2 BvF2:3 BvF3:3
## BeverFac2
## BeverFac3
## BeverFac4
## hours
## hours2
## hours3
## BeverFac2:hours
## BeverFac3:hours
```

```
## BeverFac4:hours   0.527
## BeverFac2:hours2 -0.487 -0.481
## BeverFac3:hours2 -0.958 -0.505  0.511
## BeverFac4:hours2 -0.510 -0.961  0.509  0.535
## BeverFac2:hours3  0.479  0.473 -0.996 -0.515 -0.514
## BeverFac3:hours3  0.929  0.491 -0.510 -0.996 -0.533  0.519
## BeverFac4:hours3  0.498  0.934 -0.510 -0.536 -0.996  0.519  0.539
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -2.005018720 -0.574988300 -0.005288275  0.569275357  2.348144215
##
## Residual standard error: 0.2779885
## Degrees of freedom: 108 total; 92 residual
```

```
glscshM2<-gls(
    resp~BeverFac+BeverFac*hours+BeverFac*hours2+BeverFac*hours3,
    correlation=corCompSymm(form=~1|combfacs),
    weights=varIdent(form=~1|combfacs),
    method="REML"
    )
summary(glscshM2)
```

```
## Generalized least squares fit by REML
##   Model: resp ~ BeverFac + BeverFac * hours + BeverFac * hours2 + BeverFac *      hours3
##   Data: NULL
##        AIC      BIC    logLik
##   112.7808 216.1742 -15.39042
##
## Correlation Structure: Compound symmetry
##  Formula: ~1 | combfacs
##  Parameter estimate(s):
##       Rho
## 0.7524762
## Variance function:
##  Structure: Different standard deviations per stratum
##  Formula: ~1 | combfacs
##  Parameter estimates:
##        11        12        13        14        21        22        23
## 1.0000000 1.7085892 1.5050892 1.3943657 2.1031158 1.2377060 0.8244413
##        24        31        32        33        34        41        42
## 1.0028094 2.5042074 1.3940892 1.9695143 1.0877401 0.8183035 1.5689082
##        43        44        51        52        53        54        61
## 2.2197443 1.2184968 0.9731768 0.9990605 0.9785889 0.8022278 1.1073171
##        62        63        64
## 1.4045538 1.0742013 0.4363101
##
## Coefficients:
##                      Value  Std.Error  t-value p-value
## (Intercept)       8.831592 0.09607132 91.92746  0.0000
## BeverFac2        -0.065214 0.14717206 -0.44311  0.6587
## BeverFac3         0.034849 0.13398488  0.26010  0.7954
## BeverFac4         0.011172 0.11422983  0.09781  0.9223
## hours             0.161000 0.07271219  2.21421  0.0293
```

```
## hours2           -0.048792 0.02306691 -2.11525  0.0371
## hours3            0.002915 0.00146541  1.98941  0.0496
## BeverFac2:hours  -0.221505 0.11339768 -1.95335  0.0538
## BeverFac3:hours   0.010425 0.09796734  0.10641  0.9155
## BeverFac4:hours  -0.184226 0.08829503 -2.08648  0.0397
## BeverFac2:hours2  0.068419 0.03486588  1.96235  0.0527
## BeverFac3:hours2 -0.036343 0.03101146 -1.17192  0.2443
## BeverFac4:hours2  0.049836 0.02788309  1.78734  0.0772
## BeverFac2:hours3 -0.004195 0.00218724 -1.91815  0.0582
## BeverFac3:hours3  0.002967 0.00196469  1.50991  0.1345
## BeverFac4:hours3 -0.002940 0.00176293 -1.66756  0.0988
##
##  Correlation:
##                  (Intr) BvrFc2 BvrFc3 BvrFc4 hours  hours2 hours3 BvrF2:
## BeverFac2        -0.653
## BeverFac3        -0.717  0.468
## BeverFac4        -0.841  0.549  0.603
## hours            -0.449  0.293  0.322  0.378
## hours2            0.364 -0.238 -0.261 -0.306 -0.959
## hours3           -0.337  0.220  0.241  0.283  0.931 -0.996
## BeverFac2:hours   0.288 -0.452 -0.207 -0.242 -0.641  0.615 -0.597
## BeverFac3:hours   0.334 -0.218 -0.402 -0.281 -0.742  0.712 -0.691  0.476
## BeverFac4:hours   0.370 -0.242 -0.265 -0.415 -0.824  0.790 -0.767  0.528
## BeverFac2:hours2 -0.241  0.374  0.173  0.203  0.634 -0.662  0.659 -0.965
## BeverFac3:hours2 -0.271  0.177  0.323  0.228  0.713 -0.744  0.741 -0.457
## BeverFac4:hours2 -0.301  0.197  0.216  0.338  0.793 -0.827  0.824 -0.509
## BeverFac2:hours3  0.226 -0.348 -0.162 -0.190 -0.624  0.667 -0.670  0.939
## BeverFac3:hours3  0.251 -0.164 -0.297 -0.211 -0.695  0.743 -0.746  0.445
## BeverFac4:hours3  0.280 -0.183 -0.201 -0.313 -0.774  0.828 -0.831  0.496
##                  BvrF3: BvrF4: BvF2:2 BvF3:2 BvF4:2 BvF2:3 BvF3:3
## BeverFac2
## BeverFac3
## BeverFac4
## hours
## hours2
## hours3
## BeverFac2:hours
## BeverFac3:hours
## BeverFac4:hours   0.611
## BeverFac2:hours2 -0.471 -0.523
## BeverFac3:hours2 -0.960 -0.587  0.492
## BeverFac4:hours2 -0.589 -0.963  0.547  0.615
## BeverFac2:hours3  0.463  0.514 -0.996 -0.496 -0.552
## BeverFac3:hours3  0.933  0.572 -0.492 -0.996 -0.615  0.500
## BeverFac4:hours3  0.575  0.938 -0.548 -0.616 -0.996  0.557  0.620
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.9815152 -0.5724057  0.1097721  0.6862958  1.9471685
##
## Residual standard error: 0.1948148
## Degrees of freedom: 108 total; 92 residual
```

According to the summary of the fitted model ,the last two model seem better. The time is significant only

in the last models.

Q7
Define new data set

