Analysis of Variance (ANOVA)

Chapter 5, Lab 2: Solutions

OpenIntro Biostatistics

Topics

- ANOVA *F*-test

- Adjustment for multiple comparisons

The previous lab introduced the two-group independent *t*-test as a method for comparing the means of two groups. In some settings, it is useful to compare the means across more than two groups. The methodology behind a two-group independent *t*-test can be generalized to a procedure called **analysis of variance (ANOVA)**. Assessing whether the means across several groups are equal by conducting a single hypothesis test rather than multiple two-sample tests is important for controlling the overall Type I error rate.

The material in this lab corresponds to Section 5.5 of *OpenIntro Biostatistics*.

FAMuSS: comparing change in non-dominant arm strength by *ACTN3* genotype

> *Is change in non-dominant arm strength after resistance training associated with genotype?*

In the Functional polymorphisms Associated with Human Muscle Size and Strength study (FAMuSS), researchers examined the relationship between muscle strength and genotype at a particular location on the *ACTN3* gene. The famuss dataset in the oibiostat package contains a subset of data from the study.

The percent change in non-dominant arm strength, comparing strength after resistance training to before training, is stored as ndrm.ch. There are three possible genotypes (CC, CT, TT) at the *r577x* position on the *ACTN3* gene; genotype is stored as actn3.r557x.

1. Load the data. Create a plot that shows the association between change in non-dominant arm strength and *ACTN3* genotype. Describe what you see.

   Change in non-dominant arm strength seems positively associated with presence of the T allele; TT individuals have higher median ndrm.ch than CT and CC individuals, and CT individuals have higher median ndrm.ch than CC individuals.

   ```
   #load the data
   library(oibiostat)
   data("famuss")

   #create plot
   plot(famuss$ndrm.ch ~ famuss$actn3.r577x)
   ```
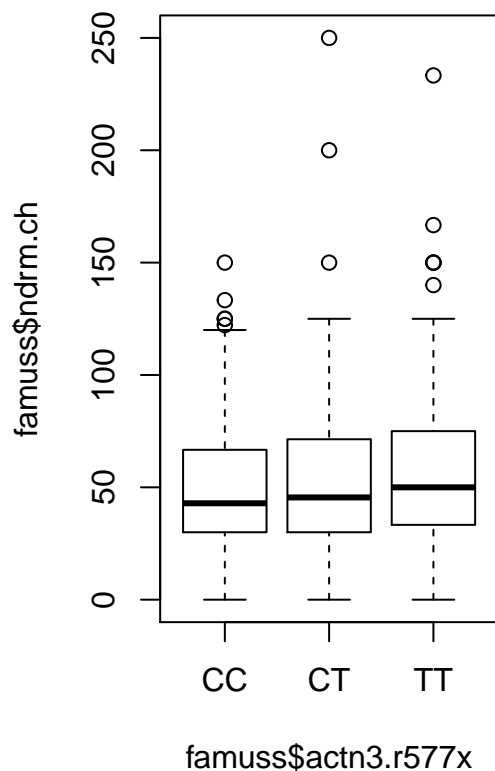
2. Assess whether the assumptions for conducting an ANOVA are reasonably satisfied: 1) observations are independent within and across groups, 2) the data within each group are nearly normal, and 3) the variability across the groups is about equal.

It is reasonable to assume that the observations are independent within and across groups; it is unlikely that participants in the study were related, or that data collection was carried out in a way that one participant's change in arm strength could influence another's. Based on the Q-Q plots, there is some evidence of skew and presence of outliers, but the data are approximately normal. The variance between groups is roughly constant.

```
#check assumptions
par(mfrow = c(1, 3))
qqnorm(famuss$ndrm.ch[famuss$actn3.r577x == "CC"], main = "Q-Q for CC genotype")
qqline(famuss$ndrm.ch[famuss$actn3.r577x == "CC"])
qqnorm(famuss$ndrm.ch[famuss$actn3.r577x == "CT"], main = "Q-Q for CT genotype")
qqline(famuss$ndrm.ch[famuss$actn3.r577x == "CT"])
qqnorm(famuss$ndrm.ch[famuss$actn3.r577x == "TT"], main = "Q-Q for TT genotype")
qqline(famuss$ndrm.ch[famuss$actn3.r577x == "TT"])
```

