

STAT5303HW10

Mingming Xu

11/20/2020

E15.1, E15.3, P15.1(a-e), P15.2(a-e), P15.3, P15.4

```
library(cfcdae)
```

```
## Registered S3 method overwritten by 'DoE.base':  
##   method          from  
##   factorize.factor conf.design
```

```
library(car)
```

```
## Loading required package: carData
```

```
## Registered S3 methods overwritten by 'car':  
##   method          from  
##   influence.merMod      lme4  
##   cooks.distance.influence.merMod lme4  
##   dfbeta.influence.merMod      lme4  
##   dfbetas.influence.merMod     lme4
```

```
library(knitr)
```

E15.1

I will consider to use the even odd method with two defining contrasts BCD and ACD to get four blocks of eight.

Table 1: Blocks

Treatments	ACD	BCD
(1)	e	e
a	o	e
b	e	o
c	o	o
d	o	o
ab	o	o
ac	e	o
ad	e	o
bc	o	e
bd	o	e
cd	e	e
abc	e	e
abd	e	e
acd	o	e
bcd	e	o
abcd	o	o

e_e	e_o	o_e	o_o
(1)	b	a	c
cd	ac	bc	d
abc	ab	bd	ab
abd	bcd	acd	abcd

E15.3

Table 3: 15 Confunded Effects

Generators	TwoWay	ThreeWay	FourWay
ABCF	CDEF	AF	ABCDEFH
ABDE	BDEF	BEFH	
ACDE	ADFH	CEFH	
BCDH	BC	DH	
	ACEF		
	ABEF		

P15.1

- In the experiment, the treatment is different spatial scales and I will specifically set the levels of treatment are a meter of each other, two meters of each other, three meters, four meters and five meters of each other. For each level, there are 10 measurement. Besides, 10 blocks of 5 units each can be used. Thus, we can do a randomized complete design.
- In this experiment, we have three factors with 2 levels so that it is a two series factorial design. So, we have 8 combinations of treatments in total and each combination has 2 measurments. Besides, we treat the variable day as a block. So, we have 4 blocks with 4 units in each block. Thus, we can do an incomplete block design.
- In this experiment, we have four different designs of arterial stints and three major coronary arteries so that we have 12 combinations. And, we can let animal be the block. It follow that we have 12 blocks with 3 units in each block. This is an incomplete block design.
- In this experiment, we have three factors and all them has 2 levels so it is a two series factorial design. In total, we have 8 combinations and each combination has 6 measurements. Because we can only make eight measurements before the equipment needs to be cleaned and recalibrated, we can treat the equipment as blocks. Thus, we have 6 blocks and each block has 8 units. It is an incomplete block design.
- In this experiment, we have three factors A,B and C, and, all factors has 2 levels. If we do not think there is large differences between devices, we can use these three factors to do ANOVA test. However, if we expect large device to device differences, we should let the device be block. Further, the block design is better.

P 15.2

- This is an incomplete block design. We have 2 blocks. One of them contains a low rigor/pharma abstract and a high rigor/NIH abstract combinations. Another one contains a low rigor/NIH abstract and a high rigor/pharma abstract.

Source DF

Block 2

Attribution 1

Rigor 1

Attribution.Rigor 1

- This is an incomplete block design. The treatment is the brand of booties with 6 different brands. The vraiable day is viewd as block. So, we have 15 blocks and each block has 4 units.

Source DF

Day 14

Brand 5

Error 39

- (c) In this experiment, we have 2 factors with 2 level, which is a two series factorial design. Four combinations of treatment and each combination has 10 measurments. Baccuse the warmer is not uniformly warm and differs slightly in temperature from back to front, there is a difference in dishes. Thus, the dishes should be blocks in this experiment because it might cause variability. This is an incomplete block design with four blocks. For each row, the two dishes are in a block.

Source DF

Dishes 3

Choline 1

Protein 1

Choline.Protein 1

Error 32

- (d) Sessions and drugs are treatments with 2 levels. When we expect there exits diffrence in subjects, it is appropriate to treat subjects as blocks. Each block contains 2 combinations of session and drug. This is an incomplete block design.

