

Chapter 8 and 9 Homework

Due 3-18-2018

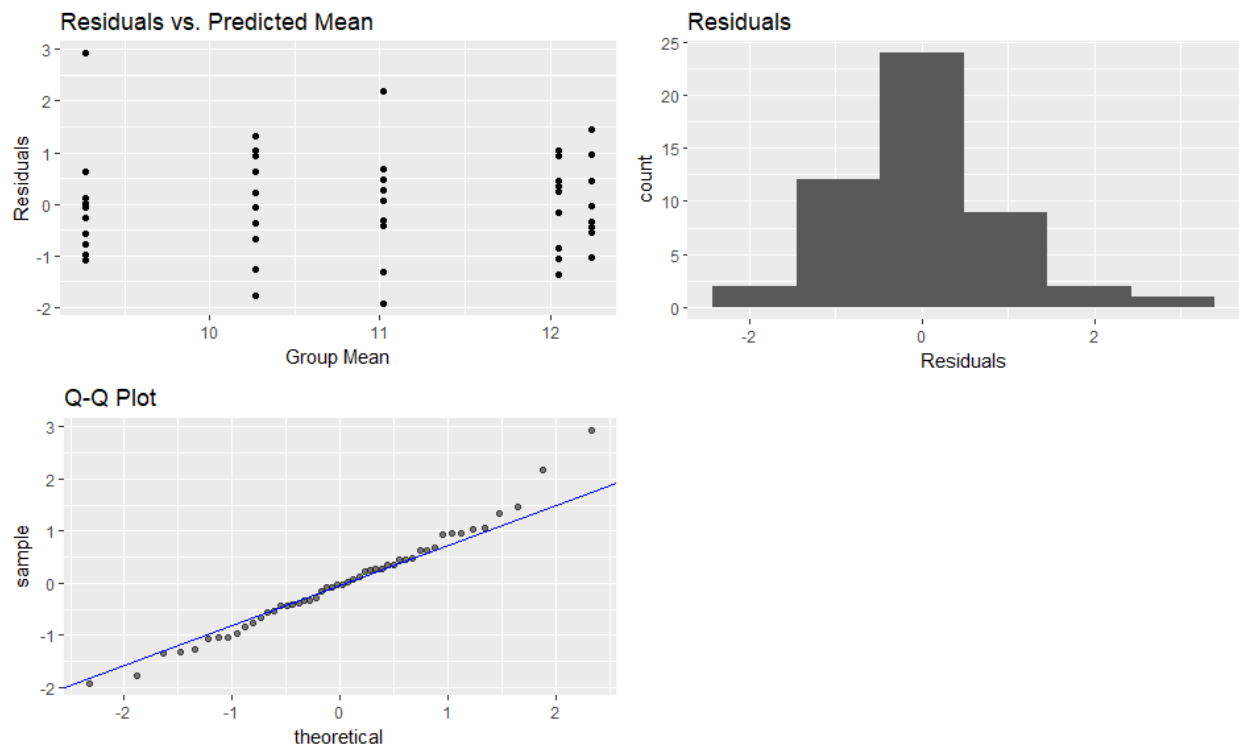
Problem 9.13

a) Assess ANOVA assumptions using the graph from PROC MIXED.

In order to proceed with ANOVA, we must check for the following pieces:

- No obvious pattern in our Residual Scatterplot
- Normal shape to the Residual Histogram.
- Residuals fit along the linear prediction of our Actual "Normality" to the Theoretical Normal Model.

Since there is no readily observable pattern in our residuals, our distribution is mound shaped, and our Residuals fit the along the Q-Q plot(although, there may be evidence of a right skew), we will proceed with ANOVA to see if at least one mean is different.



b) Perform ANOVA to determine if there is a difference among the five weight-reducing agents, $\alpha = 0.05$.