Q–Q for CC genotype   **Q–Q for CT genotype**   **Q–Q for TT genotype**
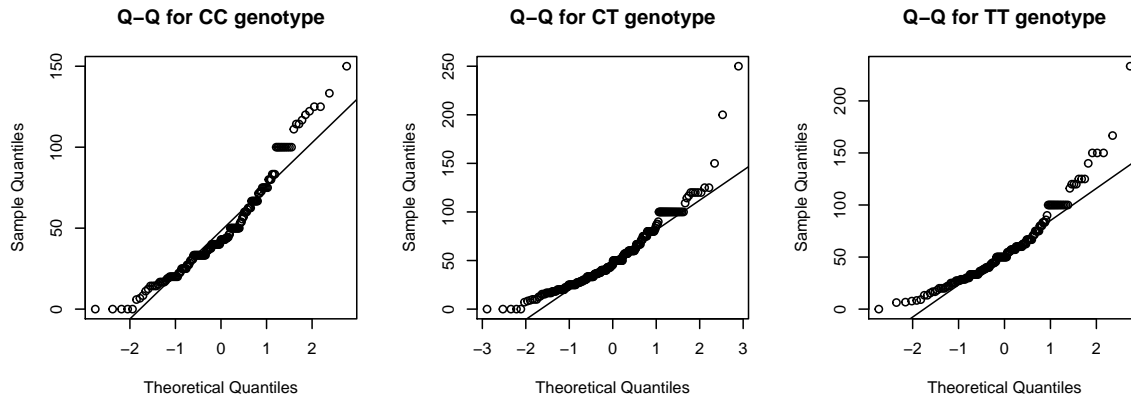
```
tapply(famuss$ndrm.ch, famuss$actn3.r577x, var)
```

```
##        CC        CT        TT
##  897.3667 1104.3311 1273.8712
```

3. Conduct a hypothesis test to address the question of interest. Let $\alpha = 0.05$.

   a) Let the parameters μ_{CC}, μ_{CT}, and μ_{TT} represent the population mean change in non-dominant arm strength for individuals of the corresponding genotype. State the null and alternative hypotheses.

   The null hypothesis is that mean percent change in non-dominant arm strength is equal across the three genotypes, $H_0 : \mu_{CC} = \mu_{CT} = \mu_{TT}$. The alternative hypothesis is that at least one group has mean percent change in non-dominant arm strength that is different from the other groups.

   b) Use `summary(aov())` to compute the F-statistic and p-value. Interpret the p-value.

   The F-statistic is 3.231, and the p-value is 0.0402. The p-value is less than 0.05; there is sufficient evidence to reject the null in favor of the alternative, and conclude that there is a difference in population mean change in non-dominant arm change between the groups.

```
#use summary(aov())
summary(aov(famuss$ndrm.ch ~ famuss$actn3.r577x))
```

```
##                     Df Sum Sq Mean Sq F value Pr(>F)
## famuss$actn3.r577x   2   7043    3522   3.231 0.0402 *
## Residuals          592 645293    1090
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

   c) Complete the analysis using pairwise comparisons.

   i. What is the appropriate significance level $\alpha^\star$ for the individual comparisons, as per the Bonferroni correction?

   $$\alpha^\star = \alpha/K, \text{where } K = \frac{k(k-1)}{2} \text{for } k \text{ groups}$$

   With $k = 3$ comparisons, $K = 3$. Thus, $\alpha^\star = 0.05/3 = 0.0167$.

3

```
#use R as a calculator
alpha = 0.05
k = 3

K = (k*(k-1))/2

alpha.star = alpha/k
alpha.star
```

ii. Use `pairwise.t.test()` to conduct the pairwise two-sample t-tests.

