Winter, 2013   Biometry   Quiz #1

1. Rummer and Bennett (2005) examined the physiological effects of swim bladder overexpansion and catastrophic decompression on red snapper (*Lutjanus campechanus*). In one part of their study, they examined the total geometric volume of the swimbladder (volume) at four experimentally controlled treatments of pressure (pressure) -- control (101.1 kPa), low (405.3 kPa), medium (608.0 kPa), and high (1215.9 kPa). In the study, each fish was randomly selected from a large population, was randomly allocated to one of the pressure treatments, and was tested individually in an experimental tank. The author's goal was to determine if the mean volume of the swimbladder was different under the different pressure levels. The data from this experiment were entered into R and analyzed with the commands on the unstapled sheet. Answer the questions below **with the fullest amount of detail that you can provide – be specific, refer to results where appropriate** (you may want to label figures and tables on the output).

   a. **[8 pts]** On the original scale only, fully assess ALL assumptions appropriate to this analysis.

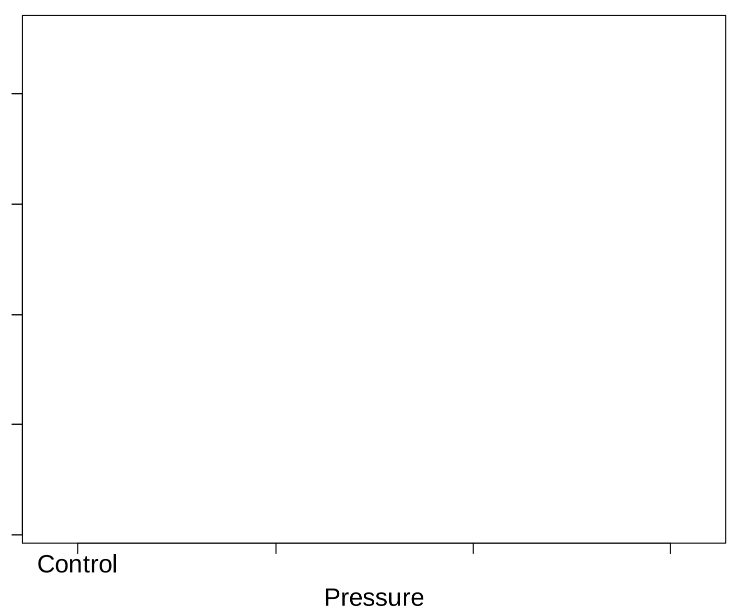
   Answers to b-e should refer to either the original or transformed results. Questions b-e will refer to "volume" but this may be interpreted as "transformed volume" if you choose to use the transformed scale. Either way, you should be very precise with your language. Your answer to f should be made on the original scale (i.e., volume).

   b. **[3 pts]** What specific conclusion about swimbladder volume can be made from the results in the ANOVA table?

   c. **[3 pts]** What is the difference in **sample** mean swimbladder volume between the control and high pressure treatments? Be sure to explicitly say which one is larger.

   d. **[4 pts]** Specifically comment on whether or not the **population** mean swimbladder volume differs significantly between EACH pair of treatments.

   e. **[4 pts]** On the schematic below, manually construct a means plot (i.e., a `fitPlot()` but without confidence intervals) from the provided results. Include letters by each mean that shows which treatments are statistically different. Make sure to label the y-axis.



   f. **[3 pts]** Specifically interpret the confidence interval for the treatments with the most different mean swimbladder volumes.

2. **[18 pts]** Sanz (2001) examined the nesting behavior and success for male and female pied flycatchers (*Ficedula hypoleuca*). He hypothesized that males would be more involved in nesting activities as the level of brood demand increased and as the "attractiveness" of the male decreased. The author manipulated brood demand by removing two eggs from randomly selected nests and placing these eggs into other randomly selected nests. This created two levels for a "clutch-size manipulation"' factor (CSM) - reduced and enlarged clutch sizes. A third level consisted of nests where the number of eggs was not manipulated (called the "control" level). The author manipulated the "attractiveness"' of the male by clipping approximately two-thirds of the white feathers on randomly selected males. Thus, the experiment consisted of two levels of the "forehead patch" factor (FH) – "unmanipulated" and "reduced." In one aspect of the experiment, Sanz recorded the number of hatchlings from each nest.

   a. The two-way ANOVA table for these data is shown below. Fill in the missing results. Note that there is no overall among source row.

Source	df	SS	MS	F	p
	2		125.178	99.2513	<2e-16
	1	0.002			0.9702
CSM:FH			0.371		0.7456
Residuals					
Total	114	388.574			

   b. What effects are AND are not evident from these results.
   c. What is the variance among individuals ignoring group membership?
   d. What is the variance among individuals within each group?
   e. How many total nests are evident in these results?

