[redacted] BK lmer


```
    Update of BK repeated measures analysis   2014 version
> library(lme4)
> #note brogkutlong restarts subject numbering at 1 for each method; brogkutlong2 numbers sequentially
> bk  = read.table(file="http://www-stat.stanford.edu/~rag/stat222/brogkutlong2.dat", header = T)
> attach(bk)
> bklist = lmList(outcome ~ time|subject, data = bk) # getting difference scores the hard way
> bklist
Call: lmList(formula = outcome ~ time | subject, data = bk)
Coefficients:
   (Intercept) time
1           54   -3
2           15   20
...
21          54  -11
# if you want the "intercept" to be level at time=1 (pretest) the
> t1 = time - 1
> bklist1 = lmList(outcome ~ t1|subject, data = bk)

> library(lattice) # make a plot for individual subjects
> xyplot(outcome ~ time|subject, groups = method, type = c("p","r"), data = bk)

# the repeated measures anova, shown in previous analysis
> bkrepaov1 = aov(outcome ~ as.factor(time)*as.factor(method)+ Error(as.factor(subject)))
> summary(bkrepaov1)
Error: as.factor(subject)
                  Df Sum Sq Mean Sq F value Pr(>F)
as.factor(method)  1    847   847.5   3.627 0.0721 .
Residuals         19   4440   233.7
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Error: Within
                                  Df Sum Sq Mean Sq F value   Pr(>F)
as.factor(time)                    1  542.9   542.9   15.14 0.000982 ***
as.factor(time):as.factor(method)  1  407.4   407.4   11.36 0.003209 **
Residuals                         19  681.2    35.9
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# as noted R does Type I SS, Type III SS for time is 317 (SAS etc); interaction is prime concern,
   that (407) matches SAS PROC GLM
```

data in stat222 web

truncated

better version

fun plots

unequal group sizes makes non-orthog design

bigger than Type III

matches SAS, publication

```
#so let's try an lmer model: level 1 outcome ~ time; level 2 slope (diff score) depends on method
> bklmera = lmer(outcome ~ I(time - 1) + I(time-1):as.factor(method) + (time|subject), data = bk)
> summary(bklmera)
Linear mixed model fit by REML
Formula: outcome ~ I(time - 1) + I(time - 1):as.factor(method) + (time |      subject)
   Data: bk
   AIC   BIC logLik deviance REMLdev
 305.7 317.9 -145.9    301.1   291.7
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 subject  (Intercept) 35.000   5.9161
          time        21.455   4.6320   0.220
 Residual             25.125   5.0124
Number of obs: 42, groups: subject, 21
Fixed effects:
                                Estimate Std. Error t value
(Intercept)                       44.619      2.112  21.130
I(time - 1)                       -5.672      1.902  -2.981
I(time - 1):as.factor(method)1     6.378      1.902   3.354

Correlation of Fixed Effects:
            (Intr) I(t-1)
I(time - 1) 0.028
I(-1):s.()1 0.000  0.238
# so interaction matches F-statistic from repeated measures anova
> 3.354^2
[1] 11.24932
# AND lmer gets the occasions (time) term "correct" in the test statistic
> 2.981^2
[1] 8.886361
# this matches F-statistic in publication (and SAS) repeated measures output of 8.86 for pre/post (time)
# whereas the aov above has F-statistic 15.1
# SS not comparable with anova because here were are modeling level 1 params, not outcome
So before looking at other small details, let us declare an lmer victory over non-othogonal designs
```

lmer rules

Model

Level 1 within subject

$t1 = time - 1$ better this way

$Y = \alpha_0 + \alpha_1 t1 + \varepsilon$ $\alpha_0 = pre$

$\alpha_1 = post - pre$

Level 2 $\alpha_0 = \gamma_{00} + u_0$

$\alpha_1 = \gamma_{10} + \gamma_{11} method + u_1$

Combined

$Y = \gamma_{00} + \gamma_{10} t1 + \gamma_{11} t1{:}method + (u_0 + u_1 + \varepsilon)$

Type III SAS

extended version posted bklmer

```
BK lmer analysis 2016 style ######################
> bk  = read.table(file="http://statweb.stanford.edu/~rag/stat222/brogkutlong2.dat", header = T)
> bk$t1 = bk$time - 1    # make 0,1 vars
> bk$G = bk$method - 1
> # 'fix' lmer 2-wave errors  by 'no 2-wave worries'
> bklmera = lmer(outcome ~ t1 + t1:as.factor(method) + (t1|subject), data = bk,
                       control = lmerControl(check.nobs.vs.nRE = "warning"))
Warning messages:
1: number of observations (=42) <= number of random effects (=42) for term (t1 | subject);
  the random-effects parameters and the residual variance (or scale parameter) are probably unidentifia
2: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv,  :
  Model is nearly unidentifiable: large eigenvalue ratio - Rescale variables?

> summary(bklmera)
Linear mixed model fit by REML ['lmerMod']
Formula: outcome ~ t1 + t1:as.factor(method) + (t1 | subject)
   Data: bk
Control: lmerControl(check.nobs.vs.nRE = "warning")
REML criterion at convergence: 290.3
Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.9936 -0.4127 -0.1596  0.4288  1.7313
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 subject  (Intercept) 66.45    8.152
          t1          17.31    4.161    0.87
 Residual             27.20    5.215
Number of obs: 42, groups:  subject, 21

Fixed effects:
                     Estimate Std. Error t value
(Intercept)           44.6190     2.1117  21.129
t1                     0.7057     2.9931   0.236
t1:as.factor(method)2 -12.7553     3.8035  -3.354
Correlation of Fixed Effects:
            (Intr) t1
t1           0.018
t1:s.fct()2  0.000 -0.787

> anova(bklmera) # put fixed effects in SS metric
Analysis of Variance Table
                    Df Sum Sq Mean Sq F value
t1                   1 411.79  411.79  15.142
t1:as.factor(method)  1 305.84  305.84  11.246  # test statistic for method effect on improvement

> confint(bklmera)
> confint(bklmera, method = "boot", nsim = 1000, boot.type = "perc")
> # bootstrap gives reasonable bounds for random effects even

> # lmer 'a' does not include pretest diffs because of random assignment, can look at that
> bklmerb = lmer(outcome ~ t1 + t1*as.factor(method) + (t1|subject), data = bk,
                         control = lmerControl(check.nobs.vs.nRE = "warning"))
Warning messages:

> anova(bklmera, bklmerb) # compare nested models
refitting model(s) with ML (instead of REML)
Data: bk
Models:
bklmera: outcome ~ t1 + t1:as.factor(method) + (t1 | subject)
bklmerb: outcome ~ t1 + t1 * as.factor(method) + (t1 | subject)
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
bklmera  7 315.10 327.26 -150.55   301.10
bklmerb  8 316.65 330.55 -150.32   300.65 0.4516      1     0.5016
> # extended model does not help
>  ##############################################################################################
> # can also use older lme
```