```
#use pairwise.t.test()
pairwise.t.test(famuss$ndrm.ch, famuss$actn3.r577x, p.adj = "none")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  famuss$ndrm.ch and famuss$actn3.r577x
##
##    CC    CT
## CT 0.179 -
## TT 0.011 0.144
##
## P value adjustment method: none
```

```
#alternatively, use pairwise.t.test() with the bonferroni correction
pairwise.t.test(famuss$ndrm.ch, famuss$actn3.r577x, p.adj = "bonf")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  famuss$ndrm.ch and famuss$actn3.r577x
##
##    CC    CT
## CT 0.537 -
## TT 0.034 0.433
##
## P value adjustment method: bonferroni
```

iii. Summarize the results.

A conclusion can be made by either comparing the unadjusted p-values to $\alpha^\star$, or comparing the adjusted p-values to α. When the Bonferroni correction is applied in `pairwise.t.test()`, R multiplies the p-value by K. Comparing a p-value to α/K is equivalent to comparing the quantity $(K \times p\text{-value})$ to α.

Only the CC versus TT comparison results in a p-value less than $\alpha^\star = 0.0167$. There is evidence that mean strength change for individuals of genotype CC and TT are different, and that individuals of TT genotype have higher mean strength change in the non-dominant arm than CC individuals. The other two comparisons did not result in p-values less than $\alpha^\star$;

NHANES: comparing BMI by educational level

Is body mass index (BMI) associated with educational attainment?

This section uses data from the National Health and Nutrition Examination Survey (NHANES), a survey conducted annually by the US Centers for Disease Control (CDC).[1] The dataset nhanes.samp.adult.500 contains data for 500 participants ages 21 years or older that were randomly sampled from the complete NHANES dataset that contains 10,000 observations.

The variable BMI contains BMI information for the study participants. The variable Education records the highest level of education obtained: 8th grade, 9th - 11th grade, high school, some college, or college degree.

4. Load the data. Create a plot that shows the association between BMI and educational level. Describe what you see.

   Median BMI is roughly similar across education categories, although median BMI looks lower in the 8th grade and college graduate groups.

```
#load the data
library(oibiostat)
data("nhanes.samp.adult.500")

#create a plot
boxplot(nhanes.samp.adult.500$BMI ~ nhanes.samp.adult.500$Education)
```


_____
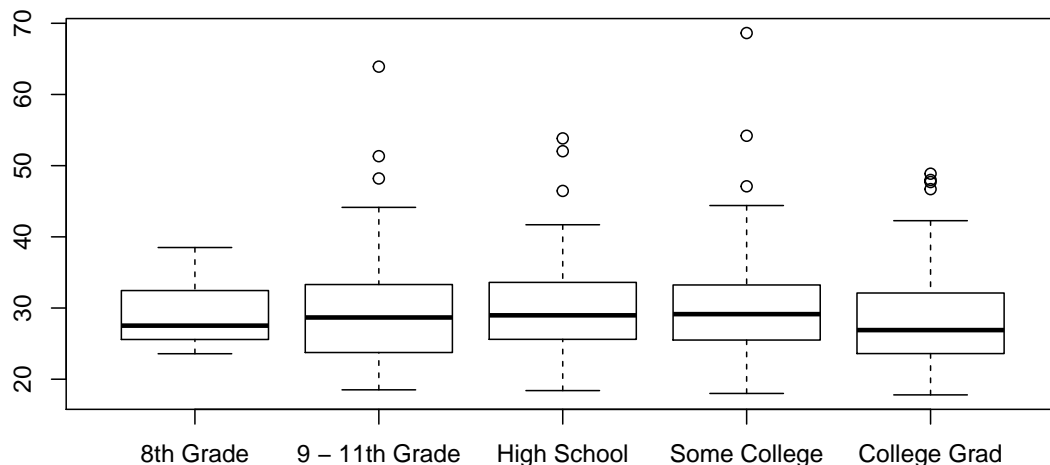
[1] The dataset was first introduced in Chapter 1, Lab 1 (Introduction to Data).

5. Examine the normality and equal variance assumptions across the groups. Explain why it is advisable to restrict the analysis to participants who have completed at least 9th grade.

In this sample, there are not many participants in the 8th grade category; this affects both the normality and equal variance assumptions. The variance in the 8th grade group is small relative to the other groups, and the center of the distribution shows departures from normality. Across the other groups, normality and equal variance are reasonably satisfied. One rule of thumb for assessing equal variance in the ANOVA context is to check the ratio of the largest variance to the smallest variance; a ratio less than 3 is considered acceptable for conducting the ANOVA F-test.