Table 1: ANOVA Table for Weight Loss Study

Row Names	SS	df	MS	F	P-Value
Treatments (T)	61.618	4	15.4045	15.6805	4.16×10^{-8}
Error (E)	44.207	45	0.9824		
Total	105.825	49			

Hypotheses:

$$H_0: \mu_{a1} = \mu_{a2} = \mu_{a3} = \mu_{a4} = \mu_s$$

$$H_1: \text{At least one mean is different.}$$

Test Statistic:

$$F = 15.6805$$

Rejection Region:

$$\text{Reject } H_0 \text{ if } F_0 > F_{\alpha, 4, 45}$$

$$F_{0.95, 4, 45} = 5.72$$

Conclusion/Interpretation:

Since our $F_0 > 5.72$, there is strong enough evidence to support rejecting the null hypothesis that the means are the same. The data provided does suggest at least one mean is different from the rest.

c) Determine significantly different pairs using Tukey's W with $\alpha = 0.05$

To find which groups are different from each other, we will utilize Tukey's W to check which means are different from one another. We will verify this by checking which one's p-values are less than 0.05 after using the TukeyHSD test on our ANOVA Model in R.

Conclusion/Interpretation

With our Tukey W test being run, the differences lie in the following means:

- S differs with A_1 , A_2 , and A_4
- A_3 differs with A_1 and A_4

d) Determine which, if any, of the new agents have significantly larger mean weight loss as compared to the standard agent; $\alpha = 0.05$

Using s as our control, we will perform Dunnett's test to determine if there's a significantly larger difference with the new agents. The first step in checking out if any of the new agents are yielding better weight loss results is to find Dunnett's D.

$$D = d_{\alpha}(k, v) \sqrt{\frac{2s_W^2}{n}}$$

$$D = d_{0.05}(4, 45) \sqrt{\frac{2(0.982)}{10}}$$

$$D = 2.23 \sqrt{\frac{1.964}{10}}$$

$$D = 2.23(0.4432)$$

$$D = 0.9888$$

Table 2: Dunnett's Control Test

Treatment	$\bar{y}_i - \bar{y}_c$	Comparison	Conclusion
a1	2.78	> D	Greater Than Control
a2	1.75	> D	Greater Than Control
a3	1.00	> D	Greater Than Control
a4	2.97	> D	Greater Than Control

Conclusion/Interpretation:

All four of the agents are statistically significant. Thus, we can reject the null hypotheses that the difference in weight loss metrics between the original agent(s) and the new agents(a_1, a_2, a_3 , and a_4) is zero. Each of the new agents appear to have an increase in weight loss over the agent s.

