HW 7 Solutions, 21 pts

(Regarding Exercise 23 on p. 173 in Chapter 6 re: *Diet Wars*) A researcher has two primary research questions: (1) Is there a difference in weight loss, on the average, between the low-fat and low-carbohydrate diets? (2) Is there a difference in mean weight loss for the low-fat and low-carbohydrate diets compared to the mean of just the Mediterranean diet group?

1. (2 pts) It is appropriate to test a few planned primary research questions using an individual CI or individual test for each (as opposed to using a family of CIs or a family of tests that maintain a family-wise confidence or significance level). This way, the full power of the study is applied to these planned questions.

2. (4 pts) An ANOVA was fit to these data (R-code is in the Appendix). The ANOVA table is:

```
Analysis of Variance Table

           Df Sum Sq Mean Sq F value  Pr(>F)
Group       2  216.9 108.430  3.2358 0.04086 *
Residuals 269 9013.9  33.509
```

   The assumptions were checked using normal probability and residual vs. fits plots.

   - Based on the normal probability plot, the residuals do not appear to deviate from **normality**.
   - Based on the residual vs. fits plot, the residuals do not appear to deviate from **constant variance**. Furthermore, using our rule of thumb, the largest SD (7.0 for Low-Carb) is not twice more than the smallest SD (4.11 for Low-Fat, which also fails to discredit the constant variance assumption.
   - The humans who participated in this study were all employees from a single workplace. Furthermore, the participating employees from this workplace were " recruited" by the researchers, so there is potential for voluntary response and non-response bias. Therefore, these data are **not a random sample** from the popultion of all humans, or of Americans, or of all employees.
   - Because the employees were randomly assigned to different groups, we will assume that the **groups are independent**.

3. (5 pts) Testing the FIRST research question:

   (a) Using the notation μ_1 for the mean weight-loss of the low-carbohydrate group, μ_2 for the mean weight loss of the low-fat group and μ_3 being the mean weight loss of the Mediterranean group, the parameter γ being tested is the linear combination

   $$\gamma = \mu_1 - \mu_2.$$

   (b) The 6 steps of the hypothesis test are:

   - The hypotheses being tested are:

     $H_0 : \gamma = \mu_1 - \mu_2 = 0$ or $H_0 : \mu_1 = \mu_2$

     versus

     $H_0 : \gamma = \mu_1 - \mu_2 \neq 0$ or $H_0 : \mu_1 \neq \mu_2$
   - The assumptions were assessed in #2 above.
   - The test statistic for the follow-up t test is $t = 2.52$ (see Appendix).
   - The two-sided p-value is 0.012 (see Appendix).
   - Because p-value $= 0.012 < \alpha = 0.05$, we reject H_0 in favor of H_a.

- The conclusion is given in #3(c)

(c) The evidence suggests that the Low Carb diet gives more weight loss on average compared to the low fat diet (t=2.52, p-value = 0.012).

4. (4 pts) Testing the SECOND research question:

   (a) Using the notation μ_1 for the mean weight-loss of the low-carbohydrate group, μ_2 for the mean weight loss of the low-fat group and μ_3 being the mean weight loss of the Mediterranean group, the parameter γ being tested is the linear combination

$$\gamma = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 - \mu_3.$$

   (b) The Low Carb and Low Fat diets are estmated to give 2.54 points kg more weight loss on average compared to the Meditarre (this is an estimate for γ). A 95% CI for γ is [0.73, 4.35] (see R-code in the Appendix).

   (c) With 95% confidence, the Low Carb and Low Fat diets give 0.73 to 4.35 kg more weight loss, on the average, compared to the Mediterranean diet.

5. (3 pts) It is appropriate to test for all pairwise differences among the group means in addition to the primary research questions because the ANOVA suggests that there is some difference in the mean weight loss among the diets (p=0.041, $F = 3.2$). Testing all pairwise comparisons will be performed using a multiple comparison procedure to maintain a family-wise significance level of 95%. Although the average weight loss of the Low-Carb diet was more than either of the other two groups in this study, the evidence only suggests that the Low Carb diet causes a higher weight loss than the Low-Fat diet in a larger population (p=0.033, Tukeys). The evidence fails to suggest a difference in the mean weight loss between the Mediterranean diet and either the Low Carb ($p = 0.566$) or Low Fat (p=0.277) diets.

6. (4 pts) *Scope of Inference* Because this was a randomized experiment, then the evidence does suggest that the Low-Carb diet <u>caused</u> more weight loss compared to the Low Fat diet. Because the humans in this study were not a random sample of all humans, and were not even a random sample of the employees from the single workplace, it is unclear what population these results can be extended to beyond the participants themselves.

