Solution to Series 4

1. **a)** Test for treatment differences without taking into account the initial hormone concentration. Estimate the treatment means.

```
> feed <- read.table(file="http://stat.ethz.ch/Teaching/Datasets/feed.txt",header=TRUE)
> feed$Feeding <- as.factor(feed$Feeding)
> modF <- aov(Final~Feeding,data=feed)
> summary(modF)
            Df Sum Sq Mean Sq F value Pr(>F)
Feeding      2   1083   541.4   0.629   0.54
Residuals   29  24971   861.1
> TukeyHSD(modF,"Feeding", conf.level=0.95)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Final ~ Feeding, data = feed)

$Feeding
         diff       lwr      upr     p adj
2-1 11.555556 -20.40036 43.51147 0.6489302
3-1 14.010101 -18.56238 46.58259 0.5446072
3-2  2.454545 -27.79581 32.70490 0.9781225
```

The treatment means are estimated as shown in the R-output below.

```
> summary.lm(aov(Final~Feeding-1,data=feed))
Call:
aov(formula = Final ~ Feeding - 1, data = feed)

Residuals:
    Min      1Q  Median      3Q     Max
-55.000 -18.614   4.778  21.136  46.000

Coefficients:
         Estimate Std. Error t value Pr(>|t|)
Feeding1  220.444      9.781   22.54   <2e-16 ***
Feeding2  232.000      8.471   27.39   <2e-16 ***
Feeding3  234.455      8.848   26.50   <2e-16 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 29.34 on 29 degrees of freedom
Multiple R-squared:  0.9854,        Adjusted R-squared:  0.9839
F-statistic: 653.4 on 3 and 29 DF,  p-value: < 2.2e-16
```

b) Carry out a one-way analysis of variance for the differences $D_i = Y_i - x_i$ of hormone measurements, where Y_i is the response after treatment and x_i the baseline measurement.

```
> modF2 <- aov((Final-Initial)~Feeding,data=feed)
> summary(modF2)
            Df Sum Sq Mean Sq F value Pr(>F)
Feeding      2    101    50.7   0.154  0.858
Residuals   29   9574   330.2
```

c) Include the baseline measurement in the model as a covariate and do an analysis of covariance for the responses Y_i. Estimate the adjusted treatment means.

```
> modF3 <- aov(Final~Feeding+Initial,data=feed)
> summary(modF3)
            Df Sum Sq Mean Sq F value   Pr(>F)
Feeding      2   1083     541   10.45 0.000408 ***
Initial      1  23520   23520  453.88  < 2e-16 ***
Residuals   28   1451      52
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The estimates for the adjusted treatments means are calculated as follows.

```
> summary.lm(aov(Final~Feeding+Initial-1,data=feed))
Call:
aov(formula = Final ~ Feeding + Initial - 1, data = feed)

Residuals:
     Min       1Q   Median       3Q      Max
-13.2306  -5.4712   0.3818   4.6875  13.0846