Source DF

Subjects 29

Sessions 1

Drugs 1

Sessions.Drugs 1

Error 26

- (e) In this experiment, we have fifteen level combinations. Each combination has 2 replications. Apparently, the day is the block. Each block has 10 units which is 10 combinations of two levels of the charge and all levels additive.

Source DF

Day 2

Charge 2

Additive 4

Charge.Additive 8

Error 13

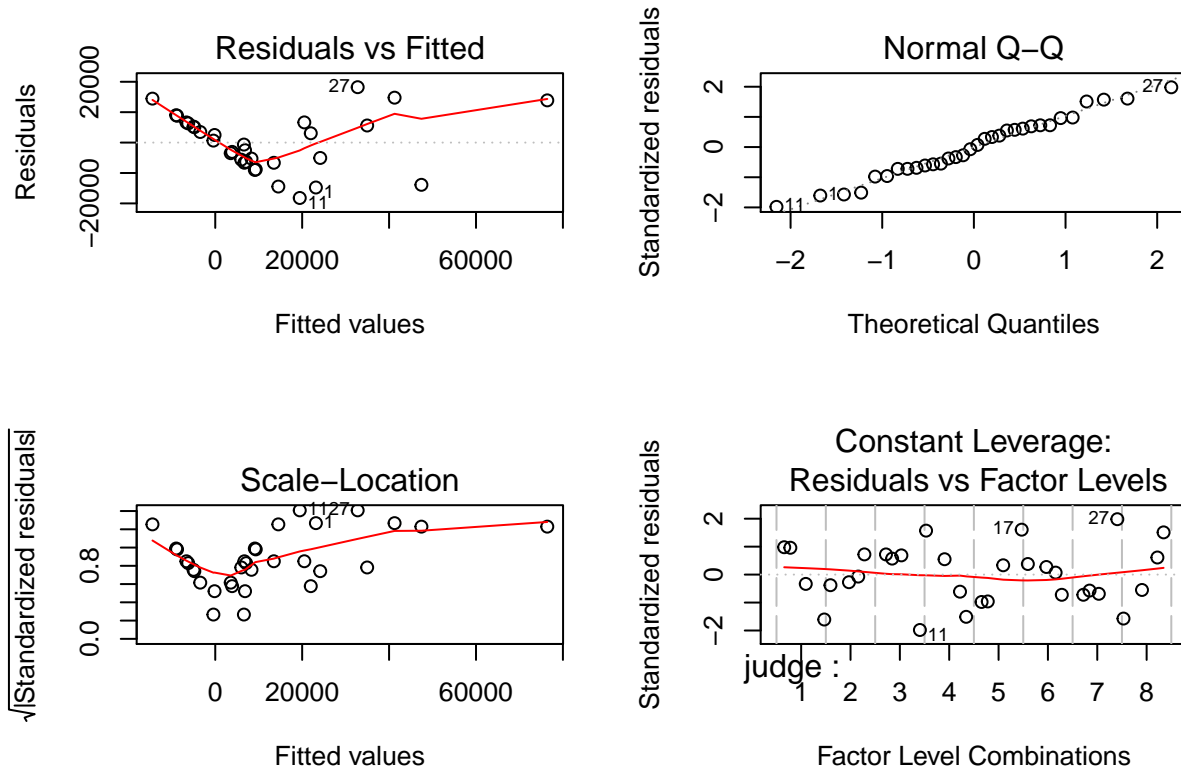
P15.3

```
data(OdorIntensity)
attach(OdorIntensity)
head(OdorIntensity)
```

