

Review Problem Set 3 Solutions

Exercise 13.3

```
Ex13.3 <- read.table("http://www.stat.umn.edu/~gary/book/fcdae.data/ex13.3",header=TRUE)
head(Ex13.3)
```

```
##   row col trt  y
## 1   1   1   4 44
## 2   1   2   2 26
## 3   1   3   3 67
## 4   1   4   1 77
## 5   1   5   2 51
## 6   1   6   4 62
```

```
Ex13.3 <- within(Ex13.3,
  {
    square <- NA
    square[col < 5] <- 1
    square[col > 4] <- 2
  })
```

```
Ex13.3$row <- as.factor(Ex13.3$row)
Ex13.3$col <- as.factor(Ex13.3$col)
Ex13.3$trt <- as.factor(Ex13.3$trt)
```

```
# -----
# Using just the first four columns
```

```
Ex13.3a <- subset(Ex13.3, square == 1)
```

```
Ex13.3a.lm1 <- lm(y ~ row + col + trt, data=Ex13.3a)
```

```
anova(Ex13.3a.lm1)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: y
```

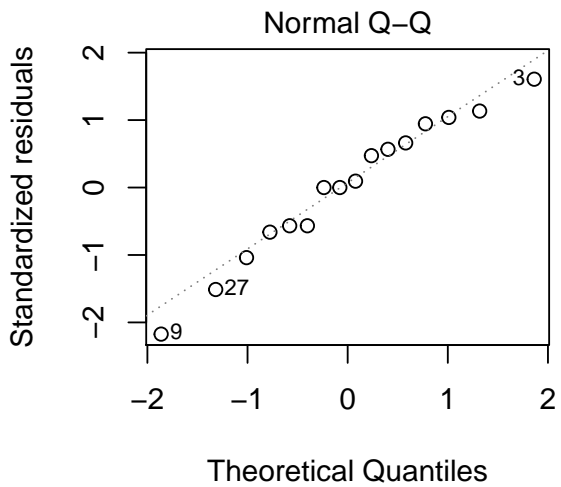
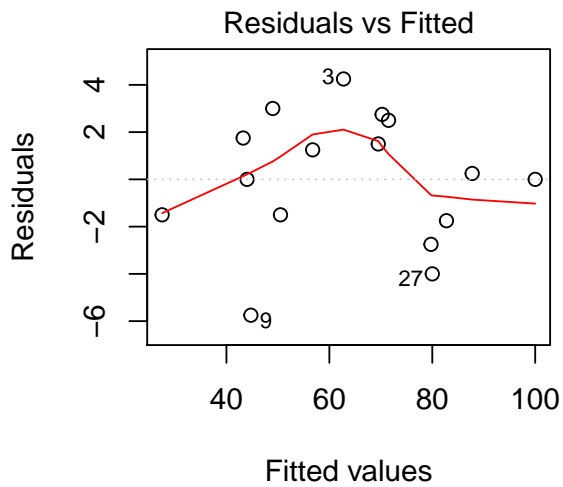
```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## row         3  1321.5   440.50  23.598 0.001012 **
## col         3  4198.5  1399.50   74.973 3.8e-05 ***
## trt         3   315.0   105.00    5.625 0.035362 *
## Residuals   6   112.0    18.67
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# check model assumptions

par(mfrow=c(1,2))
plot(Ex13.3a.lm1, which = c(1,2))
```



```
par(mfrow=c(1,1))

# Tukey tests

library(multcomp)
```

```
## Warning: package 'multcomp' was built under R version 3.6.3

## Loading required package: mvtnorm

## Warning: package 'mvtnorm' was built under R version 3.6.3

## Loading required package: survival

## Loading required package: TH.data

## Warning: package 'TH.data' was built under R version 3.6.3

## Loading required package: MASS

##

## Attaching package: 'TH.data'
```

```

## The following object is masked from 'package:MASS':
##
##      geyser

# Tukey's multiple comparison method

trta.Tukey <- glht(Ex13.3a.lm1,linfct=mcp(trt="Tukey"))

summary(trta.Tukey)

##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lm(formula = y ~ row + col + trt, data = Ex13.3a)
##
## Linear Hypotheses:
##           Estimate Std. Error t value Pr(>|t|)
## 2 - 1 == 0  -12.000     3.055  -3.928  0.0296 *
## 3 - 1 == 0   -6.000     3.055  -1.964  0.2972
## 4 - 1 == 0   -3.000     3.055  -0.982  0.7647
## 3 - 2 == 0    6.000     3.055   1.964  0.2971
## 4 - 2 == 0    9.000     3.055   2.946  0.0917 .
## 4 - 3 == 0    3.000     3.055   0.982  0.7647
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)

#plot(trta.Tukey,sub="Tukeys method")

```

From the ANOVA table we see that the row and column factors account for a lot of variation, we could calculate efficiencies to quantify their effect more precisely. The test for treatment has an $F = 5.63$ on 3 and 6 df and is significant. The model assumptions look good, with no obvious HOV problems from the residual by predicted plot and a reasonable normal plot. The Tukey multiple comparisons show only a difference between treatments 1 and 2.

```

# -----
# Using all columns as a replicated Latin square

# perform Latin Square analysis

Ex13.3.lm1 <- lm(y ~ row +col +trt, data=Ex13.3)

anova(Ex13.3.lm1)

```

```

## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## row         3  1913.6   637.88  39.254 4.187e-08 ***

```

```
## col      7 4869.4  695.62  42.808 5.997e-10 ***
## trt      3  635.4  211.79  13.033 9.167e-05 ***
## Residuals 18 292.5  16.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

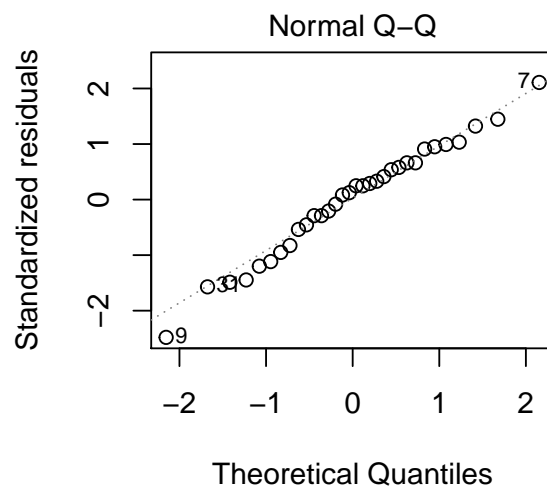
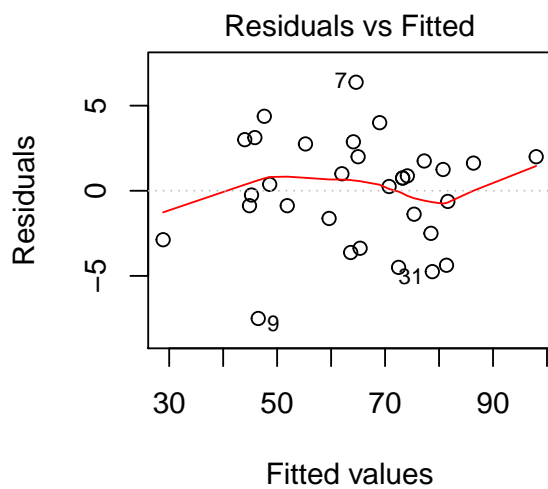
```
# alternatively
```