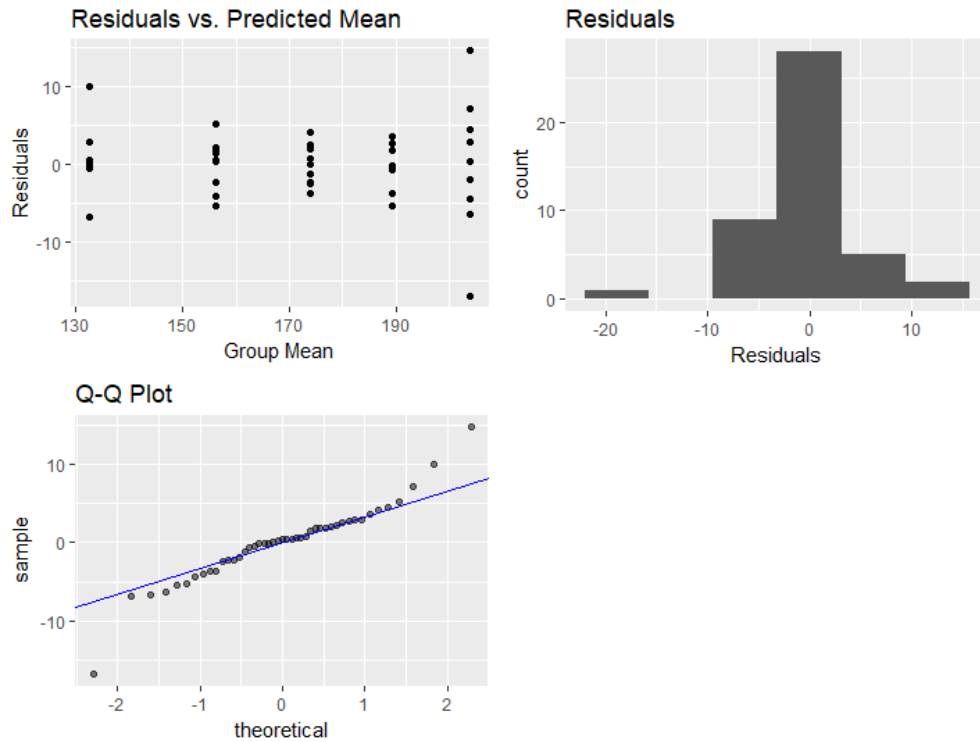
9.17 - construct the contrasts

a) Assess ANOVA assumptions using the graph from PROC MIXED

In order to proceed with ANOVA, we must check for the following pieces:

- No obvious pattern in our Residual Scatterplot
- Normal shape to the Residual Histogram.
- Residuals fit along the linear prediction of our Actual "Normality" to the Theoretical Normal Model.

There appears to be a pattern in the residuals. The best way to describe the pattern is a increasing fan shape at the edges. There is an increase in the residuals as you are on the extremes of the means of each group. Finally, a pattern in the Q-Q plot has this sample data straying from the Normal model by mismatching the Normal model outside of the -1 and + 1.5 standard deviations.



b) Run an ANOVA on the data to see if there is a difference in the suppliers

Table 3: ANOVA Table for Supplier's Variability in Power

Row Names	SS	df	MS	F	P-Value
Suppliers (S)	28024	4	7006	865.9	$< 2 \times 10^{-16}$
Error (E)	1054	40	26		
Total	29078	44			

Hypotheses:

$$H_0 : \mu_{S1} = \mu_{S2} = \mu_{S3} = \mu_{S4} = \mu_{S5}$$

H_1 : At least one mean is different.

Test Statistic:

$$F = 865.9$$

Rejection Region:

$$\text{Reject } H_0 \text{ if } F_0 > F_{\alpha, 4, 20} \quad F_{0.05, 4, 25} = 5.80$$

Conclusion/Interpretation:

With a test statistic greater than the rejection bound, there does seem to extreme enough evidence to reject the null hypothesis that the means are equal. It does appear that at least one mean is different.

c) Perform a Kruskal-Wallis test to determine if there is a difference among the suppliers; $\alpha = 0.05$

Hypothesis:

H_0 : There is no difference among the five suppliers with respect to median Variability in Power.

H_1 : At least one of the five suppliers' medians differs from the others.

Test Statistic:

$H = 41.596$ with 4 degrees of freedom and a p-value of 2.023×10^{-8}

Rejection Region:

We will reject if $H > \chi^2_{0.05,4}$
 $\chi^2_{0.05,4} = 9.488$

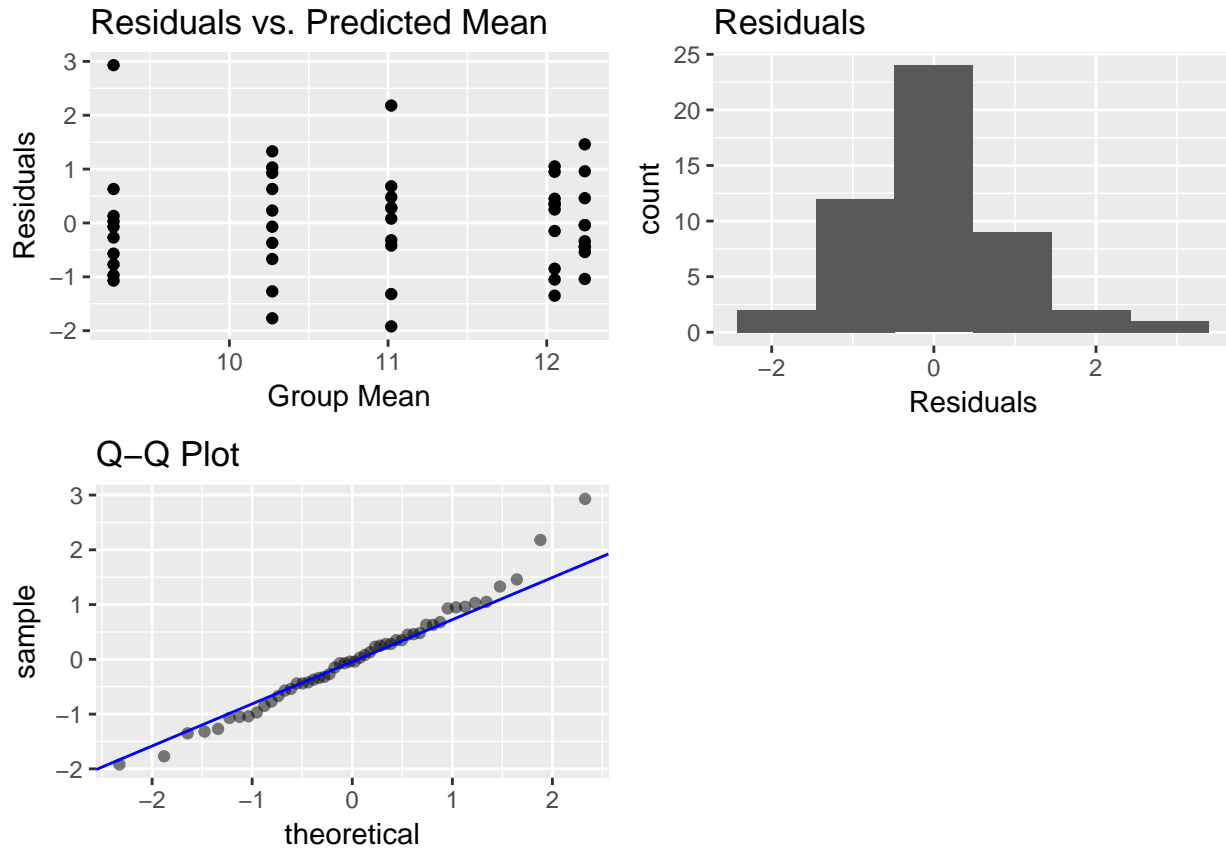
Conclusion/Interpretation

Since our Test Statistic is greater than 9.488, we have significant evidence to reject the Null Hypothesis that the all the medians are the same. There does appear to be at least one median that is different than the rest.

Chapter 8 And 9

Kyle Ligon

9.13 a) Checking the results from Proc Mixed in order to do ANOVA



9.13 b) Perform ANOVA test on the data: Show ANOVA Table First, then Run the Test

```
anova_mod
```

```
## Call:
##   aov(formula = wt_loss ~ treatment, data = gather_frame)
##
## Terms:
##               treatment Residuals
## Sum of Squares    61.618    44.207
## Deg. of Freedom      4      45
##
## Residual standard error: 0.9911497
## Estimated effects may be unbalanced
```

```
summary(anova_mod)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## treatment   4  61.62  15.404    15.68 4.16e-08 ***
## Residuals  45  44.21   0.982
## ---
```