3. **[16 pts]** Answer FOUR of the six questions below (a-f).

   a. Completely compare and contrast the meanings of MS_{Within}, MS_{Total}, and MS_{Among}. *Your statements should be general but you may refer to a specific instance as an example.*

   b. Completely compare and contrast the concepts of a "full" and a "simple" model. *Your statements should be general but you may refer to a specific instance as an example.*

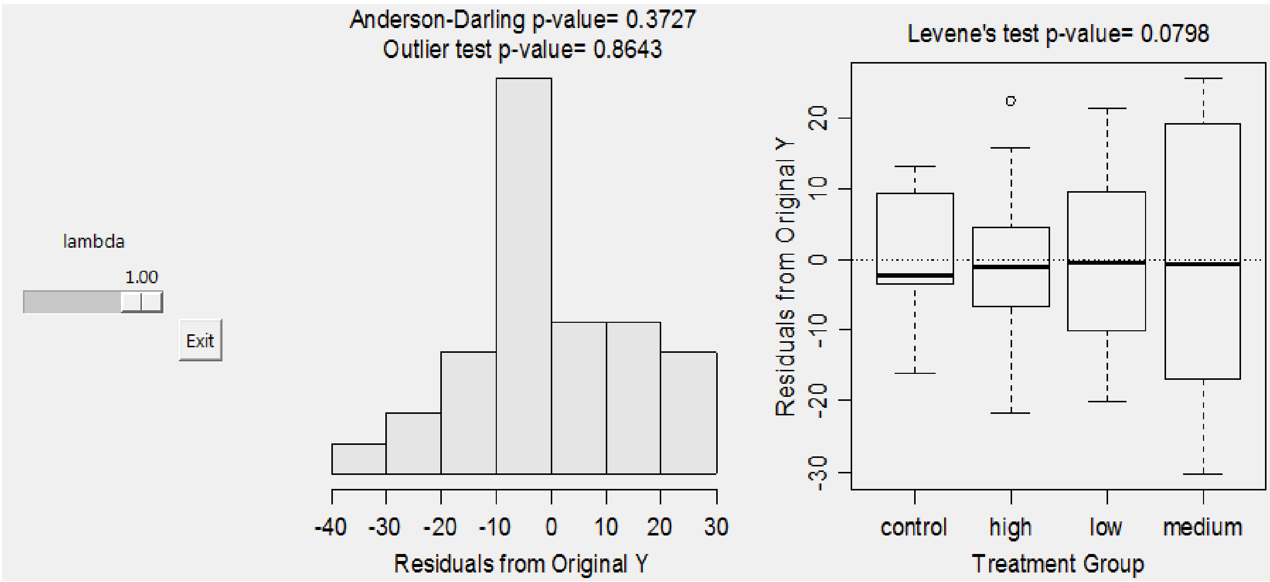
   c. Define experiment-wise and comparison-wise error rates. Identify the specific and relative sizes of each error rate.

   d. Describe when a Tukey's HSD and when a Dunnet's procedure would be appropriate to use. For the Dunnet's situation describe why the Dunnet's method would be "better" than Tukey's method in the same situation.

   e. Thoroughly explain why one experiment where two factors are simultaneously manipulated is "better" than two separate experiments where one factor at a time is manipulated. If you decided to demonstrate your points with an illustrative example, assume that there are 30 individuals available for experimentation and that one factor has two levels and the other factor has three levels.