```
Ex13.3.lm2 <- lm(y ~ square +square:col +row +trt, data=Ex13.3)
anova(Ex13.3.lm2)
```

```
## Analysis of Variance Table
##
## Response: y
##          Df Sum Sq Mean Sq F value    Pr(>F)
## square    1   36.1   36.13  2.2231  0.1533
## row       3 1913.6  637.88 39.2538 4.187e-08 ***
## trt       3  635.4  211.79 13.0333 9.167e-05 ***
## square:col 6 4833.3  805.54 49.5718 3.176e-10 ***
## Residuals 18  292.5   16.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# check model assumptions
```

```
par(mfrow=c(1,2))
plot(Ex13.3.lm1, which = c(1,2))
```



```

par(mfrow=c(1,1))

# Tukey tests

#library(multcomp)

# Tukey's multiple comparison method

trt.Tukey <- glht(Ex13.3.lm1,linfct=mcp(trt="Tukey"))

summary(trt.Tukey)

```

```

##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lm(formula = y ~ row + col + trt, data = Ex13.3)
##
## Linear Hypotheses:
##      Estimate Std. Error t value Pr(>|t|)
## 2 - 1 == 0  -12.250     2.016  -6.078 < 0.001 ***
## 3 - 1 == 0   -6.250     2.016  -3.101  0.02858 *
## 4 - 1 == 0   -3.750     2.016  -1.861  0.27926
## 3 - 2 == 0    6.000     2.016   2.977  0.03695 *
## 4 - 2 == 0    8.500     2.016   4.217  0.00278 **
## 4 - 3 == 0    2.500     2.016   1.240  0.61044
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)

```

```

#plot(trt.Tukey,sub="Tukeys method")

```

The results from the replicated Latin square are similar, but now there are more pairs declared significant by Tukey's method due to more replication.

Exercise 14.4

```

# first read the data directly from Dr. Oehlerts website

Ex14.4 <- read.table("http://www.stat.umn.edu/~gary/book/fcdae.data/ex14.4",header=TRUE)

head(Ex14.4)

```

```

##   computer treatment response
## 1         1         1       35
## 2         2         1       41
## 3         3         2       40

```

```
## 4      4      1      32
## 5      5      1      40
## 6      1      2      42
```

```
Ex14.4$computer <- as.factor(Ex14.4$computer)
Ex14.4$treatment <- as.factor(Ex14.4$treatment)

# perform BIBD analysis

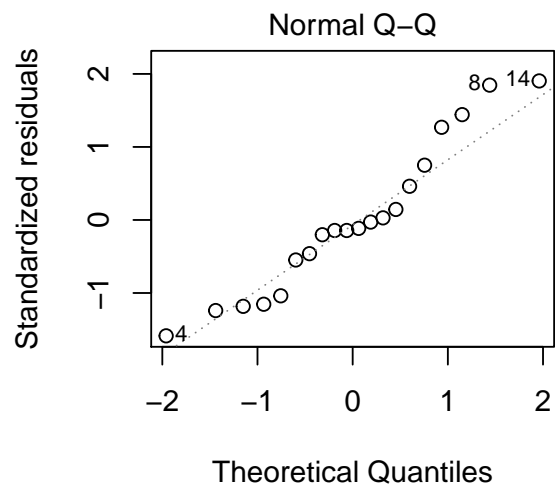
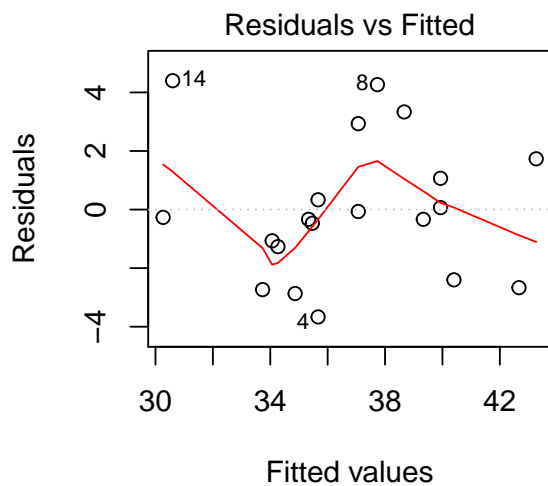
Ex14.4.lm1 <- lm(response ~ computer + treatment, data=Ex14.4)

anova(Ex14.4.lm1)
```

```
## Analysis of Variance Table
##
## Response: response
##          Df Sum Sq Mean Sq F value Pr(>F)
## computer  4   95.2  23.800   2.4513 0.10791
## treatment 4  139.2  34.800   3.5843 0.04179 *
## Residuals 11  106.8   9.709
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# check model assumptions

par(mfrow=c(1,2))
plot(Ex14.4.lm1, which = c(1,2))
```



```
par(mfrow=c(1,1))

# Tukey tests
```

```

library(multcomp)

# Tukey's multiple comparison method

trt.Tukey14.4 <- glht(Ex14.4.lm1, linfct=mcp(treatment="Tukey"))

summary(trt.Tukey14.4)

##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lm(formula = response ~ computer + treatment, data = Ex14.4)
##
## Linear Hypotheses:
## Estimate Std. Error t value Pr(>|t|)
## 2 - 1 == 0 3.333e+00 2.276e+00 1.465 0.6028
## 3 - 1 == 0 -1.600e+00 2.276e+00 -0.703 0.9515
## 4 - 1 == 0 -5.067e+00 2.276e+00 -2.227 0.2395
## 5 - 1 == 0 3.953e-16 2.276e+00 0.000 1.0000
## 3 - 2 == 0 -4.933e+00 2.276e+00 -2.168 0.2600
## 4 - 2 == 0 -8.400e+00 2.276e+00 -3.691 0.0237 *
## 5 - 2 == 0 -3.333e+00 2.276e+00 -1.465 0.6029
## 4 - 3 == 0 -3.467e+00 2.276e+00 -1.523 0.5694
## 5 - 3 == 0 1.600e+00 2.276e+00 0.703 0.9515
## 5 - 4 == 0 5.067e+00 2.276e+00 2.227 0.2395
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)

```

