Stat 209
week 9, [redacted]

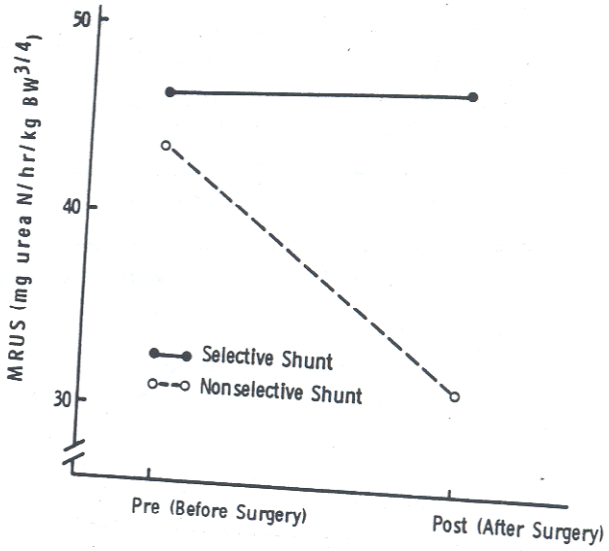
Pre-post, 2 groups Repeated Measures Anova

Brogan-Kutner example

© The American Statistician, November 1980, Vol. 34, No. 4

1. Pre and Post Maximal Rate of Urea Synthesis Level (mg urea N/hr/kg $BW^{3/4}$) and Sample Cell Means, by Group

Group	Subject	Pre	Post
Selective Shunt (new operation)	1	51	48
	2	35	55
	3	66	60
	4	40	35
	5	39	36
	6	46	43
	7	52	46
	8	42	54
Mean		$\hat{\mu}_{11} = 46.375$	$\hat{\mu}_{12} = 47.125$
Nonselective Shunt (standard operation)	9	34	16
	10	40	36
	11	34	16
	12	36	18
	13	38	32
	14	32	14
	15	44	20
	16	50	43
	17	60	45
	18	63	67
	19	50	36
	20	42	34
	21	43	32
Mean		$\hat{\mu}_{21} = 43.538$	$\hat{\mu}_{22} = 31.462$



analysis on back

Growth Curves (Group) T=4

Bock, DR MSMBR text

EXAMPLE 7.1-1 (*Mixed-model analysis of vocabulary growth*) Data for this example are drawn from test results on file in the Records Office of the Laboratory School of the University of Chicago. They consist of scores, obtained from a cohort of pupils at the eighth through eleventh grade level, on alternative forms of the vocabulary section of the Cooperative Reading Tests [Davis, 1950]. Since these data cover an age range in which physical growth is beginning to decelerate, it is of interest to inquire whether a similar deceleration can be observed in the acquisition of new vocabulary.

Table 7.2-5 MIXED-MODEL ANALYSIS OF VARIANCE OF SEX EFFECTS IN THE VOCABULARY-SCALED SCORES

Source	df	ss	F	p
Constant	1	ssm = 1,644.90		
Sex	1	ssb = .85	.06	> .5
Occasions	3	ssc = 194.18		
Sex × occasions	3	ssbc = 2.79	1.12	> .1
Subjects within groups	62	ssa = 873.00		
Occasions × subjects within groups	186	sse = 152.17		
Total	256	sst = 2,867.90		

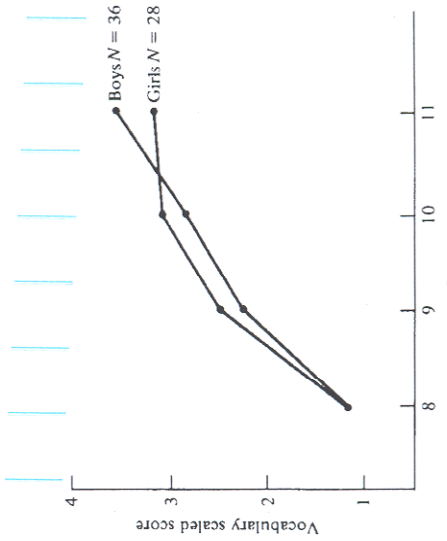


FIGURE 7.2-1
Average vocabulary scores of boys and girls in a cohort from the University of Chicago Laboratory School (longitudinal data).