Coefficients:
          Estimate Std. Error t value Pr(>|t|)
Feeding1 -280.8018    23.6499  -11.87 1.91e-12 ***
Feeding2 -302.3932    25.1696  -12.01 1.45e-12 ***
Feeding3 -311.6981    25.7274  -12.12 1.19e-12 ***
Initial     2.4254     0.1138   21.30  < 2e-16 ***
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.199 on 28 degrees of freedom
Multiple R-squared:  0.9992,        Adjusted R-squared:  0.999
F-statistic:  8257 on 4 and 28 DF,  p-value: < 2.2e-16
```

d) Compare and comment the different results.

The results from tasks a) and b) show that the factor feeding is neither significant for the final hormone concentration nor for the difference between the final and the initial concentration. This can also be seen in Figure 1.

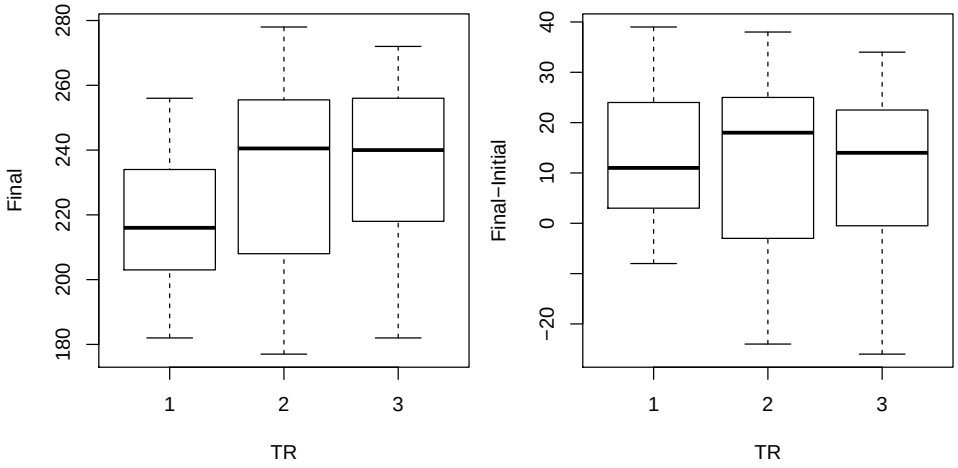

Figure 1: Final vs. treatment and final-initial vs. treatment.

On the other hand, if we include the initial hormone concentration as a covariate, the factor feeding is quite significant, as shown in task c). What is the reason for this?

The model used in b) is

$$(Final - Initial)_{ij} = \mu + A_i + \epsilon_{ij},$$

where A_i denotes the effect of feed composition i. This is equivalent to

$$Final_{ij} = \mu + 1 \cdot Initial_{ij} + A_i + \epsilon_{ij}.$$

Roughly speaking, this means that we assume that the coefficient of the covariate initial is 1. Figure 2, however, demonstrates that this is not a reasonable assumption. When ploting the final versus the initial concentration, the points clearly do not lie on the line with slope 1. This holds true for all three treatment groups. The same argumentation also shows that the model in a) is not adequate since, there, we effectively assume a coefficient of initial being 0.

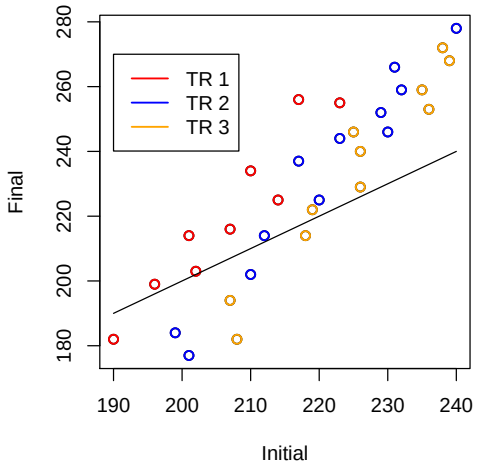

Figure 2: Final vs. initial.

Further, looking at fitted model of task c), we see that the estimated coefficient of initial is about 2.4 which is far from 1.

```
> summary.lm(modF3)
Call:
aov(formula = Final ~ Feeding + Initial, data = feed)

Residuals:
     Min       1Q   Median       3Q      Max
-13.2306  -5.4712   0.3818   4.6875  13.0846