```
#plot(trt.Tukey14.4, sub="Tukeys method")
```

This is a BIBD with $g=5$ treatment levels, $b=5$ blocks, block size of $k=4$, and each treatment has $r=4$ replicates. We can calculate the number of times each pair of treatments occur together in a block by:

$$\lambda = r(k - 1)/(g - 1) = 4 * 3/4 = 3.$$

The ANOVA table shows that the blocks account for nearly a third of the SS while only using 4 df. The test for treatment has $F=3.58$ on 4 and 11 df and is significant with $P=.042$. The residuals have a bit of a curvy shape to them, but no obvious HOV problems from the residual by predicted plot nor obvious normality issues from the QQ plot (we could try a Box-Cox analysis if we wanted to pursue the assumptions further). The Tukey results show only a difference in access times between treatments 2 and 4.

Problem 16.8

```

# first read the data directly from Dr. Oehlerts website

Pr16.8 <- read.table("http://www.stat.umn.edu/~gary/book/fcdae.data/pr16.8", header=TRUE)

```

```
head(Pr16.8)
```

```
##   sign timing interchange speed
## 1    1      1           1     13
## 2    1      2           1     25
## 3    1      3           1     26
## 4    1      1           2     24
## 5    1      2           2     35
## 6    1      3           2     37
```

```
Pr16.8$time2 <- 3*(Pr16.8$timing -1)    # convert 1, 2, 3 to 0, 3, 6
```

```
Pr16.8$sign <- as.factor(Pr16.8$sign)
```

```
Pr16.8$timing <- as.factor(Pr16.8$timing)
```

```
Pr16.8$interchange <- as.factor(Pr16.8$interchange)
```

```
# view sums of squares with model assuming fixed effects
```

```
Pr16.8.lm1 <- lm(speed ~ sign +sign:interchange +timing +sign:timing, data = Pr16.8)
```

```
anova(Pr16.8.lm1)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: speed
```

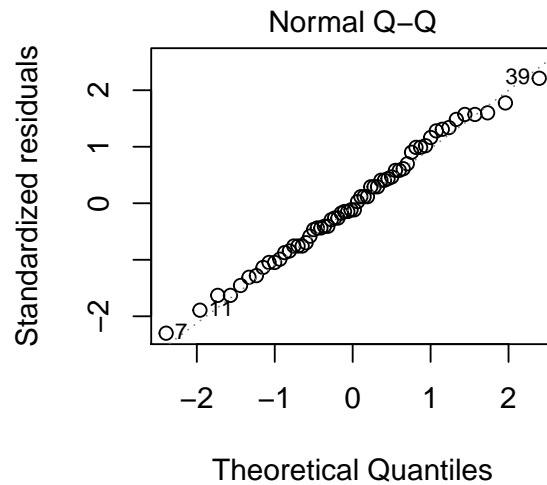
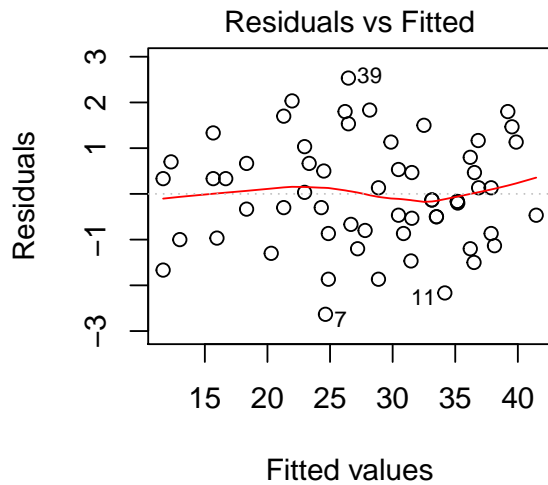
```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## sign           1  173.40   173.40  79.2854 1.255e-10 ***
## timing          2 2576.23 1288.12 588.9780 < 2.2e-16 ***
## sign:interchange 18 1042.87   57.94  26.4911 1.544e-15 ***
## sign:timing       2    3.70    1.85   0.8459   0.4375
## Residuals      36   78.73    2.19
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
par(mfrow=c(1,2))
```

```
plot(Pr16.8.lm1, which = c(1,2))
```

```
par(mfrow=c(1,1))
```

```
# The afex library calls lmer and other functions to give results fairly  
# similar to those in Proc MIXED
```

```
library(afex)
```

```
## Loading required package: lme4
```

```
## Warning: package 'lme4' was built under R version 3.6.3
```

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

```
## Registered S3 methods overwritten by 'car':
```

```
##   method                      from  
##   influence.merMod             lme4  
##   cooks.distance.influence.merMod lme4  
##   dfbeta.influence.merMod      lme4  
##   dfbetas.influence.merMod     lme4
```

```
## *****
```

```
## Welcome to afex. For support visit: http://afex.singmann.science/
```

```
## - Functions for ANOVAs: aov_car(), aov_ez(), and aov_4()
```

```
## - Methods for calculating p-values with mixed(): 'S', 'KR', 'LRT', and 'PB'
```

```
## - 'afex_aov' and 'mixed' objects can be passed to emmeans() for follow-up tests
```

```
## - NEWS: emmeans() for ANOVA models now uses model = 'multivariate' as default.
```

```
## - Get and set global package options with: afex_options()
```

```
## - Set orthogonal sum-to-zero contrasts globally: set_sum_contrasts()
```

```
## - For example analyses see: browseVignettes("afex")
```

```
## *****
```

```

##
## Attaching package: 'afex'

## The following object is masked from 'package:lme4':
##
##   lmer

library(emmeans)

Pr16.8.afex1 <- mixed(speed ~ sign +(1|sign:interchange) +timing +sign:timing, data = Pr16.8)

## Contrasts set to contr.sum for the following variables: sign, interchange, timing

## Fitting one lmer() model. [DONE]
## Calculating p-values. [DONE]

#summary(Pr16.8.afex1)

nice(Pr16.8.afex1)

## Mixed Model Anova Table (Type 3 tests, S-method)
##
## Model: speed ~ sign + (1 | sign:interchange) + timing + sign:timing
## Data: Pr16.8
##      Effect      df      F p.value
## 1      sign 1, 18.00      2.99  .101
## 2      timing 2, 36.00 588.98 ***  <.001
## 3 sign:timing 2, 36.00      0.85  .438
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1

emmeans(Pr16.8.afex1, ~sign)

## NOTE: Results may be misleading due to involvement in interactions

##   sign emmean  SE df lower.CL upper.CL
## 1      29.8 1.39 18      26.9      32.8
## 2      26.4 1.39 18      23.5      29.4
##
## Results are averaged over the levels of: timing
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95

emmeans(Pr16.8.afex1, ~timing)

## NOTE: Results may be misleading due to involvement in interactions

##   timing emmean  SE  df lower.CL upper.CL
## 1      18.9 1.02 20.8      16.8      21.1
## 2      31.6 1.02 20.8      29.5      33.8

```

```

## 3      33.8 1.02 20.8      31.7      35.9
##
## Results are averaged over the levels of: sign
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95