```
Combdata<-cbind(newresp,hours,Subject,Beverages,combfacs)
unicombfacs<-Combdata[Combdata[,2]==0,5]
newcombdata<-NULL
for (i in 1:length(unicombfacs))
{
 subcomb<-Combdata[Combdata[,5]==unicombfacs[i],]
 Yvec<-subcomb[,1]
 Hvec<-subcomb[,2]
 Zvec<-((Yvec[Hvec!=0]-Yvec[Hvec==0])>0)+0
 newcombdata<-rbind(newcombdata,cbind(Zvec,subcomb[Hvec!=0,c(2:5)]))
}

newcombdata1<-list(
                   Zvec=newcombdata[,1],
                   hours=newcombdata[,2],
                   Subject=newcombdata[,3],
                   Beverages=newcombdata[,4],
                   combfacs=newcombdata[,5]
                   )
newcombdata1$Beverages<-as.factor(newcombdata1$Beverages)
newcombdata1$hours<-as.factor(newcombdata1$hours)
```

GLMM with the default Laplace approximation

```
glmRandomIntcept1<-glmer(
   newcombdata1$Zvec~newcombdata1$Beverages+newcombdata1$hours
  +(1|newcombdata1$combfacs),family=binomial
  )
summary(glmRandomIntcept1)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## newcombdata1$Zvec ~ newcombdata1$Beverages + newcombdata1$hours +
##     (1 | newcombdata1$combfacs)
##
##      AIC      BIC   logLik deviance df.resid
##     99.1    117.8    -41.5     83.1       69
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.7177 -0.4918 -0.1585  0.4790  1.9837
##
## Random effects:
##  Groups                Name        Variance Std.Dev.
##  newcombdata1$combfacs (Intercept) 3.502    1.871
## Number of obs: 77, groups:  newcombdata1$combfacs, 21
##
```

```
## Fixed effects:
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)               2.1984     1.3218   1.663   0.0963 .
## newcombdata1$Beverages2  -2.7018     1.8156  -1.488   0.1367
## newcombdata1$Beverages3  -1.2715     1.4959  -0.850   0.3953
## newcombdata1$Beverages4  -2.4012     1.6139  -1.488   0.1368
## newcombdata1$hours2      -0.7097     0.8858  -0.801   0.4230
## newcombdata1$hours4      -2.3062     1.0033  -2.299   0.0215 *
## newcombdata1$hours12     -2.3387     1.0012  -2.336   0.0195 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) nw1$B2 nw1$B3 nw1$B4 nwc1$2 nwc1$4
## nwcmbdt1$B2 -0.658
## nwcmbdt1$B3 -0.691  0.529
## nwcmbdt1$B4 -0.714  0.547  0.582
## nwcmbdt1$h2 -0.405  0.104  0.025  0.097
## nwcmbdt1$h4 -0.451  0.189  0.093  0.179  0.507
## nwcmbdt1$12 -0.451  0.193  0.081  0.178  0.514  0.552
```

Another way to fit using Laplace approximation

```
glmRandomIntcept2<-glmmML(
  newcombdata1$Zvec~newcombdata1$Beverages+newcombdata1$hours,
  family=binomial,cluster=newcombdata1$combfacs)
summary(glmRandomIntcept2)
```