```
#check assumptions
table(nhanes.samp.adult.500$Education)
```

```
##
##      8th Grade 9 - 11th Grade    High School   Some College    College Grad
##             32             68             94            148             157
```

```
tapply(nhanes.samp.adult.500$BMI, nhanes.samp.adult.500$Education, var, na.rm = TRUE)
```

```
##      8th Grade 9 - 11th Grade    High School   Some College    College Grad
##       16.66860       72.91831       44.00945       46.71782       42.17102
```

```
par(mfrow = c(3, 2))
qqnorm(nhanes.samp.adult.500$BMI[nhanes.samp.adult.500$Education == "8th Grade"],
    main = "Q-Q for 8th Grade")
qqline(nhanes.samp.adult.500$BMI[nhanes.samp.adult.500$Education == "8th Grade"])

qqnorm(nhanes.samp.adult.500$BMI[nhanes.samp.adult.500$Education == "9 - 11th Grade"],
    main = "Q-Q for 9 - 11th Grades")
qqline(nhanes.samp.adult.500$BMI[nhanes.samp.adult.500$Education == "9 - 11th Grade"])

qqnorm(nhanes.samp.adult.500$BMI[nhanes.samp.adult.500$Education == "High School"],
    main = "Q-Q for High School")
qqline(nhanes.samp.adult.500$BMI[nhanes.samp.adult.500$Education == "High School"])

qqnorm(nhanes.samp.adult.500$BMI[nhanes.samp.adult.500$Education == "Some College"],
    main = "Q-Q for Some College")
qqline(nhanes.samp.adult.500$BMI[nhanes.samp.adult.500$Education == "Some College"])