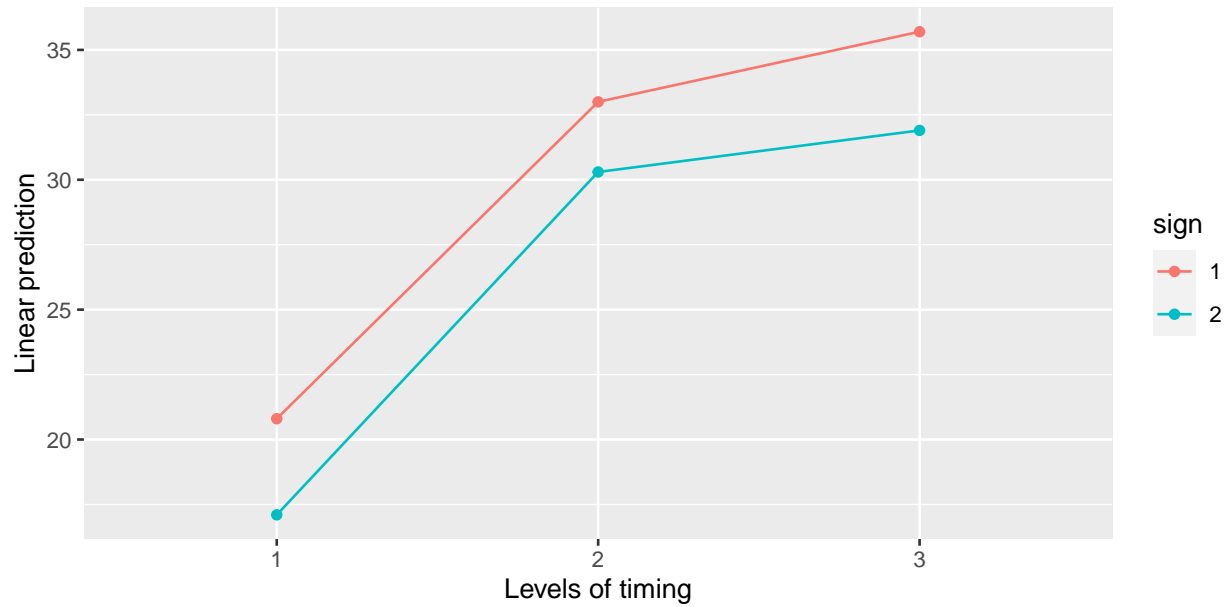
# multiple comparisons with Tukeys method
emmeans(Pr16.8.afex1, list(pairwise ~ timing), adjust = "tukey")

## NOTE: Results may be misleading due to involvement in interactions

## $'emmeans of timing'
## timing emmean SE df lower.CL upper.CL
## 1      18.9 1.02 20.8      16.8      21.1
## 2      31.6 1.02 20.8      29.5      33.8
## 3      33.8 1.02 20.8      31.7      35.9
##
## Results are averaged over the levels of: sign
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
##
## $'pairwise differences of timing'
## 1 estimate SE df t.ratio p.value
## 1 - 2 -12.70 0.468 36 -27.157 <.0001
## 1 - 3 -14.85 0.468 36 -31.754 <.0001
## 2 - 3 -2.15 0.468 36 -4.597 0.0001
##
## Results are averaged over the levels of: sign
## Degrees-of-freedom method: kenward-roger
## P value adjustment: tukey method for comparing a family of 3 estimates

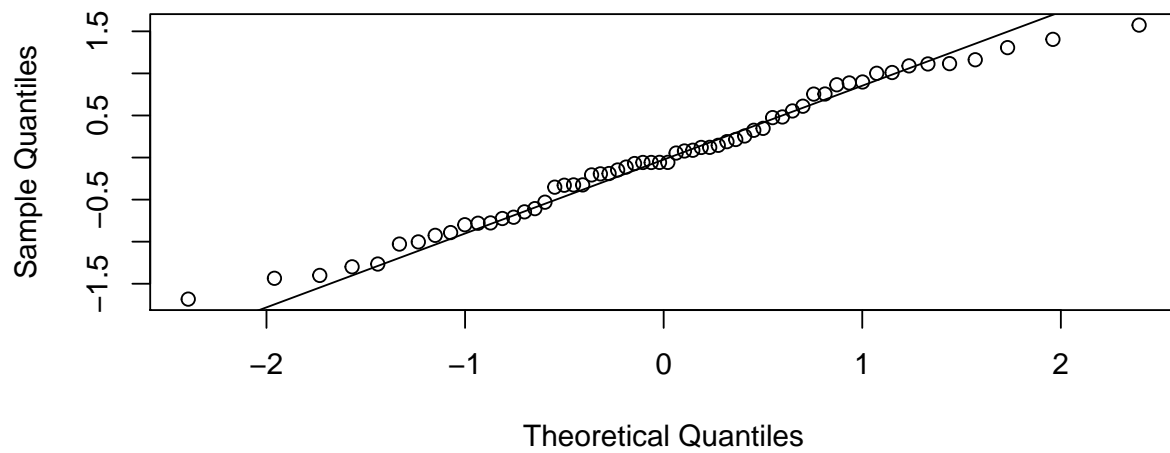
# examine potential interactions
lsmip(Pr16.8.afex1, sign ~ timing)

```

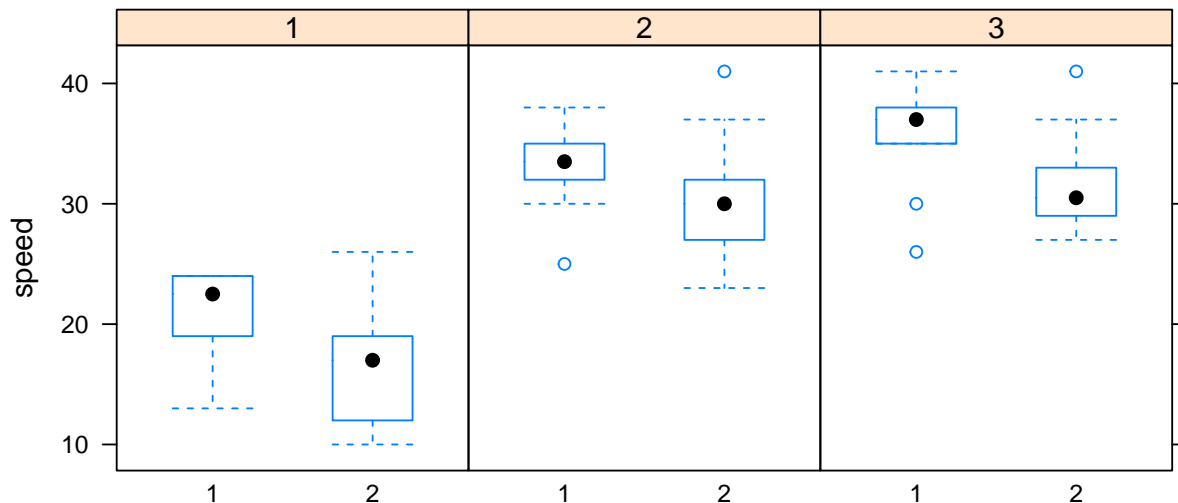