Appendix

Housekeeping

```r
source("http://www.math.montana.edu/parker/courses/STAT411/diagANOVA.r")
library(Sleuth3)
library(gmodels)
```

Problem 1

Get, summarize and plot the data:

```r
summary(ex0623)
```

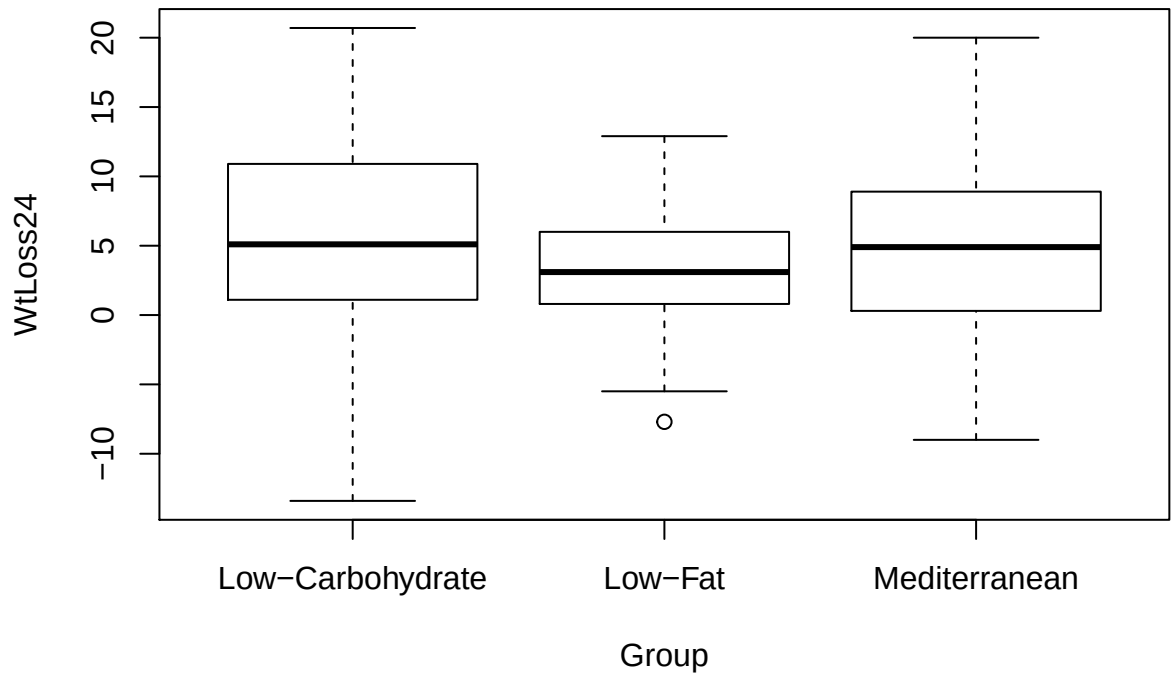
```
##     Subject                     Group       WtLoss24
##  Min.   :  1.00   Low-Carbohydrate:85   Min.   :-13.40
##  1st Qu.: 68.75   Low-Fat         :94   1st Qu.:  0.60
##  Median :146.50   Mediterranean   :93   Median :  4.10
##  Mean   :148.04                         Mean   :  4.43
```

```
##  3rd Qu.:230.25                          3rd Qu.:  8.50
##  Max.   :298.00                          Max.   : 20.70
```

```r
dim(ex0623) # n = 272
```

```
## [1] 272   3
```

```r
plot(WtLoss24 ~ Group,data=ex0623)
```



Descriptibe stats, fit an ANOVA, perform Tukeys

```r
tapply(ex0623$WtLoss24,ex0623$Group,mean)
```

```
## Low-Carbohydrate          Low-Fat    Mediterranean
##         5.487059         3.304255         4.602151
```

```r
tapply(ex0623$WtLoss24,ex0623$Group,sd)
```

```
## Low-Carbohydrate          Low-Fat    Mediterranean
##         7.004604         4.112554         6.006844
```