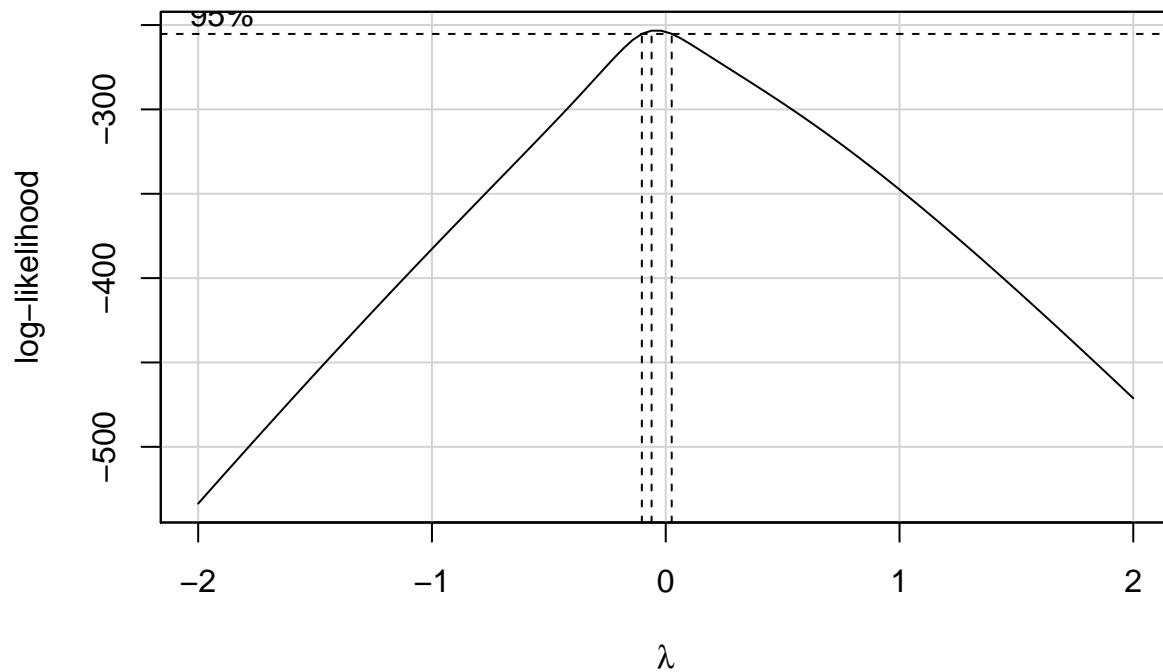
```
##   judge    A    B    C    D concentration
## 1     1 low low low low           8389
## 2     1 low high high low           816
## 3     1 high high low high           4
## 4     1 high low high high           46
```

```
## 5    2 high low low low    4351
## 6    2 high high high low    78
```

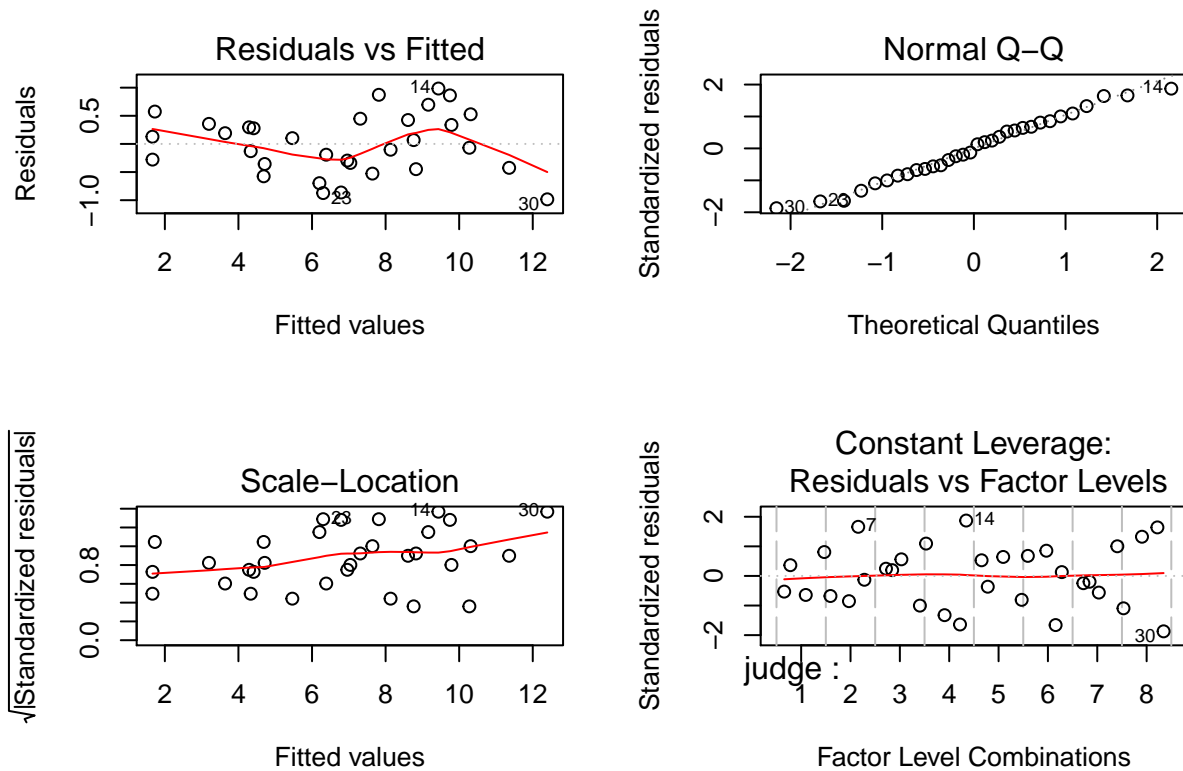
```
judge<-as.factor(judge)
##First fitting a full model with judge:
mod<-lm(concentration~judge+A*B*C*D,data=OdorIntensity )
##Checking assumptions:
par(mfrow=c(2,2))
plot(mod)
```



```
##Apparently, the constant variance is not good since it has a curve.
par(mfrow=c(1,1))
boxCox(mod)
```



```
##BoxCox suggests log
##Refitting the model:
mod_1<-lm(log(concentration)~judge+A*B*C*D,data=OdorIntensity)
##Checking assumptions again:
par(mfrow=c(2,2))
plot(mod_1)
```



```
##It is abvious that the constant variance and normality are improved.
```

```
par(mfrow=c(1,1))
mod_1$coefficients
```

```
## (Intercept)      judge1      judge2      judge3      judge4      judge5
## 6.845578696 -1.607072018  1.062719503 -2.080618471 -0.806662546  1.141405457
##      judge6      judge7      A1      B1      C1      D1
## -0.451579675  0.589800195 -1.750230497 -1.524401815 -0.164119967 -0.580632787
##      A1:B1      A1:C1      B1:C1      A1:D1      B1:D1      C1:D1
## -0.270995112 -0.187488653  0.143915100      NA -0.008112085 -0.260149961
##      A1:B1:C1      A1:B1:D1      A1:C1:D1      B1:C1:D1      A1:B1:C1:D1
##      NA      0.191649287 -0.013750007      NA      0.112794774
```

```
anova(mod_1)
```