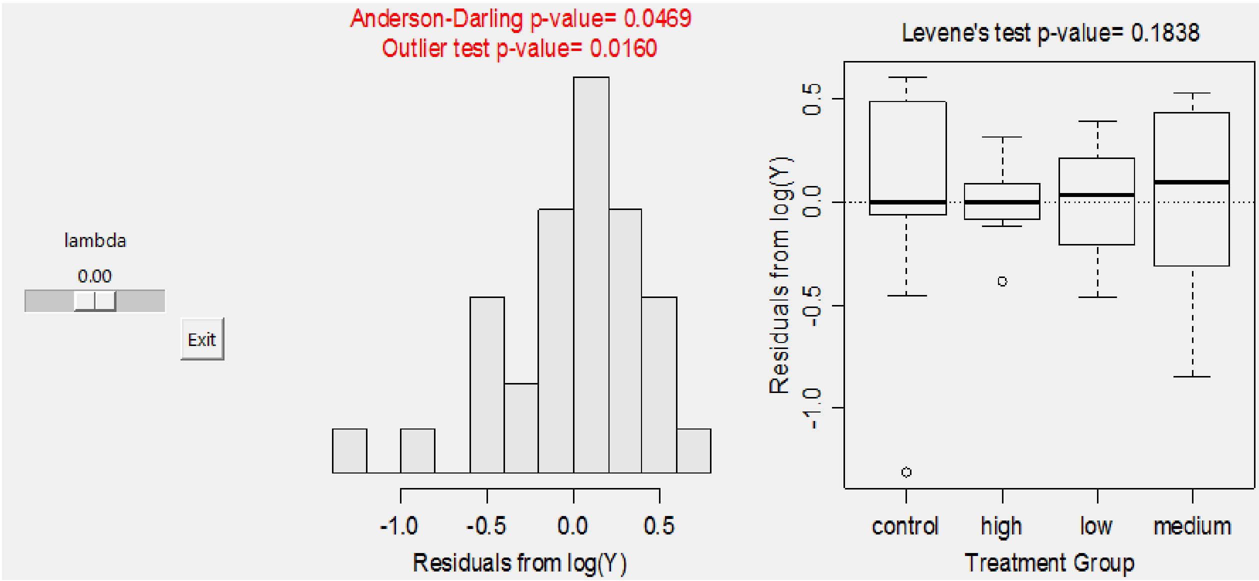
   f. Mathematically show that the difference in two means of a log-transformed variable becomes a RATIO of two means on the original scale. *You do not need complete sentences for this question but you should mathematically show each step in the proof.*

```
> library(NCStats)
> library(multcomp)
> d <- read.table("RedSnapper.txt",header=TRUE)
> str(d)
    'data.frame':   34 obs. of  2 variables:
     $ volume  : num  18.03 5.03 19.2 11.83 17.45 ...
     $ pressure: Factor w/ 4 levels "control","high",..: 1 1 1 1 1 1 1 1 1 1 ...
> lm1 <- lm(volume~pressure,data=d)
> transChooser(lm1)
```




```
> anova(lm1)
              Df   Sum Sq Mean Sq F value     Pr(>F)
    pressure   3   9880.5  3293.5  14.865 4.116e-06
    Residuals 30   6646.8   221.6
    Total     33 16527.4
> summary(lm1)
                    Estimate Std. Error t value Pr(>|t|)
    (Intercept)       21.000      4.707   4.461 0.000106
    pressurehigh      44.399      6.839   6.492 3.56e-07
    pressurelow       29.600      7.061   4.192 0.000225
    pressuremedium    28.399      7.335   3.871 0.000543
> mc1 <- glht(lm1,mcp(pressure="Tukey"))
> summary(mc1)
                           Estimate Std. Error t value Pr(>|t|)
    high - control == 0      44.399      6.839   6.492  < 0.001
    low - control == 0       29.600      7.061   4.192  0.00122
    medium - control == 0    28.399      7.335   3.871  0.00300
    low - high == 0         -14.799      7.233  -2.046  0.19378
    medium - high == 0      -16.000      7.501  -2.133  0.16610
    medium - low == 0        -1.201      7.704  -0.156  0.99861
> confint(mc1)
                           Estimate lwr      upr
    high - control == 0     44.3989  25.8094  62.9884
    low - control == 0      29.6000  10.4087  48.7913
    medium - control == 0   28.3986   8.4603  48.3369
    low - high == 0        -14.7989 -34.4583   4.8605
    medium - high == 0     -16.0003 -36.3896   4.3890
    medium - low == 0       -1.2014 -22.1408  19.7380
```

```
> transChooser(lm1)
```




```
> d$logvol <- log(d$volume)
> lm2 <- lm(logvol~pressure,data=d)
> anova(lm2)
          Df  Sum Sq Mean Sq F value    Pr(>F)
pressure   3  8.1608 2.72026  14.875 4.091e-06
Residuals 30  5.4862 0.18287
Total     33 13.6470
> summary(lm2)
               Estimate Std. Error t value Pr(>|t|)
 (Intercept)     2.9277     0.1352  21.650  < 2e-16
pressurehigh     1.2344     0.1965   6.282 6.35e-07
pressurelow      0.9556     0.2028   4.711 5.26e-05
pressuremedium   0.8688     0.2107   4.122 0.000272
> mc2 <- glht(lm2,mcp(pressure="Tukey"))
> summary(mc2)
                       Estimate Std. Error t value Pr(>|t|)
high - control == 0     1.23440    0.19649   6.282  < 0.001
low - control == 0      0.95562    0.20285   4.711  < 0.001
medium - control == 0   0.86878    0.21074   4.122  0.00144
low - high == 0        -0.27878    0.20779  -1.342  0.54408
medium - high == 0     -0.36562    0.21551  -1.697  0.34260
medium - low == 0      -0.08684    0.22132  -0.392  0.97908
> ( ci2 <- confint(mc2) )
                       Estimate lwr      upr
high - control == 0     1.23440  0.70054  1.76825
low - control == 0      0.95562  0.40448  1.50675
medium - control == 0   0.86878  0.29619  1.44137
low - high == 0        -0.27878 -0.84337  0.28580
medium - high == 0     -0.36562 -0.95116  0.21992
medium - low == 0      -0.08684 -0.68818  0.51450
> exp(ci2$confint)
                Estimate       lwr      upr
high - control 3.4363079 2.0148414 5.860616
low - control  2.6002696 1.4985180 4.512059
medium - control 2.3839929 1.3447192 4.226475
low - high     0.7567045 0.4302601 1.330827
medium - high  0.6937658 0.3862910 1.245980
medium - low   0.9168253 0.5024899 1.672807
```