qqnorm(nhanes.samp.adult.500$BMI[nhanes.samp.adult.500$Education == "College Grad"],
    main = "Q-Q for College Grad")
qqline(nhanes.samp.adult.500$BMI[nhanes.samp.adult.500$Education == "College Grad"])
```
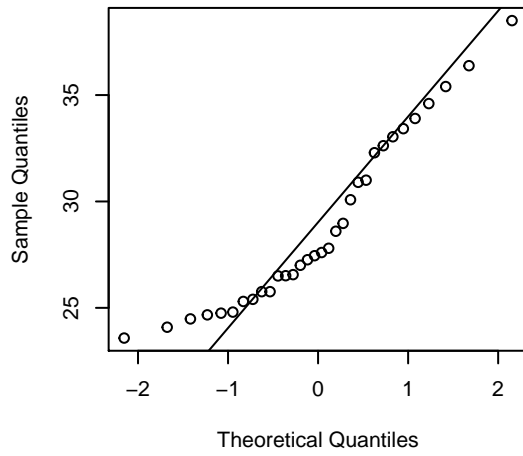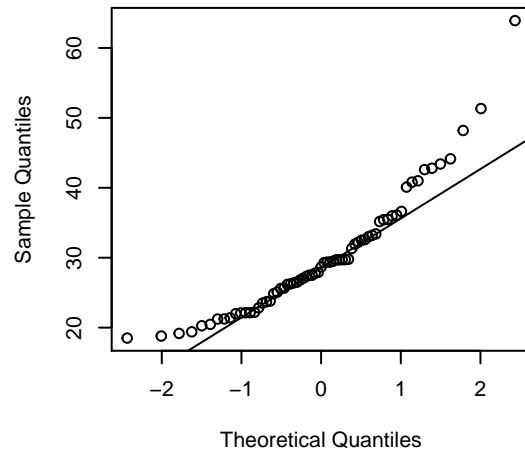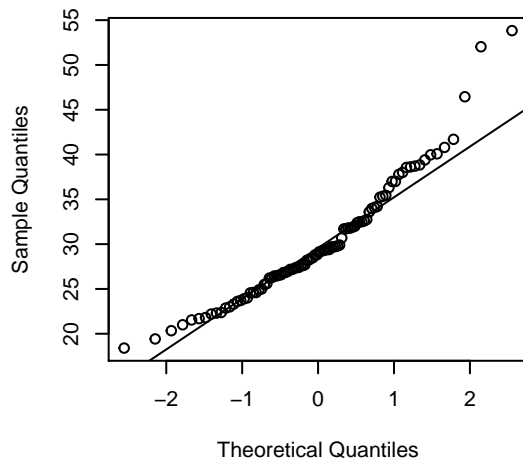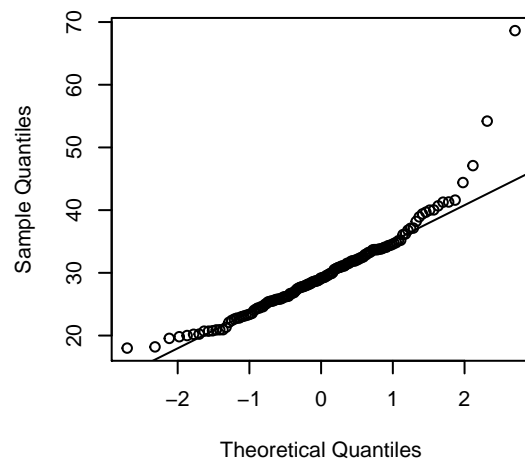
Q–Q for 8th Grade
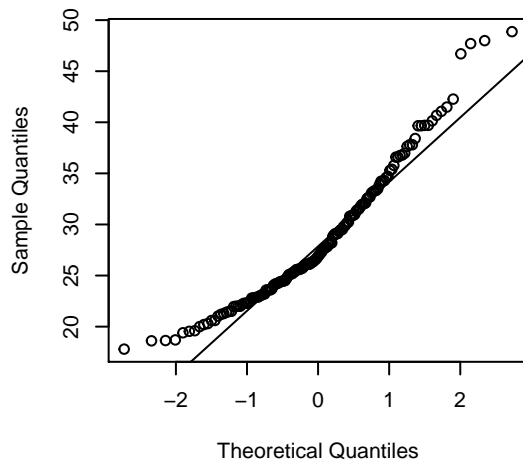

Q–Q for 9 – 11th Grades

Q–Q for High School

Q–Q for Some College

Q–Q for College Grad

6. Conduct a hypothesis test to address the question of interest. Let $\alpha = 0.05$. Summarize the conclusions.

The p-value is 0.30, which is greater than $\alpha = 0.05$. There is not evidence to suggest that BMI is associated with educational level for individuals who have completed at least 9th grade. The observed discrepancy in the mean BMI across groups is not implausible if the population means are equal (i.e., $\mu_{9-11} = \mu_{HS} = \mu_{SC} = \mu_{CG}$).

Note that since the F-test did not reject the null hypothesis, the analysis does not proceed to pairwise testing.

```
#conduct hypothesis test
summary(aov(nhanes.samp.adult.500$BMI ~ nhanes.samp.adult.500$Education))

##                                 Df Sum Sq Mean Sq F value Pr(>F)
## nhanes.samp.adult.500$Education   4    226   56.40   1.216  0.303
## Residuals                       491  22775   46.38
## 4 observations deleted due to missingness
```

Chick weights: comparing weight across feed supplements

Chicken farming is a multi-billion dollar industry, and any methods that increase the growth rate of young chicks can reduce consumer costs while increasing company profits. An experiment was conducted to measure and compare the effectiveness of various feed supplements on the growth rate of chicks. Newly hatched chicks were randomly allocated into groups, and each group was given a different feed supplement.

The chickwts dataset available in the datasets package contains the weight in grams of chicks at six weeks of age. For simplicity, this analysis will be limited to four types of feed supplements: linseed, meatmeal, soybean, and sunflower.

7. Run the following code to load the chickwts dataset and subset the data for the four feed supplements of interest.