```
# for a normal plot, use the qqnorm and qqline functions:
qqnorm(summary(Pr16.8.afex1)$residuals)
qqline(summary(Pr16.8.afex1)$residuals)
```

Normal Q-Q Plot



```
library(lattice)
bwplot(speed ~ sign | timing, data = Pr16.8)
```



```
library(performance)
check_homogeneity(Pr16.8.afex1)
```

```
## OK: There is not clear evidence for different variances across groups (Bartlett Test, p = 0.803).
```

From the separate randomization of signs to interchanges and timing within interchanges, we see that this is a split-plot design. The whole plot treatment has a completely randomized structure, as we randomize signs to interchanges without any blocking involved. We first look at a fixed-effects ANOVA and its residuals to get a sense of things. The split-plot tests here are accurate but not the whole-plot test. We see that timing is very significant but the interaction is not. The correct F test for sign would use the interchange(sign) nested term as the denominator (coded here as sign:interchange) and looks to be approximately equal to 3 (173.4/57.9). Both residual plots look pretty good.

Using the afex package we perform the correct split-plot analysis. Notice that the tests for split-plot results are the same as with `lm()`, and the $F=2.99$ on 1 and 18 df for the whole plot has the correct denominator df but is not significant with $P=.10$. Looking at pairwise comparisons all three timing levels are different. The interaction plot agrees with the test showing substantial parallelism. The normal QQ plot looks fine and the latticed boxplots as well as the Bartlett test do not reveal any problems with homogeneity of variance.

Exercise 15.2

For 2^4 experiment confound ACD and ABD to get blocks of size 4 (generalized interaction = BC)

ACD even: (1), AC, AD, CD, B, ABC, ABD, BCD

and ABD even: (1), AD, ABC, BCD

and ABD odd: AC, CD, B, ABD

ACD odd: A, C, D, ACD, AB, BC, BD, ABCD

and ABD even: C, ACD, AB, BD

and ABD odd: A, D, BC, ABCD

Problem 15.1b

Since temperatures vary from front to back in the growth chamber, we at least want to block on front to back distance. If we used a randomized complete block design with 8 blocks, we would have:

Source df

Soil 7

Block 7

Error 49

Total 63

Problem 15.4

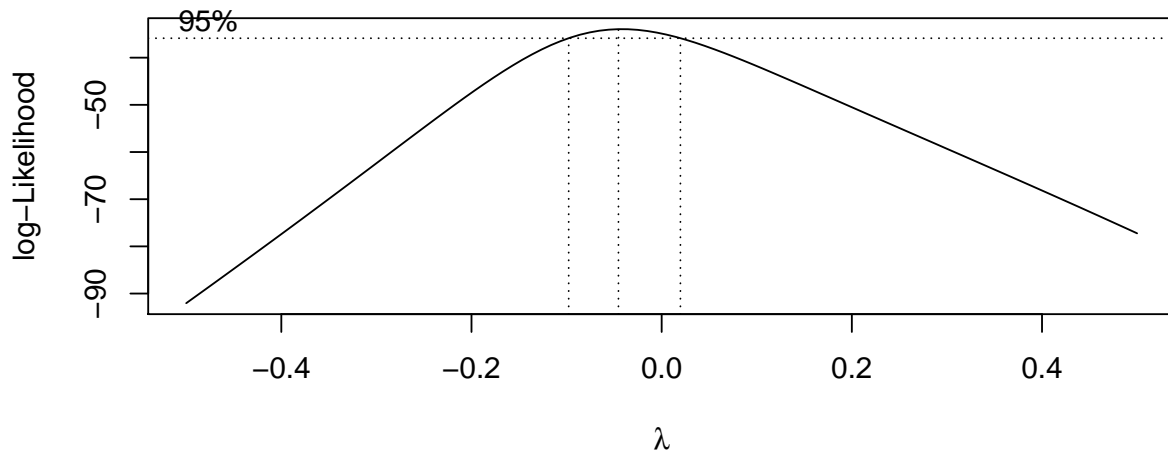
```
# first read the data directly from Dr. Oehlerts website

Prb15.4 <- read.table("http://www.stat.umn.edu/~gary/book/fcdae.data/pr15.4",header=TRUE)

Prb15.4$judge <- as.factor(Prb15.4$judge)
Prb15.4$A <- as.factor(Prb15.4$A)
Prb15.4$B <- as.factor(Prb15.4$B)
Prb15.4$C <- as.factor(Prb15.4$C)
Prb15.4$D <- as.factor(Prb15.4$D)

Prb15.4$logy <- log(Prb15.4$y)

# ----- use Box-Cox transformation -----
library(MASS)
#boxcox(y ~ judge + A*B*C*D, data=Prb15.4, lambda = seq(-2.00, 2.00, length = 50))
boxcox(y ~ judge + A*B*C*D, data=Prb15.4, lambda = seq(-0.50, 0.50, length = 50))
```



```
# -----

library(emmeans)

Prb15.4.lm1 <- lm(logy ~ judge + A*B*C*D, data = Prb15.4)

anova(Prb15.4.lm1)

## Analysis of Variance Table
##
## Response: logy
##          Df Sum Sq Mean Sq  F value    Pr(>F)
## judge      7  60.710   8.673  11.7590 0.0001586 ***
## A           1  98.026  98.026 132.9079 7.548e-08 ***
## B           1  74.362  74.362 100.8229 3.426e-07 ***
## C           1   0.862   0.862   1.1686 0.3009266
## D           1 10.788  10.788  14.6273 0.0024193 **
## A:B         1   2.350   2.350   3.1863 0.0995380 .
## A:C         1   1.125   1.125   1.5251 0.2404777
## B:C         1   0.663   0.663   0.8986 0.3618536
## B:D         1   0.002   0.002   0.0029 0.9582659
## C:D         1   2.166   2.166   2.9364 0.1122962
## A:B:D       1   1.175   1.175   1.5936 0.2307968
## A:C:D       1   0.006   0.006   0.0082 0.9293287
## A:B:C:D     1   0.407   0.407   0.5520 0.4717972
## Residuals 12   8.851   0.738
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
emmeans(Prb15.4.lm1, ~A)
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```
## A emmean SE df lower.CL upper.CL
## 1 8.6 0.215 12 8.13 9.06
## 2 5.1 0.215 12 4.63 5.56
##
## Results are averaged over the levels of: judge, B, C, D
## Confidence level used: 0.95
```

```
emmeans(Prb15.4.lm1, ~B)
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