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -280.8018    23.6499 -11.873 1.91e-12 ***
Feeding2     -21.5914     3.5351  -6.108 1.37e-06 ***
Feeding3     -30.8963     3.8616  -8.001 1.03e-08 ***
Initial        2.4254     0.1138  21.304  < 2e-16 ***
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.199 on 28 degrees of freedom
Multiple R-squared:  0.9443,        Adjusted R-squared:  0.9383
F-statistic: 158.3 on 3 and 28 DF,  p-value: < 2.2e-16
```

In other words, a model which does not include the initial concentration as covariate has residuals that are correlated. Consequently, the entire analysis, and in particular significance tests, is flawed. To see that there is correlation in the residuals, we look at the initial division of the animals in Figure 3. We can see that animals are not really divided randomly (in the first group example we have smaller animals). With a randomized division of the animals we probably would have obtained better results even for the ANOVA-table. When designing such a study aim at dividing probands at random into the different groups. In any case you have to avoid that all probands with a common feature are in the same group. This could lead to not noticing relevant effects or worse merging effects that do not exist!

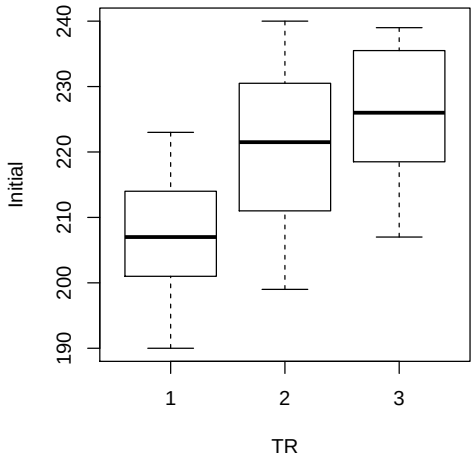

Figure 3: Initial vs. treatment.

2. We take the first replicate of the dataset softdrinkANOVA.txt, i.e.

	score	sugar	soda	water	temp	BLOCK
1	159	1	1	1	1	2
3	168	2	1	1	1	1
5	158	1	1	2	1	1
7	166	2	1	2	1	2
9	175	1	2	1	1	1
11	179	2	2	1	1	2
13	173	1	2	2	1	2
15	179	2	2	2	1	1
17	164	1	1	1	2	1
19	187	2	1	1	2	2
21	163	1	1	2	2	2
23	185	2	1	2	2	1
25	168	1	2	1	2	2
27	197	2	2	1	2	1
29	170	1	2	2	2	1
31	194	2	2	2	2	2

We have $16 = 2^4$ observations.
We wants to divide the observations in $\frac{16}{8} = 2$ different blocks such that we have a new factor (BLOCK) with 2 levels.
Construction of the experiment:
call:

```
A=sugar-effect
B=soda-effect
C=water-effect
D=temp-effect
E=BLOCK-effect
```

The values of A, B, C and D are 1 or -1 (or equivalently $+$ or $-$).
We just have to find the values of the column E to construct our experiment. We know that $E = A{\cdot}B{\cdot}C{\cdot}D$ (because $ABCD$ confounded) hence, the column BLOCK will be determined by multiplying the column of A, B, C and D. We let 1 correspond to the $first\ block$ and 2 correspond to the $second\ block$. We obtain:

```
> softBL
```

	score	sugar	soda	water	temp	BLOCK
1	159	1	1	1	1	2
3	168	2	1	1	1	1

```
5    158     1    1     2    1     1
7    166     2    1     2    1     2
9    175     1    2     1    1     1
11   179     2    2     1    1     2
13   173     1    2     2    1     2
15   179     2    2     2    1     1
17   164     1    1     1    2     1
19   187     2    1     1    2     2
21   163     1    1     2    2     2
23   185     2    1     2    2     1
25   168     1    2     1    2     2
27   197     2    2     1    2     1
29   170     1    2     2    2     1
31   194     2    2     2    2     2
```

Note that there is no reason to divide an already performed experiment in different blocks, but if we have to redo the experiment and can, for example, just test 8 combinations per day, the above division in blocks is useful.

Now we make an analysis of variance of the data with the block factor:

```
> sB.fit <- aov(score~sugar+soda+water+temp+BLOCK,data=softBL)
> summary(sB.fit)

            Df Sum Sq Mean Sq F value   Pr(>F)
sugar        1  976.6   976.6  26.138 0.000456 ***
soda         1  451.6   451.6  12.086 0.005956 **
water        1    5.1     5.1   0.135 0.720472
temp         1  315.1   315.1   8.433 0.015728 *
BLOCK        1    3.1     3.1   0.082 0.780493
Residuals   10  373.6    37.4
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

We conclude that sugar and soda are relevant at a 1% level (and temperature is relevant at a 5% level). If, additionally, we want to compute the 2-way effects we just have to type

```
aov(score~(sugar+soda+water+temp+BLOCK)^2,data=softBL)
```

If we want to compute all the n-way effects we just have to type

```
> sB.2k <- aov(score~sugar*soda*water*temp*BLOCK,data=softBL)
```

In this case 3 and 4-way effects are confounded, it follows that we obtain the same result as in the previous two function calls.
Furthermore, if we want to do an analysis of variance we can not look at all the 1 and 2-way effects because otherwise we lose all the degrees of freedom for the residuals!