```
#load the data
library(datasets)
data("chickwts")

#subset the four feed supplements
keep = (chickwts$feed == "linseed" | chickwts$feed == "meatmeal" |
  chickwts$feed == "soybean" | chickwts$feed == "sunflower")
chickwts = chickwts[keep, ]

#eliminate unused levels
chickwts$feed <- droplevels(chickwts$feed)
```

8. Analyze the data and report the results. Using language accessible to a non-statistician, discuss which feed supplement(s) is/are the most effective for increasing chick weight.

The equal variance assumption is reasonably satisfied. Based on the design of the experiment, it is reasonable to assume that the observations are independent within and between

groups. Since sample size is small, it is difficult to assess normality; however, physical measures like weight and height are usually normally distributed.

The F-test p-value is highly significant; there is sufficient evidence to reject the null that all group means are the same. Two of the pairwise tests result in signficant p-values: sunflower versus linseed and sunflower versus soybean.

Based on these results, it is possible to recommend sunflower over either linseed or soybean; chicks fed on sunflower are about 80-100 g heavier at six weeks of age than those fed on linseed or soybean. There is less clear evidence in favor of sunflower over meatmeal, as the 50 g difference in average chick weight between the sunflower and meatmeal groups observed in this study may simply be due to random fluctuation.

Note: A more detailed recommendation to a chicken farmer could discuss additional points, such as whether sunflower or meatmeal is more cost-effective as a feed supplement. A follow-up experiment might be conducted with a larger sample size and designed to specifically assess sunflower and meatmeal.

```r
#summary statistics
table(chickwts$feed)
```

```
##
##   linseed  meatmeal   soybean sunflower
##        12        11        14        12
```

```r
tapply(chickwts$weight, chickwts$feed, var)
```

```
##   linseed  meatmeal   soybean sunflower
##  2728.568  4212.091  2929.956  2384.992
```

```r
tapply(chickwts$weight, chickwts$feed, mean)
```

```
##   linseed  meatmeal   soybean sunflower
##  218.7500  276.9091  246.4286  328.9167
```