Repeated Measures Brogan-Kutner ix p. 2

D. Rogosa

model

$$X_{ijk} = \mu + \alpha_i + \Pi_{k(i)} + \beta_j + \alpha\beta_{ij} + \beta\Pi_{jk(i)} + \epsilon_{m(ijk)} \quad (3.1)$$

$j = 1, 2$ (pretest = 1, posttest = 2),

$i = 1, 2$ (group 1 = 1, group 2 = 2),

$k = 1, 2, \ldots, n_i$, $m = 1$,

where X_{ijk} is the observed value of subject k within group i at time j,

μ is the overall mean,
α_i is the effect of group i,
$\Pi_{k(i)}$ is the effect of subject k nested within group i,
β_j is the effect of the repeated-measures variable j (*i.e.*, pretest and posttest),
$\alpha\beta_{ij}$ is the interaction of group i with level j of the repeated measures factor,

$\beta\Pi_{jk(i)}$ is the interaction of subject k within group i with level j of the repeated-measures factor.

2. Repeated-Measures Analysis of Variance for Maximal Rate of Urea Synthesis Level

Source of Variation	df	Sum of Squares	Mean Squares	F Ratio
Between Subjects	20 (n − 1)			
Groups	1	847.48	847.48 (MS_G)	3.63 (MS_G/MS_E)
Subjects Within Groups	19 (n − 2)	4440.00	233.68 (MS_E)	
Within Groups	21 (n)			
Pre/Post	1	317.69	317.69 (MS_P)	8.86 (MS_P/MS_{PE})
Groups x Pre/Post	1	407.41	407.4 (MS_{GP})	11.36 (MS_{GP}/MS_{PE})
(Pre/Post) x Subjects Within Groups	19 (n − 2)	681.21	35.85 (MS_{PE})	

Did the groups change differentially?

SAS or minitab does it

(R has problem w/ imbalance anova /error TBD

```
proc glm data=brogk;
class grp;
model m1--m2 = grp /nouni;
repeated Time 2 (1 2) / summary printe;
run;
```

OUTPUT (selected)

```
         The SAS System                 16:13 Tuesday, May 16, 2000  35

                                The GLM Procedure
                        Repeated Measures Analysis of Variance
                     Tests of Hypotheses for Between Subjects Effects

Source        DF     Type III SS      Mean Square    F Value     Pr > F

grp            1      847.476190      847.476190       3.63     0.0721
Error         19     4440.000000      233.684211

                                The GLM Procedure
                        Repeated Measures Analysis of Variance
                   Univariate Tests of Hypotheses for Within Subject Effects

Source          DF     Type III SS      Mean Square    F Value     Pr > F

Time             1     317.6932234      317.6932234       8.86     0.0078
Time*grp         1     407.4075092      407.4075092      11.36     0.0032
Error(Time)     19     681.2115385       35.8532389
```

Repeated Measures Anova / more BK

Week 9 Stat 209

```
# Brogan-Kutner Data see http://www-stat.stanford.edu/~rag/ed351longit/brogkut.dat

# Cell means
> tapply(urea, list(method, prepost), mean)
         1        2
1 46.37500 47.12500
2 43.53846 31.46154

# Recreate repeated measures anova (nesting)
# within-groups anova to obtain the 2 error terms

#within group 1 subjXtime
> bkrepaovW1 = aov(urea[method == "1"] ~ as.factor(prepost[method == "1"])*as.factor(subj[method == "1"]))
> summary(bkrepaovW1)
                                                                  Df Sum Sq Mean Sq
as.factor(prepost[method == "1"])                                  1   2.25    2.25
as.factor(subj[method == "1"])                                     7 915.00  130.71
                          piece of subjects within groups  Between subjects error term
as.factor(prepost[method == "1"]):as.factor(subj[method == "1"])   7 331.75   47.39
   piece of subjectsxrepeated measure  within group interaction   Within subjects error term

#within group 2 subjXtime
> bkrepaovW2 = aov(urea[method == "2"] ~ as.factor(prepost[method == "2"])*as.factor(subj[method == "2"]))
> summary(bkrepaovW2)
                                                                  Df Sum Sq Mean Sq
as.factor(prepost[method == "2"])                                  1  948.0   948.0
as.factor(subj[method == "2"])                                    12 3525.0   293.7
                          piece of subjects within groups  Between subjects error term
as.factor(prepost[method == "2"]):as.factor(subj[method == "2"]) 12  349.5    29.1
   piece of subjectsxrepeated measure  within group interaction   Within subjects error term

#               915 + 3525 = 4440 (and 7 + 12 = 19df)   Between subjects SS error term
#               331.7 + 349.5 = 681.2 (and 7 + 12 = 19df) Within subjects SS error term

# ignore within-subjects, get
> bkrepaovBase = aov(urea ~ as.factor(prepost)*as.factor(method))
> summary(bkrepaovBase)
                                 Df Sum Sq Mean Sq F value  Pr(>F)
as.factor(prepost)                1  542.9   542.9  4.0282 0.05190 .  #repeated measure (Within subj part)
as.factor(method)                 1  847.5   847.5  6.2884 0.01654 *  #Group  (Between subjects part)
as.factor(prepost):as.factor(method) 1 407.4  407.4  3.0230 0.09019 .  #GroupxRepeated measure Interaction
Residuals                        38 5121.2   134.8                        (Within subjects part)
```