```
##
## Call:  glmmML(formula = newcombdata1$Zvec ~ newcombdata1$Beverages +      newcombdata1$hours, family
##
##
##                             coef se(coef)       z Pr(>|z|)
## (Intercept)               2.1983   1.3063  1.6829   0.0924
## newcombdata1$Beverages2 -2.7019   1.7897 -1.5097   0.1310
## newcombdata1$Beverages3 -1.2713   1.4882 -0.8543   0.3930
## newcombdata1$Beverages4 -2.4011   1.5927 -1.5076   0.1320
## newcombdata1$hours2     -0.7095   0.8834 -0.8032   0.4220
## newcombdata1$hours4     -2.3063   0.9878 -2.3348   0.0196
## newcombdata1$hours12    -2.3388   0.9861 -2.3717   0.0177
##
## Scale parameter in mixing distribution:  1.872 gaussian
## Std. Error:                              0.6353
##
##         LR p-value for H_0: sigma = 0:  0.001518
##
## Residual deviance: 83.07 on 69 degrees of freedom    AIC: 99.07
```

Fit the model using the penalized quasi-likelihood

```
newcombdata2<-data.frame(newcombdata1)
glmRandomIntcept3<-glmmPQL(
                          Zvec~Beverages+hours,
```

```
                          random=~1|combfacs,family=binomial,
                          data=newcombdata2
                          )
```

```
## iteration 1

## iteration 2

## iteration 3

## iteration 4

## iteration 5
```

```
summary(glmRandomIntcept3)
```

```
## Linear mixed-effects model fit by maximum likelihood
##  Data: newcombdata2
##    AIC BIC logLik
##     NA  NA     NA
##
## Random effects:
##  Formula: ~1 | combfacs
##         (Intercept)  Residual
## StdDev:    1.916686 0.7552666
##
## Variance function:
##  Structure: fixed weights
##  Formula: ~invwt
## Fixed effects: Zvec ~ Beverages + hours
##                  Value Std.Error DF    t-value p-value
## (Intercept)  2.0951310 1.1708682 53  1.7893825  0.0793
## Beverages2  -2.4573363 1.6274440 17 -1.5099360  0.1494
## Beverages3  -1.1602524 1.4179107 17 -0.8182831  0.4245
## Beverages4  -2.1760696 1.4622332 17 -1.4881823  0.1550
## hours2      -0.7267728 0.6931104 53 -1.0485670  0.2991
## hours4      -2.2640438 0.7463470 53 -3.0335001  0.0037
## hours12     -2.2785129 0.7461441 53 -3.0537169  0.0035
##  Correlation:
##             (Intr) Bvrgs2 Bvrgs3 Bvrgs4 hours2 hours4
## Beverages2 -0.635
## Beverages3 -0.691  0.497
## Beverages4 -0.701  0.488  0.553
## hours2     -0.319  0.053 -0.021  0.049
## hours4     -0.341  0.077  0.019  0.076  0.498
## hours12    -0.341  0.081  0.012  0.076  0.504  0.519
##
## Standardized Within-Group Residuals:
##        Min         Q1        Med         Q3        Max
## -2.6510114 -0.5505002 -0.1868421  0.5903681  2.3682410
##
## Number of Observations: 77
## Number of Groups: 21
```

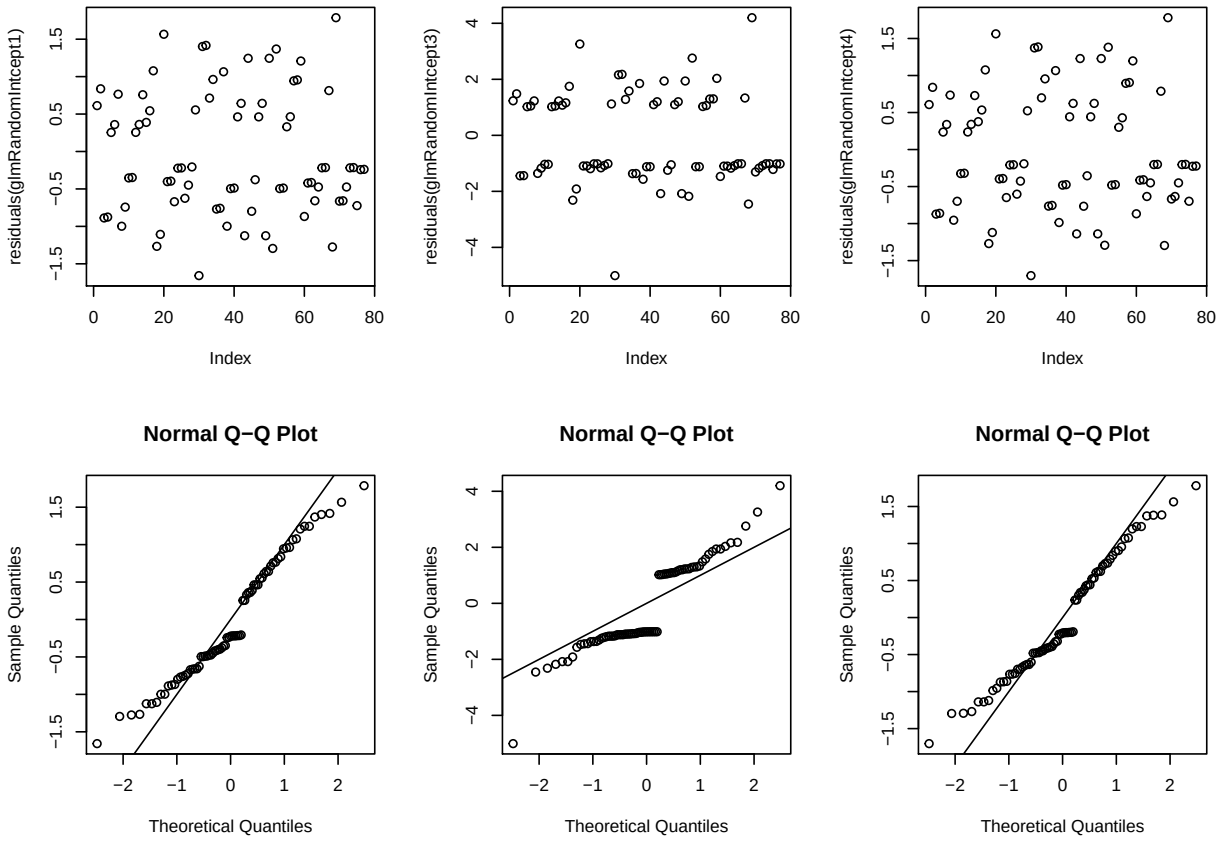
Fit the model using adaptive Gauss-Hermite qudrature(AGQ)
nAGQ: the number of points per axis for evaluating the adaptive Gauss-Hermite approximation to the log-likelihood.