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

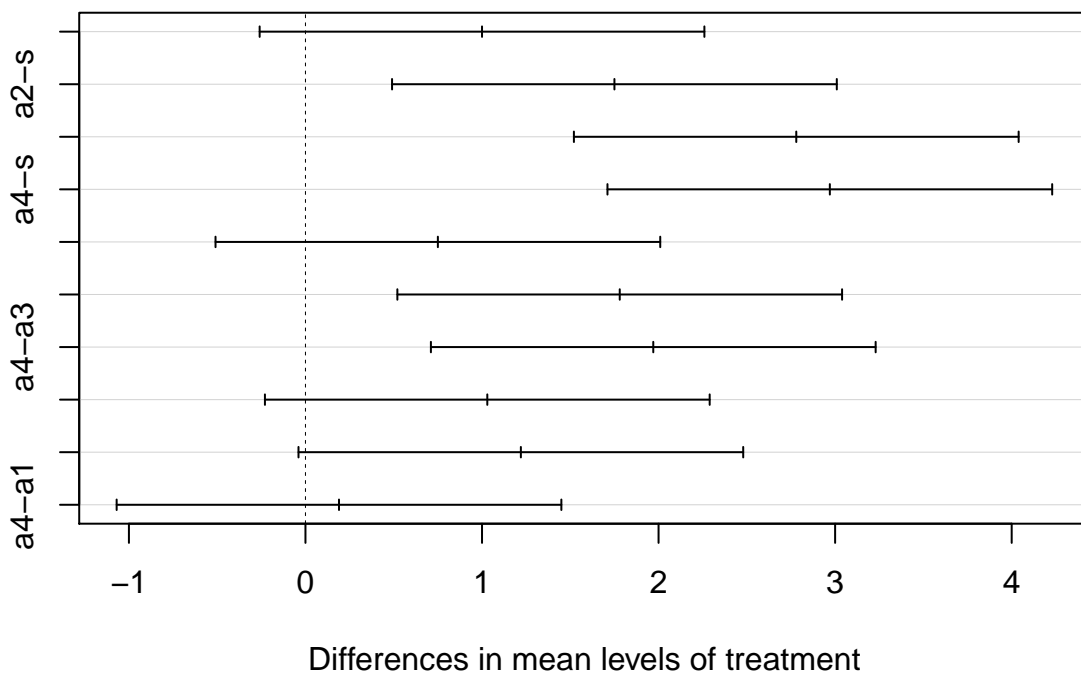
9.13 c) Perform Tukey's W on the significant pairs

```
real_w <- TukeyHSD(anova_mod, ordered = TRUE)
real_w$treatment
```

```
##      diff      lwr      upr      p adj
## a3-s  1.00 -0.2594887 2.259489 1.784060e-01
## a2-s  1.75  0.4905113 3.009489 2.428628e-03
## a1-s  2.78  1.5205113 4.039489 1.200843e-06
## a4-s  2.97  1.7105113 4.229489 2.780828e-07
## a2-a3 0.75 -0.5094887 2.009489 4.490082e-01
## a1-a3 1.78  0.5205113 3.039489 1.980323e-03
## a4-a3 1.97  0.7105113 3.229489 5.243121e-04
## a1-a2 1.03 -0.2294887 2.289489 1.563263e-01
## a4-a2 1.22 -0.0394887 2.479489 6.176067e-02
## a4-a1 0.19 -1.0694887 1.449489 9.927171e-01
```

```
plot(real_w)
```