Remark
With `sB.2k$coef` we can see that the 3 and 4-way effects are confounded (Effects are market with `NA`).

3. We have the following:

- $8 = 2^3 = 2^{k-l}$ runs,
- 5 two-level factors, thus: $k = 5$,
- consequently we need $l = 5 - 3 = 2$ "confounding relations".

Solution:
STEP 1:
Write down the complete 2^3 table.

A	B	C
−	−	−
+	−	−
−	+	−
+	+	−
−	−	+
+	−	+
−	+	+
+	+	+

STEP 2:
Define the "confounding relations". (If not specified otherwise confounding relations can be chosen quite freely).
We try to maximize the resolution without prior information on the dataset and choose: $D = -A \cdot B$ and $E = -A \cdot C$ (The $-$ is not necessary, but doing so our first run will be (1)).
We obtain:

A	B	C	D	E
−	−	−	−	−
+	−	−	+	+
−	+	−	+	−
+	+	−	−	+
−	−	+	−	+
+	−	+	+	−
−	+	+	+	+
+	+	+	−	−

STEP 3:
Now read every row of the matrix marking the factors with $+$ for high level:

A	B	C	D	E	Treatm.
−	−	−	−	−	(1)
+	−	−	+	+	ade
−	+	−	+	-	bd
+	+	−	−	+	abe
−	−	+	−	+	ce
+	−	+	+	−	acd
−	+	+	+	+	bcde
+	+	+	−	−	abc

Which effects are confounded with each other?
Because $l = 2$ every effect is confounded with $2^l = 2^2 = 4$ effects.
We know: $D = -A \cdot B$, so the effects of D and AB are not distinguishable (we write $D \cong AB$).

From $D \cong AB$ and $E \cong AC$ we get:

- $I \cong ABD \cong ACE \cong BCDE$

By multiplication we find:

- $A \cong BD \cong CE \cong ABCDE$
- $B \cong AD \cong ABCE \cong CDE$
- $C \cong AE \cong BDE \cong ABCD$
- $D \cong AB \cong BCE \cong ACDE$
- $E \cong AC \cong BCD \cong ABDE$
- $BC \cong ED \cong ABE \cong ACD$

- $BE \cong CD \cong ABC \cong AED$

Remark
The resolution of the experiment can be calculated as follows: take two effects which are confounded and count the number of letters you have. The minimal result you can obtain is the resolution. In our case:

- B & $AD \rightarrow 3$ (letters)
- D & $AB \rightarrow 3$ (letters)
- BCD & $ACDE \rightarrow 7$ (letters)
- ...

The resolution is 3 (not very high).

Can we improve the resolution by changing the relationships $D \cong AB$ and $E \cong AC$?[1]
Let us think about it:
We can make 8 observations (7 degrees of freedom). If we want a resolution of 4 there has to be no confounding between the main effects (with 1 letter) and the 2-way effects (with 2 letters). Naturally we can not have that $A \cong D$ or something similar because otherwise the resolution would be 2. Also we can not have that $AB \cong AC$ because then $B \cong C$. Consequently 3 different 2-way effects can not be confounded all together without having the undesirable consequence that two main effects are confounded.

Summarising: <u>If we want a resolution of 4:</u>
We have 5 main effects and at least 10/2=5 2-way effect which can NOT be confounded! This makes 10 in total. There are just 7 degrees of freedom (we can look at most at 7 different effects), therefore it is impossible to find a structure with resolution 4!

4. **a)** Find k and l for this $2^{k-l}-$ design. We have: $8 = 2^3 = 2^{k-l}$ observations. Furthermore we have $k = 4$ different factors and $l = 4 - 3 = 1$.

b) Determine the alias structure of this design. Let us call the effects of "Side-to-side", "Yarn type", "Pick density" and "Air pressure" A, B, C and D respectively. Then we have

A	B	C	D	Treatm.	Strength
$-$	$-$	$-$	$-$	(1)	24.50
$+$	$-$	$-$	$+$	ad	22.05
$-$	$+$	$-$	$+$	bd	24.52
$+$	$+$	$-$	$-$	ab	25.00
$-$	$-$	$+$	$+$	cd	25.68
$+$	$-$	$+$	$-$	ac	24.51
$-$	$+$	$+$	$-$	bc	24.68
$+$	$+$	$+$	$+$	abcd	24.23

with the alias $D = ABC$.
To find out which terms are aliased together it is enough to multiply the terms by $I = ABCD$. So

$$\begin{aligned} D &= ABC \\ A &= BCD \\ B &= ACD \\ C &= ABD \\ AB &= CD \\ AC &= BD \\ AD &= BC \end{aligned}$$