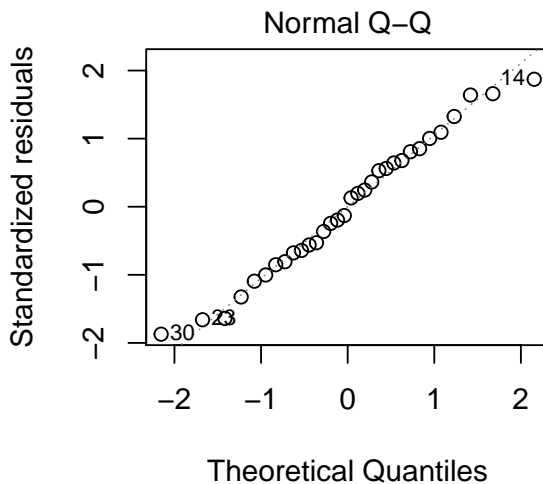
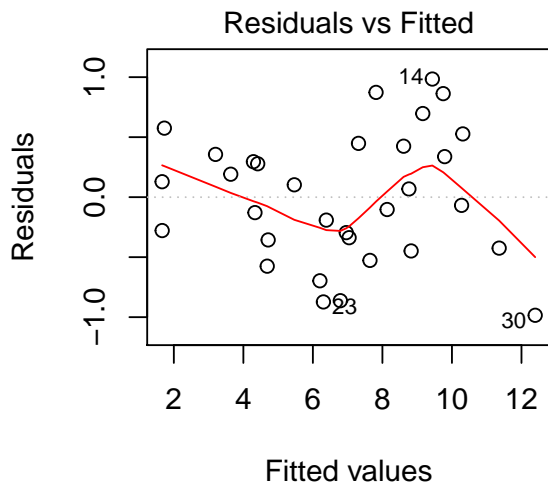
```
## B emmean SE df lower.CL upper.CL
## 1 8.37 0.215 12 7.90 8.84
## 2 5.32 0.215 12 4.85 5.79
##
## Results are averaged over the levels of: judge, A, C, D
## Confidence level used: 0.95
```

```
emmeans(Prb15.4.lm1, ~D)
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```
## D emmean SE df lower.CL upper.CL
## 1 7.43 0.215 12 6.96 7.89
## 2 6.26 0.215 12 5.80 6.73
##
## Results are averaged over the levels of: judge, A, B, C
## Confidence level used: 0.95
```

```
par(mfrow=c(1,2))
plot(Prb15.4.lm1, which = c(1,2))
```




```
par(mfrow=c(1,1))

# examine potential interactions
#lsmip(Prb15.4.lm1, A ~ B)
```

This is a confounded block design where we have taken a quarter fraction of the 2^4 factorial to get blocks of size four, by confounding the ABC and BCD interactions and their generalized interaction $(ABC)(BCD)=AD$. From the analysis we see that only the main effects of A, B, and D are significant at the .05 level. The main effect means are printed to show how the levels differ. The normal QQ plot looks good and although the residual by predicted plot has an odd curved shape it does not show evidence of variance heterogeneity. One potential problem with the design is that both the A and D effects were significant, but we are unable to investigate the potential AD interaction because it is confounded.

Problem 16.8

```
library(afex)

library(emmeans)

#-----#
# Mauchly test and downweighted degrees of freedom #
#-----#

Pr16.8.afex1 <- aov_ez(id="interchange",dv="speed",data=Pr16.8, between=c("sign"),within=c("timing"))

## Contrasts set to contr.sum for the following variables: sign

# the summary function produces Type III anova results assuming sphericity, and includes the Mauchly t

summary(Pr16.8.afex1)

## Warning in summary.Anova.mlm(object$Anova, multivariate = FALSE): HF eps > 1
## treated as 1

##
## Univariate Type III Repeated-Measures ANOVA Assuming Sphericity
##
##           Sum Sq num Df Error SS den Df  F value Pr(>F)
## (Intercept) 47489     1 1042.87   18 819.6668 <2e-16 ***
## sign         173     1 1042.87   18   2.9929 0.1007
## timing      2576     2   78.73   36 588.9780 <2e-16 ***
## sign:timing    4     2   78.73   36   0.8459 0.4375
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Mauchly Tests for Sphericity
##
##           Test statistic p-value
```

```
## timing          0.98788 0.90152
## sign:timing      0.98788 0.90152
##
##
## Greenhouse-Geisser and Huynh-Feldt Corrections
## for Departure from Sphericity
##
##          GG eps Pr(>F[GG])
## timing    0.98802    <2e-16 ***
## sign:timing 0.98802    0.4364
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##          HF eps  Pr(>F[HF])
## timing    1.109115 3.146242e-28
## sign:timing 1.109115 4.375261e-01
```

```
# the nice function produces Type III anova results correcting for
# possible lack of sphericity by downweighting split plot degrees of freedom
```

```
nice(Pr16.8.afex1)
```

```
## Anova Table (Type 3 tests)
##
## Response: speed
##      Effect      df  MSE      F ges p.value
## 1      sign         1, 18 57.94    2.99 .134  .101
## 2      timing 1.98, 35.57 2.21 588.98 *** .697  <.001
## 3 sign:timing 1.98, 35.57 2.21    0.85 .003  .436
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
##
## Sphericity correction method: GG
```

Since this data came from a split-plot experiment, it is unlikely that any repeated measures analysis will be needed, but we look at them as examples. The `summary()` output from our `aov_ez` model shows that the test for sphericity is not significant ($P = .90$) and that the epsilon adjustments are either .988 or 1.11, making no change to the original conclusions. We would not use the HF value if it were above 1, but the output has a message telling us that it set the HF value equal to 1.

```
#-----#
# Random coefficient models #
#-----#
```

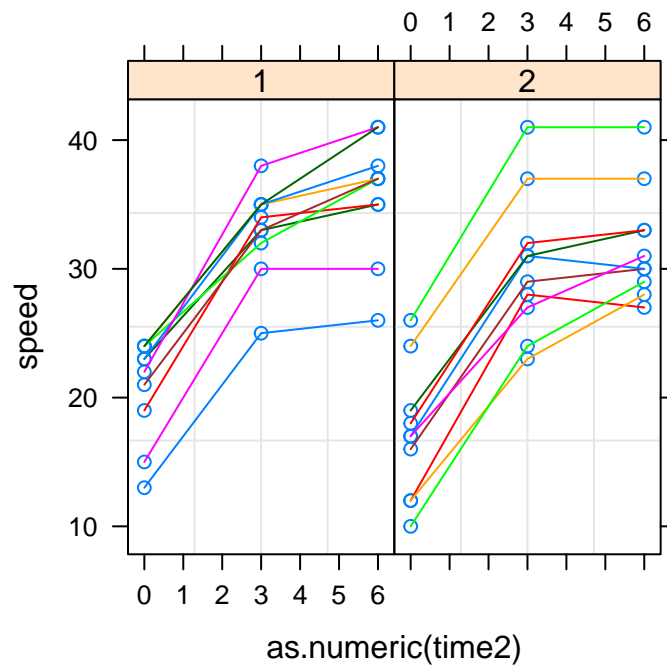
```
library(nlme)
```

```
##
## Attaching package: 'nlme'

## The following object is masked from 'package:lme4':
##
##      lmList
```

```
# first we use the groupedData function to create a data structure appropriate for our plot
Pr16.8.longplot <- groupedData(speed ~ as.numeric(time2) | interchange, outer = ~ sign, data = Pr16.8)
plot(Pr16.8.longplot, outer = ~ sign)
```