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

```
##
```

```
## Response: log(concentration)
```

```
##      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
```

This is a two series factorial design with four factors and judge is as block. The ANOVA table shows the terms A,B,D and judge have significant difference so there is no significant interaction term.

P15.3

```
data("Milfoil")
attach(Milfoil)
```

```
## The following objects are masked from OdorIntensity:
```

```
##
```

```
##      A, B, C
```

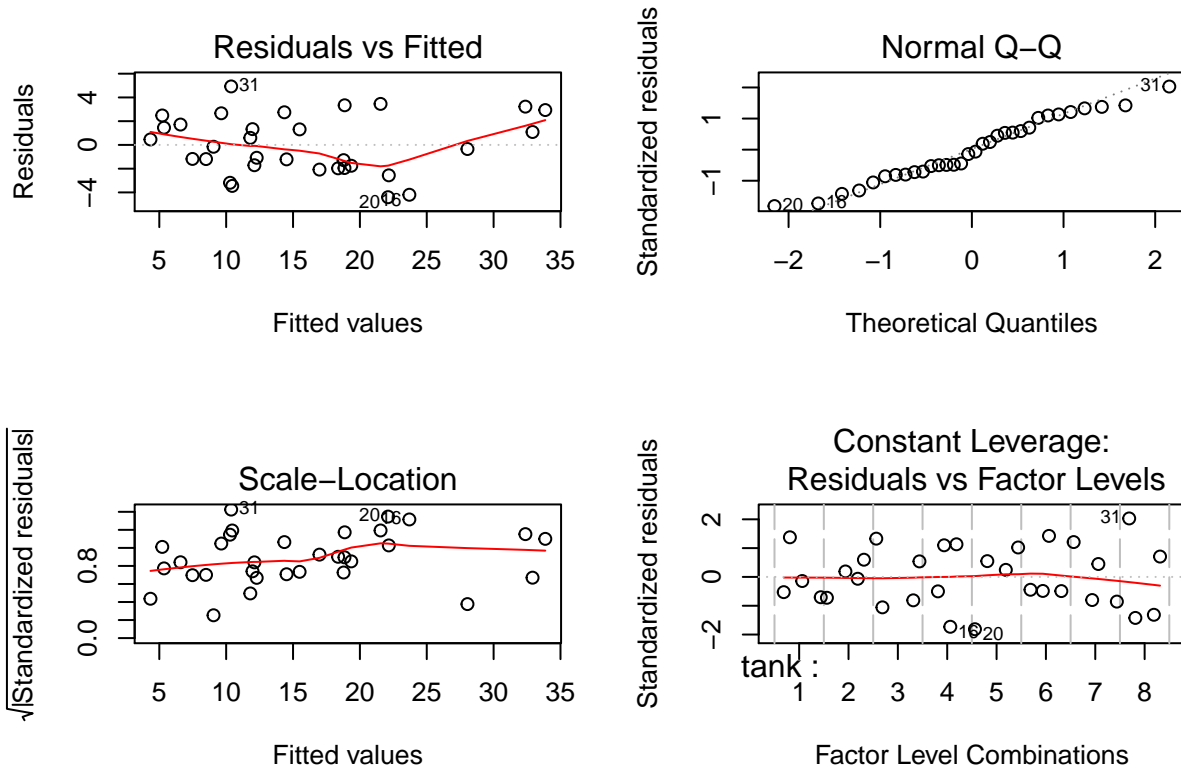
```
head(Milfoil)
```

```
##      tank      A      B      C biomass
## 1      1 low low low 10.4
## 2      1 high high low 17.5
## 3      1 high low high 22.2
## 4      1 low high high 27.7
## 5      2 high low low 4.8
## 6      2 low high low 8.9
```

```

tank<-as.factor(tank)
##First fitting a full model with judge:
lmod<-lm(biomass~tank+A*B*C,data=Milfoil )
##Checking assumptions:
par(mfrow=c(2,2))
plot(lmod)

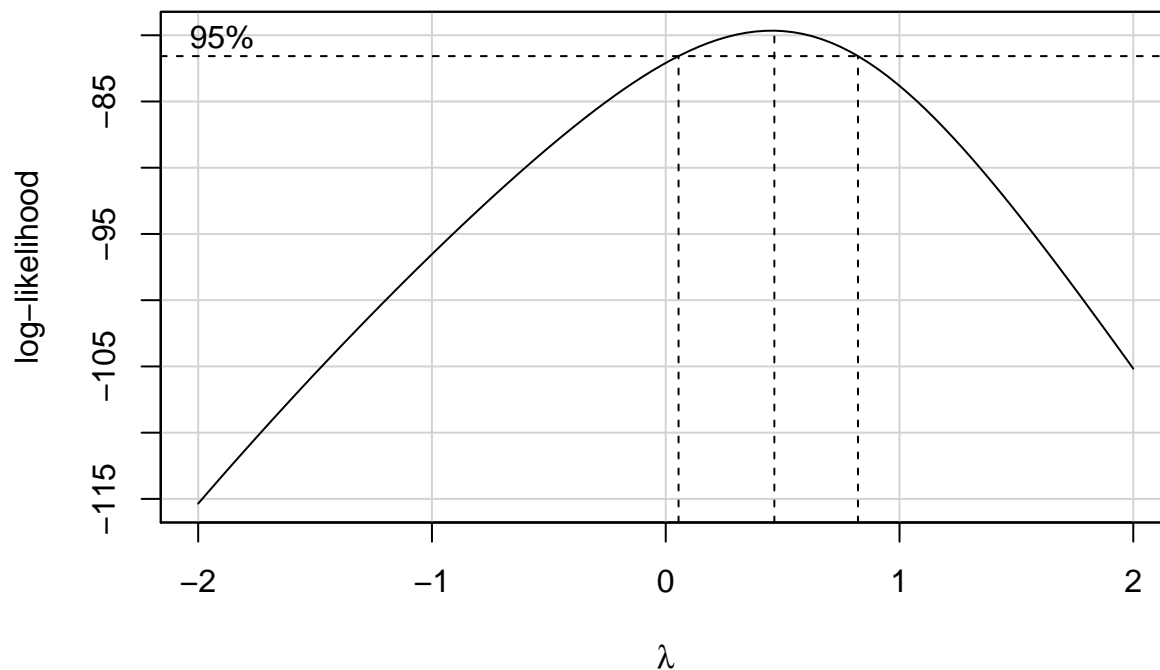
```



