EDDA - assignment1

Geoffrey

February 16, 2020

Exercise 5. (Chick weights)

a) The two groups are not paired. We first test normallity and their variance:

```
x = chickwts
mm = subset(x, feed=="meatmeal")["weight"]
sf = subset(x, feed=="sunflower")["weight"]
mm1 = as.numeric(unlist(mm))
sf1 = as.numeric(unlist(sf))

var.test(mm1, sf1)
```

```
##
##  F test to compare two variances
##
## data:  mm1 and sf1
## F = 1.7661, num df = 10, denom df = 11, p-value = 0.3645
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.5009206 6.4725366
## sample estimates:
## ratio of variances
##            1.766081
```

```
par(mfrow=c(1,2))
qqnorm(mm1)
qqnorm(sf1)
```

Normal Q–Q Plot	Normal Q–Q Plot
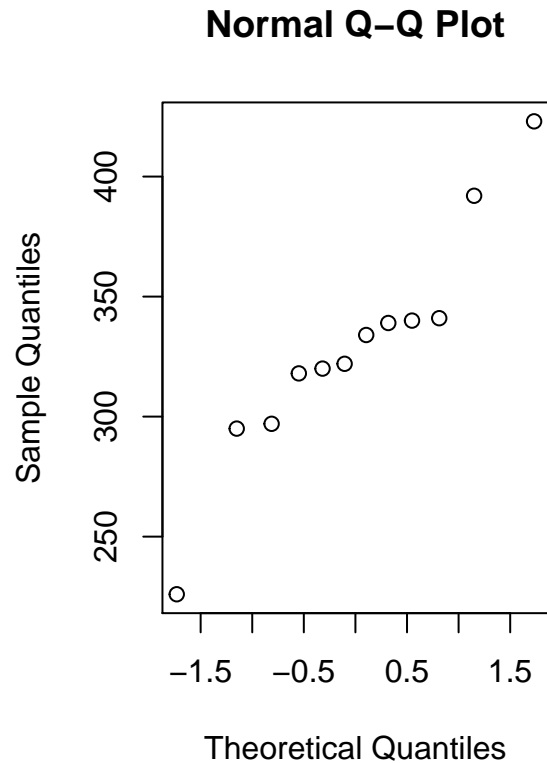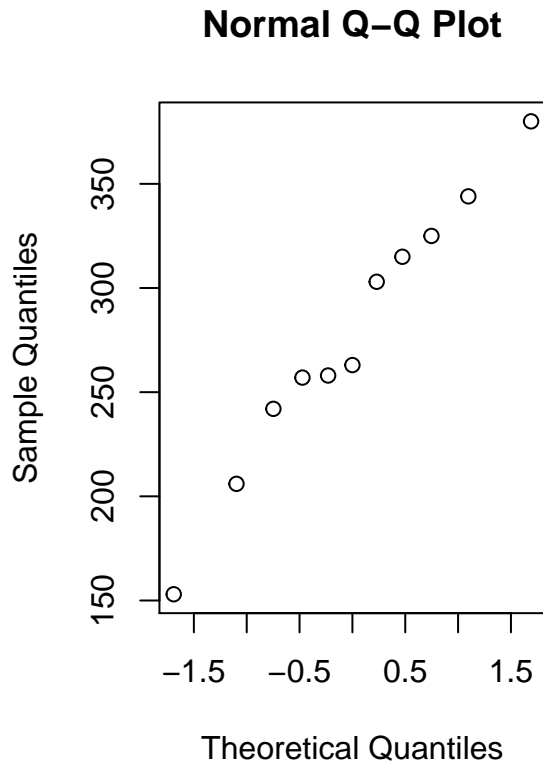
From the variance test we can conclude that their is no difference. And from the qqplots we can conclude that they are normal. From the following code we conclude that according to 1) the two sample t test there is a significant difference in their mean, 2) the mann whitney test there is not a sifnificant difference in distribution and 3) the kalmogorov-smirnov test there is not a difference in the distribution.

```
t.test(mm1, sf1, alternative = "two.sided", var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  mm1 and sf1
## t = -2.1838, df = 21, p-value = 0.04047
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -101.53409   -2.48106
## sample estimates:
## mean of x mean of y
##  276.9091  328.9167
```

```
wilcox.test(mm1, sf1, alternative = "two.sided", var.equal = TRUE)
```

```
##
##  Wilcoxon rank sum test
##
## data:  mm1 and sf1
## W = 36, p-value = 0.06882
## alternative hypothesis: true location shift is not equal to 0
```

```
ks.test(mm1, sf1)
```

```
##
##  Two-sample Kolmogorov-Smirnov test
```

```
##
## data:  mm1 and sf1
## D = 0.47727, p-value = 0.1085
## alternative hypothesis: two-sided
```

We conclude that there is a significant difference in the weight of the chicken groups. Because the sample is normaly distributed and therefor we use the regular t.test.