[1] The answer to this question is not required to solve the exercise and it is not trivial.

c) Calculate estimates of the effects. Estimates:

$$\begin{aligned}
\hat{A} &= \frac{1}{4}(-24.5 + 22.05 - 24.52 + 25 - 25.68 + 24.51 - 24.68 + 24.23) = -0.8975 \\
\hat{B} &= \frac{1}{4}(-24.5 - 22.05 + 24.52 + 25 - 25.68 - 24.51 + 24.68 + 24.23) = 0.4225 \\
\hat{C} &= \frac{1}{4}(-24.5 - 22.05 - 24.52 - 25 + 25.68 + 24.51 + 24.68 + 24.23) = 0.7575 \\
\hat{D} &= \frac{1}{4}(-24.5 + 22.05 + 24.52 - 25 + 25.68 - 24.51 - 24.68 + 24.23) = -0.5525 \\
\hat{AB} &= \hat{CD} = \frac{1}{4}(+24.5 - 22.05 - 24.52 + 25 + 25.68 - 24.51 - 24.68 + 24.23) = 0.9125 \\
\hat{AC} &= \hat{BD} = \frac{1}{4}(+24.5 - 22.05 + 24.52 - 25 - 25.68 + 24.51 - 24.68 + 24.23) = 0.0875 \\
\hat{AD} &= \hat{BC} = \frac{1}{4}(+24.5 + 22.05 - 24.52 - 25 - 25.68 - 24.51 + 24.68 + 24.23) = -1.0625
\end{aligned}$$

d) Suppose that additional experimentation shows that only effects whose magnitudes exceed 0.35 are important. Wich factors or interactions have a practically significant effect on fabric strength?

One way to test practical significance is to test if a confidence interval around our effect is bounded away from the critical value 0.35, in absolute value (so either all the way to the left or to the right). Let's fit the model with R and compute these CI:

```
> y <- c(24.5, 22.05, 24.52, 25, 25.68, 24.51, 24.68, 24.23)
> A <- as.factor(rep(c(1,2), 4))
> B <- as.factor(rep(c(1,1,2,2), 2))
> C <- as.factor(rep(c(1,2), each=4))
> D <- as.factor(c(1,2,2,1,2,1,1,2))
> dat <- data.frame(A=A,B=B,C=C,D=D,y)
> fit <- lm(y~., data=dat)
> anova(fit)
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value Pr(>F)
A          1 1.6110 1.61101  1.2271 0.3488
B          1 0.3570 0.35701  0.2719 0.6381
C          1 1.1476 1.14761  0.8742 0.4188
D          1 0.6105 0.61051  0.4650 0.5442
Residuals  3 3.9384 1.31281
> confint(fit)
                2.5 %    97.5 %
(Intercept) 21.648530 27.413970
A2          -3.475883  1.680883
B2          -2.155883  3.000883
C2          -1.820883  3.335883
D2          -3.130883  2.025883
```

From the ANOVA table we already see that no effect is significant. If you can't reject the null hypothesis that an effect is zero, there is no hope to not reject the null hypothesis that an effect is bigger than a value (0.35). You could just stop here, but to show how one would proceed, we look at the CI and see that indeed, either +0.35 or -0.35 is always included in the CI, therefore none of the effect is practically significant.

e) Suppose that additional experiments show that the AB and AD interactions are not significant. If the objective of the study is to maximize fabric strength, what setting of of each factor do you recommend? With the assumption and the calculations of point **??** we do not care about the two (or more)-way-effects.
The effect of A is $-0.8975 < 0$ which means that by changing the level of the factor A from low to high we lose strength. Consequently we choose:

Effect	Estimate	$\gtreqless 0$	Best
A	-0.8975	< 0	low
B	0.4225	>0	high
C	0.7575	> 0	high
D	-0.5525	< 0	low