cf main BK handout

data observations as rows (42)

Do repeated measures anova by crossed designs on subjects.

```
# Brogan-Kutner Section 5 Equivalences
   # Groups, pooling over occasion
> sumtime = pre + post
> t.test(sumtime ~ as.factor(method), var.equal = TRUE)
    Two Sample t-test data:  sumtime by as.factor(method)
t = 1.9044, df = 19, p-value = 0.07212
95 percent confidence interval: -1.832786 38.832786
mean in group 1 mean in group 2
          93.5            75.0
> 1.904^2  [1] 3.625216  # matches F-stat for Groups (bet subj)

> imp = post - pre
> t.test(imp ~ as.factor(method), var.equal = TRUE)
  Two Sample t-test data:  imp by as.factor(method)
t = 3.3709, df = 19, p-value = 0.003209
95 percent confidence interval:  4.862645 20.791201
mean in group 1 mean in group 2
        0.75000       -12.07692
> 3.3709^2  [1] 11.36297  # matches F-stat for Groups X prepost

> t.test(imp)
 One Sample t-test data:  imp
t = -3.1581, df = 20, p-value = 0.004947
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval: -11.939835  -2.441117
mean of x -7.190476
> 3.1581^2 [1] 9.973596  # equiv to prepost, no differential change
                          BK p.232
```

main event differential change by t-test

subj as rows format

```
> bksubj
   pre post method
1   51   48      1
2   35   55      1
3   66   60      1
4   40   35      1
5   39   36      1
6   46   43      1
7   52   46      1
8   42   54      1
9   34   16      2
10  40   36      2
11  34   16      2
12  36   18      2
13  38   32      2
14  32   14      2
15  44   20      2
16  50   43      2
17  60   45      2
18  63   67      2
19  50   36      2
20  42   34      2
21  43   32      2
```

```
> bkrepaov1 = aov(urea ~ as.factor(prepost)*as.factor(method)+ Error(as.factor(subj)))
> summary(bkrepaov1)
Error: as.factor(subj)
                  Df Sum Sq Mean Sq F value  Pr(>F)
as.factor(method)  1  847.5   847.5  3.6266 0.07212 .
Residuals         19 4440.0   233.7
---
Error: Within
                                    Df Sum Sq Mean Sq F value    Pr(>F)
as.factor(prepost)                   1 542.88  542.88  15.142 0.0009823 ***
as.factor(prepost):as.factor(method)  1 407.41  407.41  11.363 0.0032085 **
Residuals                            19 681.21   35.85
```