b) Using one way anova we can conclude that the means are not all equal:

```
res.aov = aov(weight ~ feed, data = x)
summary(res.aov)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## feed         5 231129   46226   15.37 5.94e-10 ***
## Residuals   65 195556    3009
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
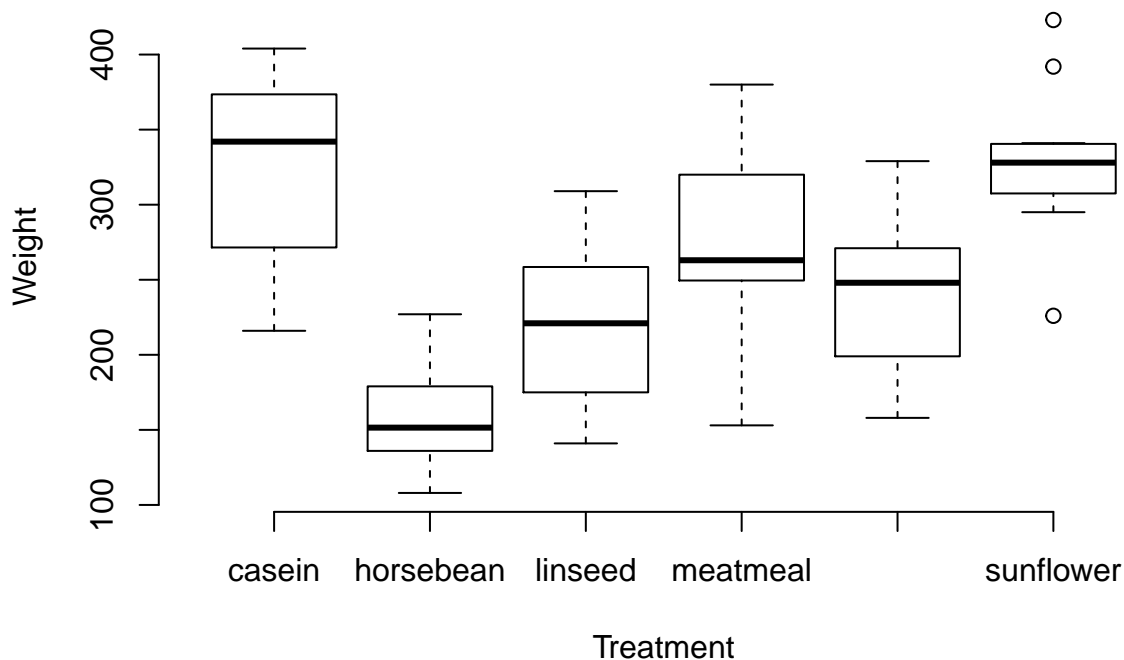
We can further conclude that sunflowers breat the heaviest chicks:

```
tapply(x$weight, x$feed, mean)
```

```
##    casein horsebean   linseed  meatmeal   soybean sunflower
##  323.5833  160.2000  218.7500  276.9091  246.4286  328.9167
```

c)By visualising in the boxplots we can see small differences in the variances.

```
boxplot(weight ~ feed, data = x,
        xlab = "Treatment", ylab = "Weight",
        frame = FALSE)
```


We assume that all the groups are normally distributed. However we do check whether their variances are homogenious.

We first visualise the residuals by the fitted values, which suggests that they are homogenious. And then we do it formally with the bartlett test. Which results in a p value 0.66, after which we can conclude that there is no significant difference in the variance.