- | | | | | |
|------|-----|------|------|------|
| — 1 | — 2 | — 3 | — 13 | — 11 |
| — 10 | — 4 | — 7 | — 19 | — 18 |
| — 6 | — 9 | — 12 | — 16 | — 20 |
| — 8 | — 5 | — 14 | — 17 | — 15 |



```
Pr16.8.afexrc1 <- mixed(speed ~ sign +time2 +sign:time2 +(1 + time2|interchange),
  data = Pr16.8)
```

```
## Contrasts set to contr.sum for the following variables: sign, interchange

## Numerical variables NOT centered on 0: time2
## If in interactions, interpretation of lower order (e.g., main) effects difficult.

## Fitting one lmer() model.

## boundary (singular) fit: see ?isSingular

## [DONE]
## Calculating p-values. [DONE]
```

```
summary(Pr16.8.afexrc1)
```

```

## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: speed ~ sign + time2 + sign:time2 + (1 + time2 | interchange)
## Data: data
##
## REML criterion at convergence: 346.2
##
## Scaled residuals:
## Min 1Q Median 3Q Max
## -1.3522 -0.5540 -0.2293 0.5882 2.0181
##
## Random effects:
## Groups Name Variance Std.Dev. Corr
## interchange (Intercept) 1.561e+01 3.95086
## time2 1.344e-04 0.01159 -1.00
## Residual 1.193e+01 3.45407
## Number of obs: 60, groups: interchange, 20
##
## Fixed effects:
## Estimate Std. Error df t value Pr(>|t|)
## (Intercept) 20.708333 1.130297 19.807539 18.321 6.92e-14 ***
## sign1 1.675000 1.130297 19.807539 1.482 0.154
## time2 2.475000 0.182064 37.985746 13.594 3.60e-16 ***
## sign1:time2 0.008333 0.182064 37.985746 0.046 0.964
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
## (Intr) sign1 time2
## sign1 0.000
## time2 -0.494 0.000
## sign1:time2 0.000 -0.494 0.000
## convergence code: 0
## boundary (singular) fit: see ?isSingular

```

```
nice(Pr16.8.afexrc1)
```

```

## Mixed Model Anova Table (Type 3 tests, S-method)
##
## Model: speed ~ sign + time2 + sign:time2 + (1 + time2 | interchange)
## Data: Pr16.8
## Effect df F p.value
## 1 sign 1, 19.81 2.20 .154
## 2 time2 1, 37.99 184.80 *** <.001
## 3 sign:time2 1, 37.99 0.00 .964
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1

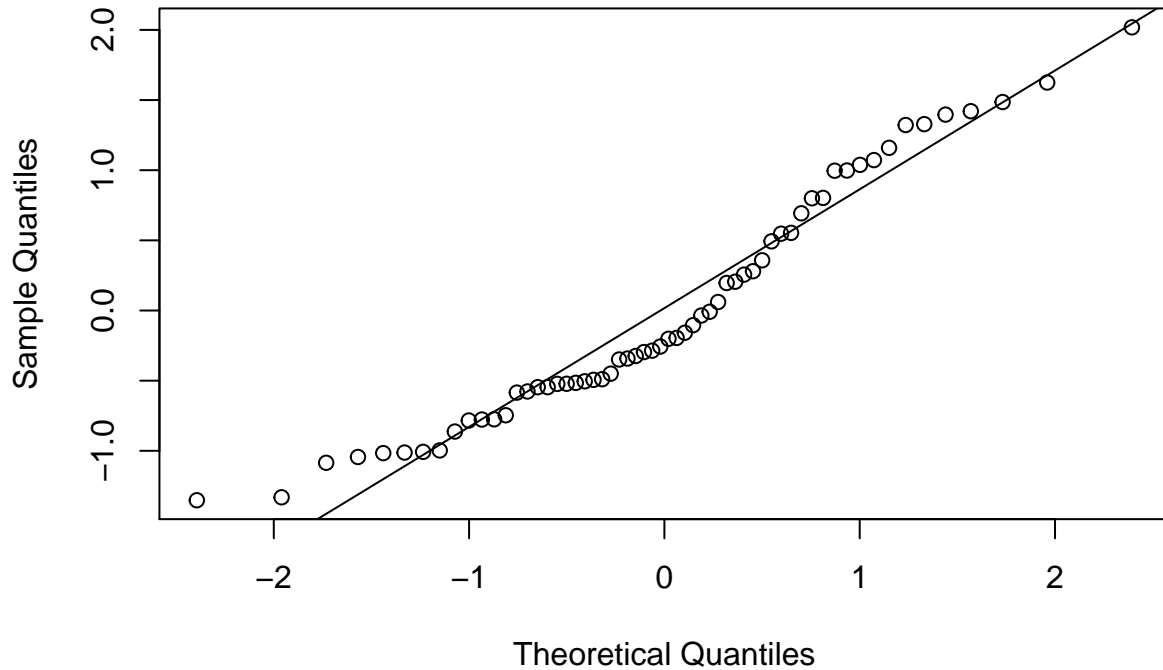
```

```

qqnorm(summary(Pr16.8.afexrc1)$residuals)
qqline(summary(Pr16.8.afexrc1)$residuals)