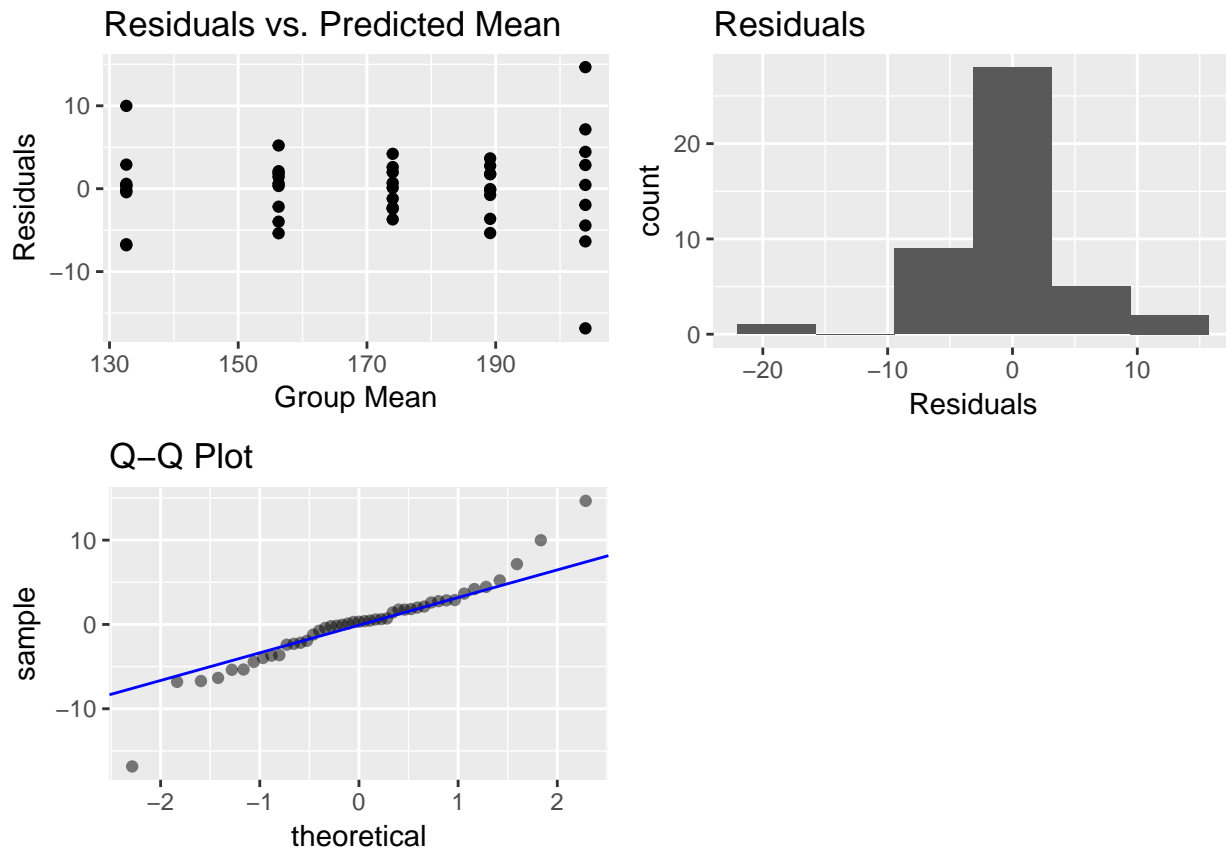
95% family-wise confidence level



9.13 d) Use Dunnett's to see if any of the new agents have significantly larger mean weights loss as compared to the standard agent. $\alpha = 0.05$

```
##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: User-defined Contrasts
##
```

```
##
## Fit: aov(formula = wt_loss ~ treatment, data = gather_frame)
##
## Linear Hypotheses:
##           Estimate Std. Error t value Pr(>|t|)
## a1 - s == 0   2.7800    0.4433   6.272 < 0.001 ***
## a2 - s == 0   1.7500    0.4433   3.948 0.00105 **
## a3 - s == 0   1.0000    0.4433   2.256 0.09304 .
## a4 - s == 0   2.9700    0.4433   6.700 < 0.001 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```



b) Perform an Anova

```
anova_lenses
```

```
## Call:
##   aov(formula = lov ~ Supplier, data = gather_lenses)
##
## Terms:
##           Supplier Residuals
## Sum of Squares 28024.350 1053.789
## Deg. of Freedom      4      40
##
## Residual standard error: 5.132711
## Estimated effects may be unbalanced
```



```
summary(anova_lenses)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Supplier    4  28024    7006   265.9 <2e-16 ***
## Residuals   40   1054     26
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
kruskal.test(x = gather_lenses$lov, g = gather_lenses$Supplier)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  gather_lenses$lov and gather_lenses$Supplier
## Kruskal-Wallis chi-squared = 41.596, df = 4, p-value = 2.023e-08
```