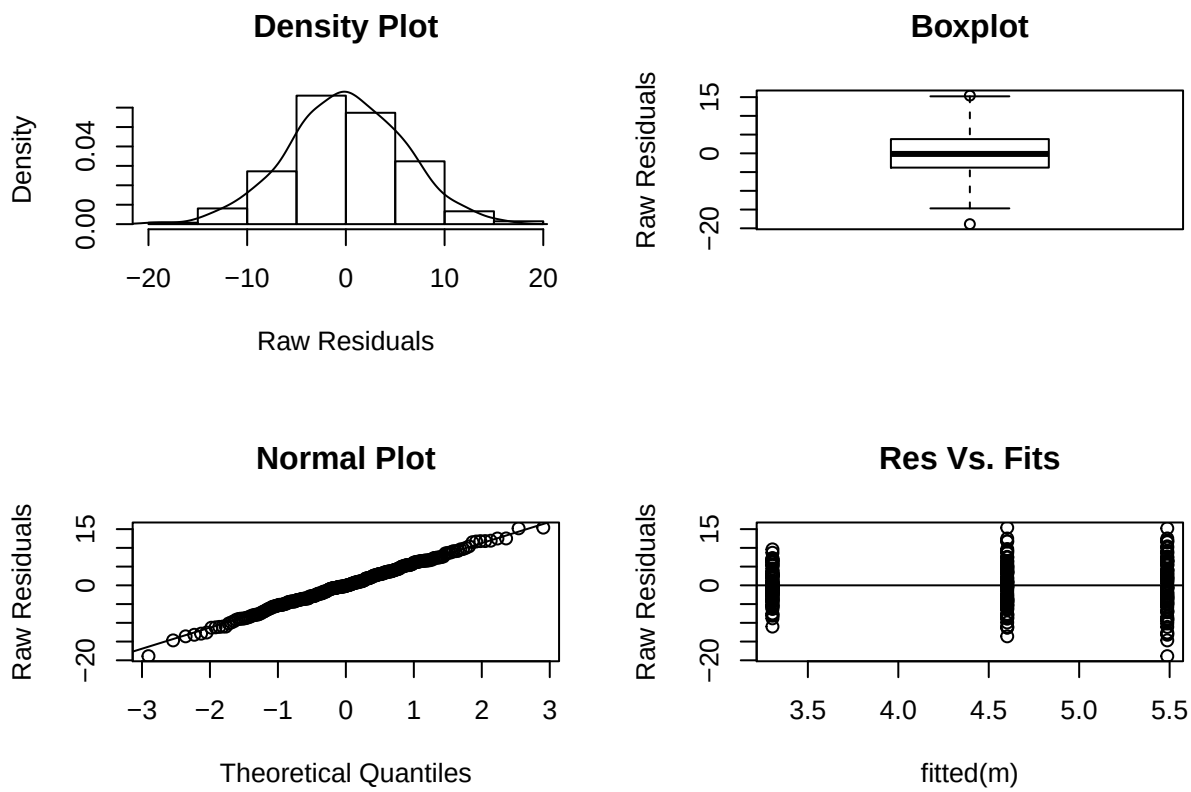
```r
m1=lm(WtLoss24 ~ Group,data=ex0623)
anova(m1)
```

```
## Analysis of Variance Table
##
## Response: WtLoss24
##            Df Sum Sq Mean Sq F value  Pr(>F)
## Group       2  216.9 108.430  3.2358 0.04086 *
## Residuals 269 9013.9  33.509
```

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

Assess model fit - the residual plots are gorgeous, indicating that the ANOVA assumptions of normality and constant variance are met. This is because the book says that they generated these data.

```
diagANOVA(m1)
```



Problem 3

Estimate the linear combination associated with the primary research question.

```
estimable(lm(WtLoss24 ~ Group -1 ,data=ex0623),c(1,-1,0),conf.int=.95)
```

```
##          Estimate Std. Error  t value  DF   Pr(>|t|)  Lower.CI Upper.CI
## (1 -1 0) 2.182804  0.8664294 2.519309 269 0.01233846 0.4769582 3.888649
```

Problem 4

Estimate the linear combination associated with the secondary research question.

```
estimable(lm(WtLoss24 ~ Group - 1,data=ex0623),c(1/2,1/2,-1),conf.int=.95)
```

```
##                Estimate Std. Error   t value  DF  Pr(>|t|)  Lower.CI
## (0.5 0.5 -1) -0.2064935  0.7402606 -0.278947 269 0.7804998 -1.663935
##              Upper.CI
## (0.5 0.5 -1) 1.250948
```

Problem 5

Perform all pairwise comparisons.

```r
TukeyHSD(aov(WtLoss24 ~ Group,data=ex0623))
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = WtLoss24 ~ Group, data = ex0623)
##
## $Group
##                                     diff       lwr        upr     p adj
## Low-Fat-Low-Carbohydrate       -2.1828035 -4.224771 -0.1408361 0.0329364
## Mediterranean-Low-Carbohydrate -0.8849083 -2.932082  1.1622656 0.5656813
## Mediterranean-Low-Fat           1.2978952 -0.697418  3.2932084 0.2771180
```