```

Normal Q-Q Plot



The plot shows an increasing trend for speed at each individual interchange as timing increases. The trends are usually not exactly linear but contain a strong linear pattern. The tests for the fixed effect coefficients show that the slope for time is significant ($F = 184.8$ on 1 and 18 df) but that intercepts and slopes do not differ by sign ($P = .16$ and $.96$ respectively).

```
#-----#  
# Modeling covariance structure #  
#-----#  
  
# compound symmetric covariance structure  
  
Pr16.8.cs <- gls(speed ~ sign * timing, data = Pr16.8,  
               corr = corCompSymm(, form= ~ 1 | interchange) )  
#summary(Pr16.8.cs)  
summary(Pr16.8.cs)$logLik ; summary(Pr16.8.cs)$AIC ; summary(Pr16.8.cs)$BIC
```

```
## [1] -134.1505
```

```
## [1] 284.301
```

```
## [1] 300.2129
```

```
anova(Pr16.8.cs)
```

```
## Denom. DF: 54
```

```
##           numDF  F-value p-value
## (Intercept)      1 819.6682 <.0001
## sign             1   2.9929 0.0893
## timing           2 588.9775 <.0001
## sign:timing       2   0.8459 0.4348
```

```
# order 1 autoregressive covariance structure
```

```
Pr16.8.ar1 <- gls(speed ~ sign * timing, data = Pr16.8,
  corr = corAR1(, form= ~ 1 | interchange) )
```

```
#summary(Pr16.8.ar1)
```

```
summary(Pr16.8.ar1)$logLik ; summary(Pr16.8.ar1)$AIC ; summary(Pr16.8.ar1)$BIC
```

```
## [1] -137.1468
```

```
## [1] 290.2936
```

```
## [1] 306.2054
```

```
anova(Pr16.8.ar1)
```

```
## Denom. DF: 54
```

```
##           numDF  F-value p-value
## (Intercept)      1 774.2975 <.0001
## sign             1   3.7185 0.0591
## timing           2 368.7212 <.0001
## sign:timing       2   1.1385 0.3279
```

```
# order 1 heterogeneous autoregressive covariance structure
```

```
Pr16.8.arh1 <- gls(speed ~ sign * timing, data = Pr16.8,
  corr = corAR1(form = ~ 1 | interchange), weights = varIdent(form = ~ 1 | timing) )
```

```
#summary(Pr16.8.arh1)
```

```
summary(Pr16.8.arh1)$logLik ; summary(Pr16.8.arh1)$AIC ; summary(Pr16.8.arh1)$BIC
```

```
## 'log Lik.' -137.0877 (df=2)
```

```
## [1] 294.1754
```

```
## [1] 314.0653
```

```
anova(Pr16.8.arh1)
```

```
## Denom. DF: 54
```

```
##           numDF  F-value p-value
## (Intercept)      1 767.8704 <.0001
## sign             1   4.1547 0.0464
## timing           2 369.4014 <.0001
## sign:timing       2   1.1134 0.3359
```

```

# an unstructured covariance matrix

Pr16.8.un <- gls(speed ~ sign * timing, data = Pr16.8,
  corr=corSymm(form = ~ 1 | interchange), weights = varIdent(form = ~ 1 | timing) )

#summary(Pr16.8.un)
summary(Pr16.8.un)$logLik ; summary(Pr16.8.un)$AIC ; summary(Pr16.8.un)$BIC

## 'log Lik.' -134.0063 (df=2)

## [1] 292.0126

## [1] 315.8805

```

```
anova(Pr16.8.un)
```

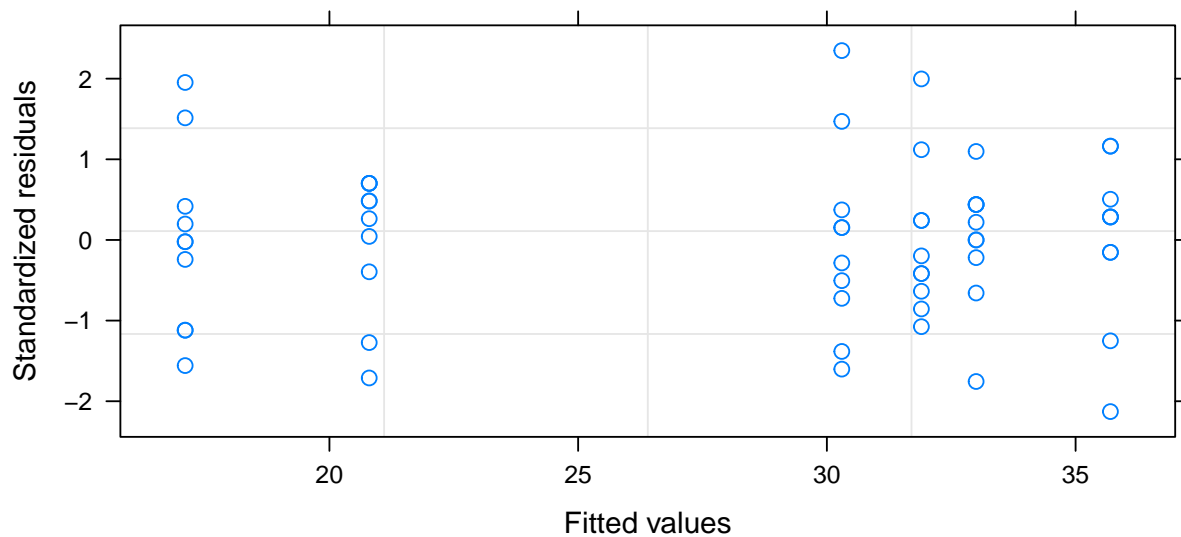
```

## Denom. DF: 54
##          numDF  F-value p-value
## (Intercept)      1 865.4615 <.0001
## sign              1  3.2325  0.0778
## timing            2 626.1664 <.0001
## sign:timing        2  0.7728  0.4668

```

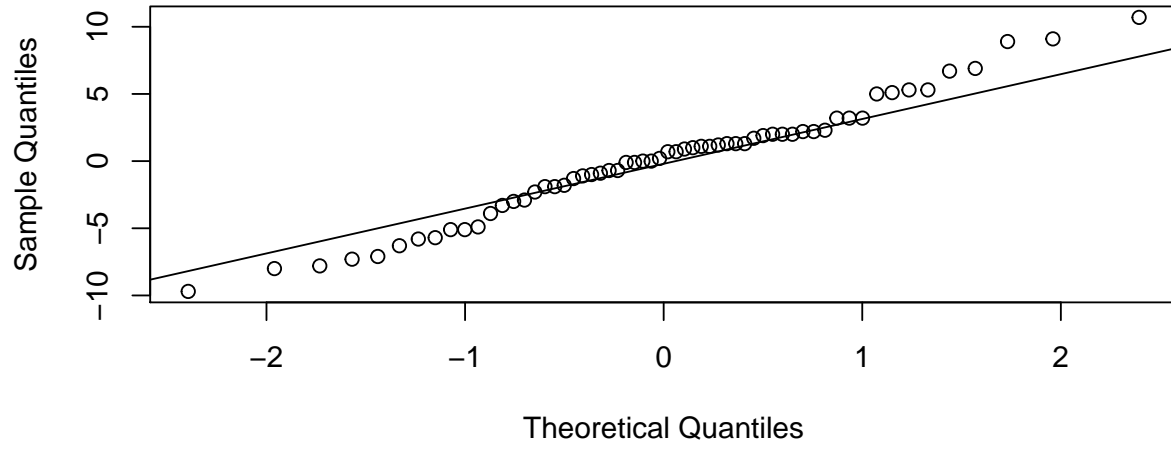
```
# get diagnostic plots for the CS model
```

```
plot(Pr16.8.cs)
```



```
qqnorm(Pr16.8.cs$residuals) ; qqline(Pr16.8.cs$residuals)
```

Normal Q-Q Plot



The best fitting model according to both AIC and BIC is the CS model. The residual plots look good.