```r
#check normality
par(mfrow = c(2, 2))
qqnorm(chickwts$weight[chickwts$feed == "linseed"],
   main = "Q-Q for linseed")
qqline(chickwts$weight[chickwts$feed == "linseed"])
qqnorm(chickwts$weight[chickwts$feed == "meatmeal"],
   main = "Q-Q for meatmeal")
qqline(chickwts$weight[chickwts$feed == "meatmeal"])
qqnorm(chickwts$weight[chickwts$feed == "soybean"],
   main = "Q-Q for soybean")
qqline(chickwts$weight[chickwts$feed == "soybean"])
qqnorm(chickwts$weight[chickwts$feed == "sunflower"],
   main = "Q-Q for sunflower")
qqline(chickwts$weight[chickwts$feed == "sunflower"])
```

Q–Q for linseed

Q–Q for meatmeal

Q–Q for soybean

Q–Q for sunflower
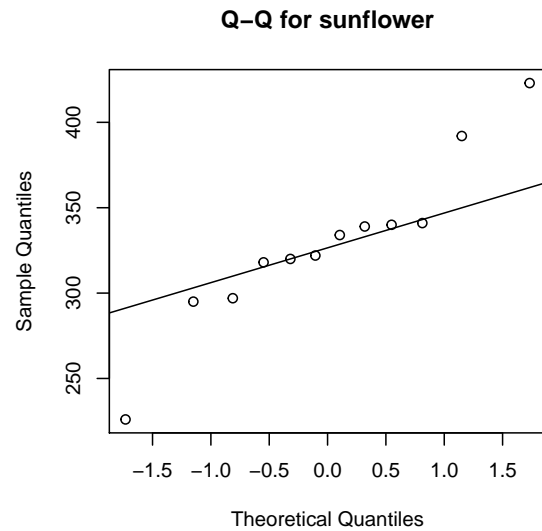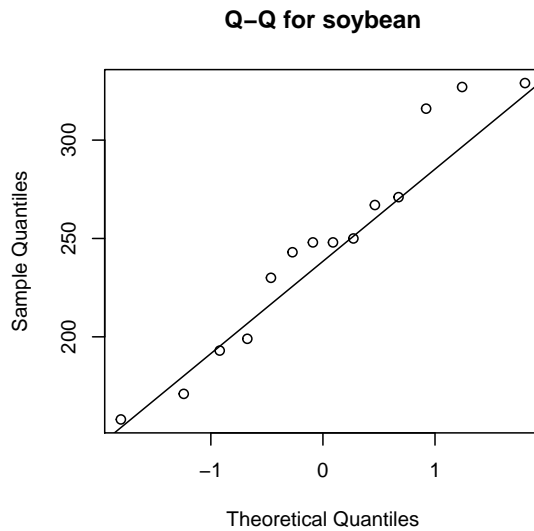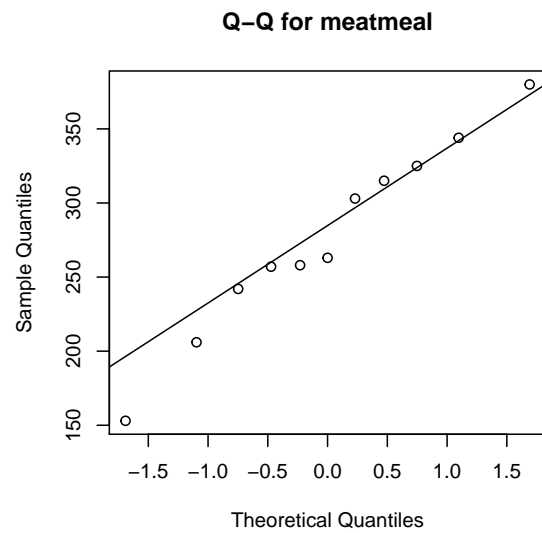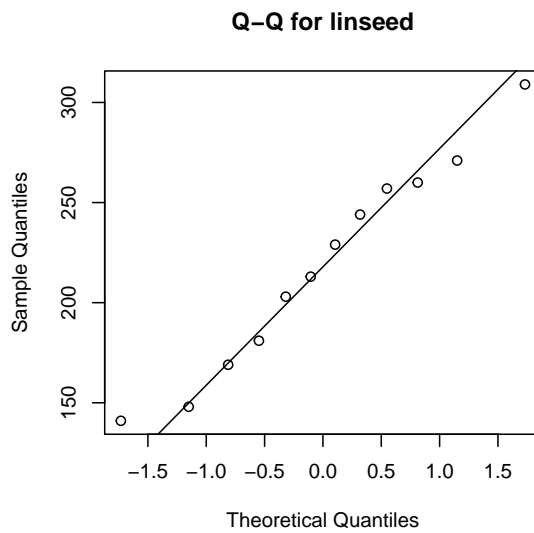

```r
#conduct analysis
summary(aov(chickwts$weight ~ chickwts$feed))
```

```
##               Df Sum Sq Mean Sq F value   Pr(>F)
## chickwts$feed  3  80941   26980   8.897 9.66e-05 ***
## Residuals     45 136460    3032
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
pairwise.t.test(chickwts$weight, chickwts$feed, p.adj = "bonf")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  chickwts$weight and chickwts$feed
```

```
##
##           linseed meatmeal soybean
## meatmeal  0.0898  -        -
## soybean   1.0000  1.0000   -
## sunflower 7.7e-05 0.1712   0.0025
##
## P value adjustment method: bonferroni
```