```
glmRandomIntcept4<- glmer(
  newcombdata1$Zvec~newcombdata1$Beverages+newcombdata1$hours+(1|newcombdata1$combfacs),
  nAGQ=8,family=binomial)
summary(glmRandomIntcept4)
```

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 8) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## newcombdata1$Zvec ~ newcombdata1$Beverages + newcombdata1$hours +
##     (1 | newcombdata1$combfacs)
##
##      AIC      BIC   logLik deviance df.resid
##     98.2    117.0    -41.1     82.2       69
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.8063 -0.4718 -0.1481  0.4632  1.9703
##
## Random effects:
##  Groups                 Name        Variance Std.Dev.
##  newcombdata1$combfacs (Intercept) 4.102    2.025
## Number of obs: 77, groups:  newcombdata1$combfacs, 21
##
## Fixed effects:
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)                2.2550     1.3793   1.635   0.1021
## newcombdata1$Beverages2   -2.7558     1.8881  -1.460   0.1444
## newcombdata1$Beverages3   -1.3029     1.5720  -0.829   0.4072
## newcombdata1$Beverages4   -2.4497     1.6815  -1.457   0.1452
## newcombdata1$hours2       -0.7365     0.8993  -0.819   0.4128
## newcombdata1$hours4       -2.3671     1.0285  -2.301   0.0214 *
## newcombdata1$hours12      -2.3966     1.0266  -2.334   0.0196 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) nw1$B2 nw1$B3 nw1$B4 nwc1$2 nwc1$4
## nwcmbdt1$B2 -0.660
## nwcmbdt1$B3 -0.690  0.524
## nwcmbdt1$B4 -0.717  0.539  0.577
## nwcmbdt1$h2 -0.394  0.105  0.018  0.098
## nwcmbdt1$h4 -0.449  0.190  0.093  0.183  0.510
## nwcmbdt1$12 -0.449  0.194  0.082  0.181  0.516  0.569
```

```
oldpar <- par(mfrow = c(2, 3))
plot(residuals(glmRandomIntcept1))
plot(residuals(glmRandomIntcept3))
plot(residuals(glmRandomIntcept4))
```

```
qqnorm(residuals(glmRandomIntcept1))
abline(0,1)
qqnorm(residuals(glmRandomIntcept3))
abline(0,1)
qqnorm(residuals(glmRandomIntcept4))
abline(0,1)
```




```
par(oldpar)
```

Compared among the 3 models, only time effects are signficant in all the models. The penalized quasi-likelihood model abd model using adaptive Gauss-Hermite qudrature have lower correlation of fixed effects. What's more, we failed to get the residuals of the second model because the infinite value.
The first model and the forth model has the same residual plot and qqnorm-plot, they both shows higher normaility compared with the third model.