subj x prepost x method

R does the repeated meas design See Baron+Li

Sequential SS issue w/ prepost SS

```
# Brogan-Kutner Data see http://www-stat.stanford.edu/~rag/ed351longit/brogkut.dat

# Cell means
> tapply(urea, list(method, prepost), mean)
         1        2
1 46.37500 47.12500
2 43.53846 31.46154
```

```
# Recreate repeated measures anova (nesting)
# within-groups anova to obtain the 2 error terms

#within group 1 subjXtime
> bkrepaovW1 = aov(urea[method == "1"] ~ as.factor(prepost[method == "1"])*as.factor(subj[method == "1"]))
> summary(bkrepaovW1)
                                                          Df Sum Sq Mean Sq
as.factor(prepost[method == "1"])                          1   2.25    2.25
as.factor(subj[method == "1"])                             7 915.00  130.71
                        piece of subjects within groups  Between subjects error term
as.factor(prepost[method == "1"]):as.factor(subj[method == "1"])  7 331.75   47.39
   piece of subjectsxrepeated measure  within group interaction   Within subjects error term

#within group 2 subjXtime
> bkrepaovW2 = aov(urea[method == "2"] ~ as.factor(prepost[method == "2"])*as.factor(subj[method == "2"]))
> summary(bkrepaovW2)
                                                          Df Sum Sq Mean Sq
as.factor(prepost[method == "2"])                          1  948.0   948.0
as.factor(subj[method == "2"])                            12 3525.0   293.7
                       piece of subjects within groups  Between subjects error term
as.factor(prepost[method == "2"]):as.factor(subj[method == "2"]) 12  349.5    29.1
   piece of subjectsxrepeated measure  within group interaction   Within subjects error term

#              915 + 3525 = 4440 (and 7 + 12 = 19df)   Between subjects SS error term
#              331.7 + 349.5 = 681.2 (and 7 + 12 = 19df) Within subjects SS error term

# ignore within-subjects, get
> bkrepaovBase = aov(urea ~ as.factor(prepost)*as.factor(method))
> summary(bkrepaovBase)
                                  Df Sum Sq Mean Sq F value  Pr(>F)
as.factor(prepost)                 1  542.9   542.9  4.0282 0.05190 .  #repeated measure (Within subj part)
as.factor(method)                  1  847.5   847.5  6.2884 0.01654 *  #Group  (Between subjects part)
as.factor(prepost):as.factor(method)  1  407.4   407.4  3.0230 0.09019 .  #GroupxRepeated measure Interaction
Residuals                         38 5121.2   134.8                          (Within subjects part)
```

```
# Brogan-Kutner Section 5 Equivalences
   # Groups, pooling over occasion
> sumtime = pre + post
> t.test(sumtime ~ as.factor(method), var.equal = TRUE)
    Two Sample t-test data:  sumtime by as.factor(method)
t = 1.9044, df = 19, p-value = 0.07212
95 percent confidence interval: -1.832786 38.832786
mean in group 1 mean in group 2
           93.5            75.0
> 1.904^2  [1] 3.625216  # matches F-stat for Groups (bet subj)

> imp = post - pre
> t.test(imp ~ as.factor(method), var.equal = TRUE)
  Two Sample t-test data:  imp by as.factor(method)
t = 3.3709, df = 19, p-value = 0.003209
95 percent confidence interval:  4.862645 20.791201
mean in group 1 mean in group 2
        0.75000       -12.07692
> 3.3709^2  [1] 11.36297  # matches F-stat for Groups X prepost

> t.test(imp)
 One Sample t-test data:  imp
t = -3.1581, df = 20, p-value = 0.004947
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval: -11.939835  -2.441117
mean of x -7.190476
> 3.1581^2 [1] 9.973596  # equiv to prepost, no differential change
                            BK p.232
```

```
> bksubj
   pre post method
1   51   48      1
2   35   55      1
3   66   60      1
4   40   35      1
5   39   36      1
6   46   43      1
7   52   46      1
8   42   54      1
9   34   16      2
10  40   36      2
11  34   16      2
12  36   18      2
13  38   32      2
14  32   14      2
15  44   20      2
16  50   43      2
17  60   45      2
18  63   67      2
19  50   36      2
20  42   34      2
21  43   32      2
```

```
> bkrepaov1 = aov(urea ~ as.factor(prepost)*as.factor(method)+ Error(as.factor(subj)))
> summary(bkrepaov1)
Error: as.factor(subj)
                  Df Sum Sq Mean Sq F value  Pr(>F)
as.factor(method)  1  847.5   847.5  3.6266 0.07212 .
Residuals         19 4440.0   233.7
---
Error: Within
                                  Df Sum Sq Mean Sq F value    Pr(>F)
as.factor(prepost)                 1 542.88  542.88  15.142 0.0009823 ***Type III SS(prepost) = 317
as.factor(prepost):as.factor(method)  1 407.41  407.41  11.363 0.0032085 **
Residuals                         19 681.21   35.85
```

[redacted] BK lmer

```
    Update of BK repeated measures analysis
> 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
```

truncated

better version

fun plots

unequal group sizes makes non-orthog design

bigger than Type I

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
```

Type II SAS

lmer rules

Model

Level 1 within subject

t1 = time - 1 better this way

$y = \alpha_0 + \alpha_1 t1 + \varepsilon$

α_0 = pre

α_1 = post - pre

Level 2

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

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

Combined

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

extended version posted bklmer