```
plot(res.aov, 1)
```

Residuals vs Fitted
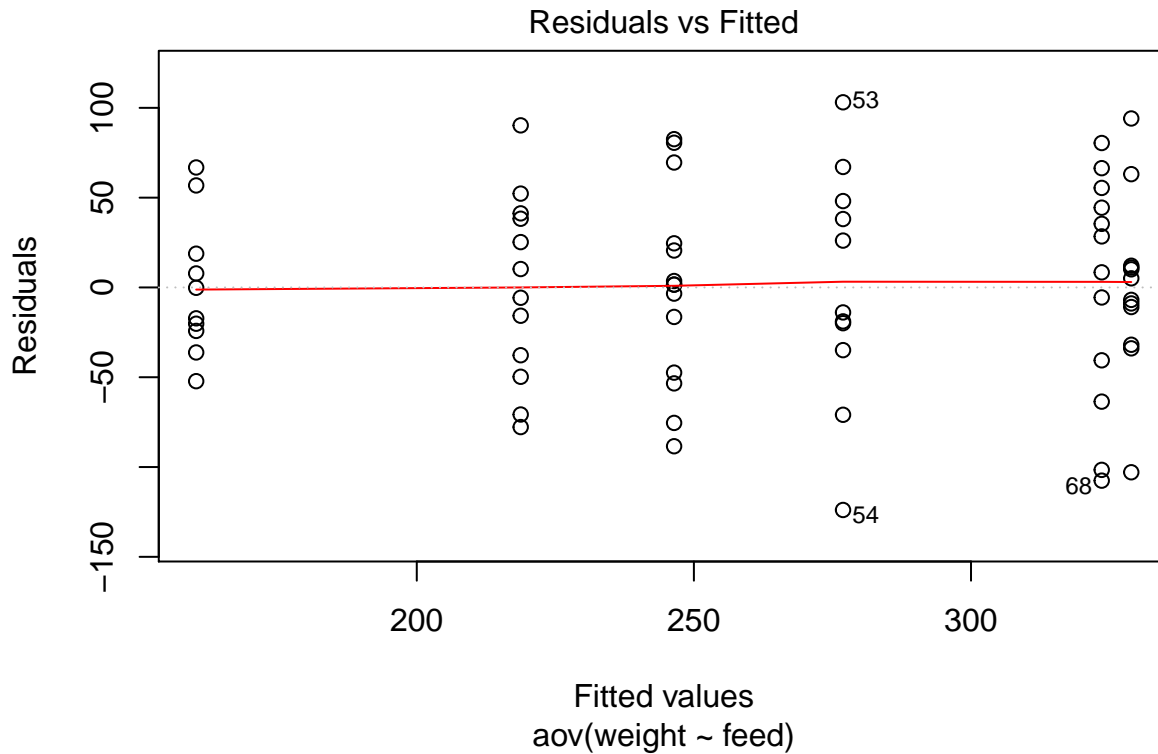

```
bartlett.test(weight ~ feed, data = x)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  weight by feed
## Bartlett's K-squared = 3.2597, df = 5, p-value = 0.66
```

d) Kruskal wallis test also comes up with a very small p number, which is also smaller than 0.05. Therefor it has the same conclusion as the one way anova test: the means are not the same and thus the treatment matters.

```
kruskal.test(weight ~ feed, data = x)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  weight by feed
## Kruskal-Wallis chi-squared = 37.343, df = 5, p-value = 5.113e-07
```

4

R Markdown <— example stuff

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

Including Plots

You can also embed plots, for example:
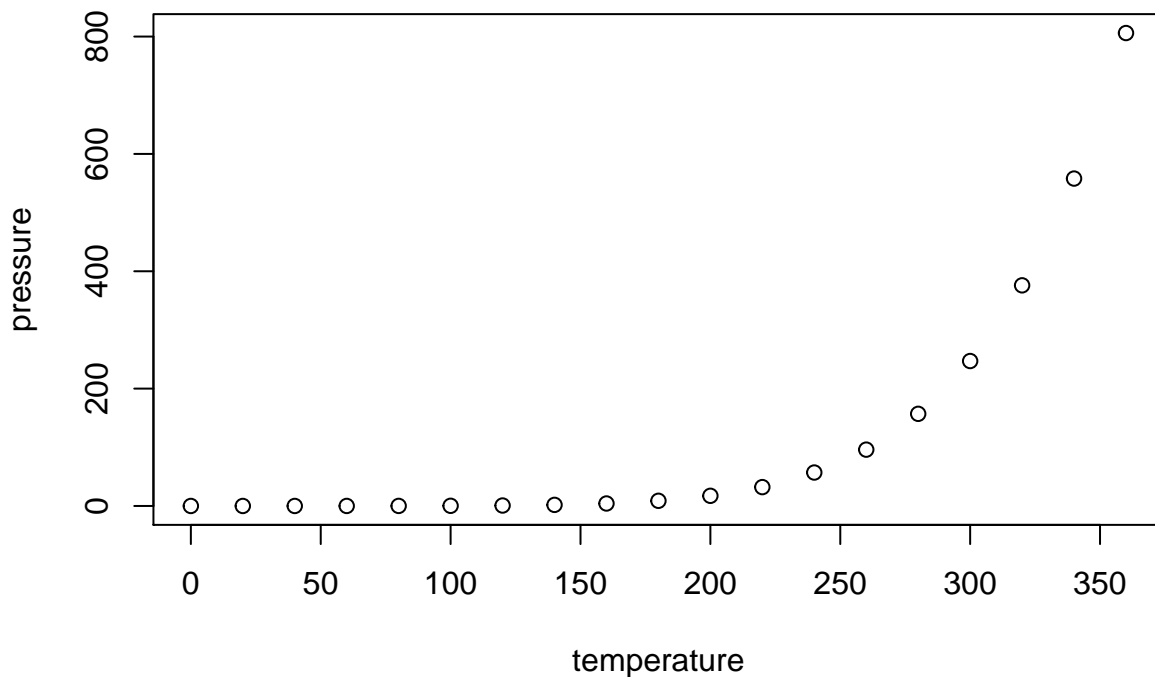
Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.