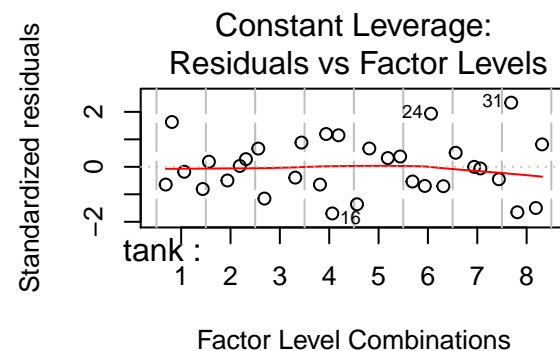
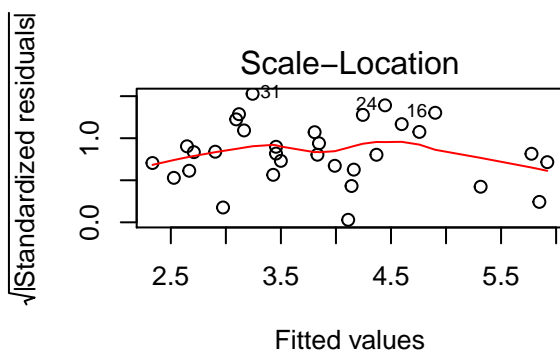
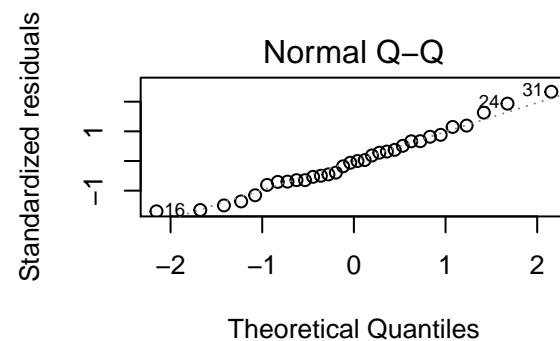
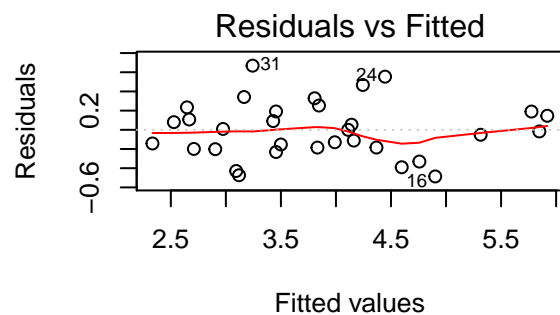
```

## The constant variance and normality is not bad but the boxCox suggest 0.5 power better.
par(mfrow=c(1,1))
boxCox(lmod)

```

```
##Refitting the model:
lmod_1<-lm(biomass^(0.5) ~tank+A*B*C,data=Milfoil )
##Checking assumptions again:
par(mfrow=c(2,2))
plot(lmod_1)
```



```
## The constant variance and normality have been improved.
par(mfrow=c(1,1))
```

```
anova(lmod_1)
```

```
## Analysis of Variance Table
##
## Response: biomass^(0.5)
##           Df  Sum Sq Mean Sq F value    Pr(>F)
## tank       7 15.3510   2.1930 14.2064 5.597e-06 ***
## A          1  0.3323   0.3323  2.1527  0.16057
## B          1  8.9774   8.9774 58.1561 6.961e-07 ***
## C          1  4.7941   4.7941 31.0566 3.367e-05 ***
## A:B        1  0.0522   0.0522  0.3380  0.56863
## A:C        1  0.0280   0.0280  0.1815  0.67543
## B:C        1  0.4796   0.4796  3.1070  0.09593 .
## A:B:C      1  0.0687   0.0687  0.4448  0.51376
## Residuals 17  2.6242   0.1544
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lmod_1$coefficientsc
```

```
## NULL
```

This is a two series factorial design with 3 factors and tank is as block. The ANOVA table shows only the terms B,C and tank have significant difference and none of the interaction terms is significant.