**Winter, 2010 Biometry Quiz #1**

Cross (2006; American Biology Teacher, 68:347-349,351-352) measured the area (mm2) of cells in the webs of spiders. The cells were measured both before and after the web had been rebuilt following the web being sprayed with vodka, a fly being introduced to the web, the spider wrapping the fly (and consuming the vodka), and the original web being destroyed. The author’s analysis of these data was repeated in R with the results shown below.

**> library(NCStats)**

**> sp <- read.table("SpidersAlcohol.txt",head=TRUE)**

**> str(sp)**

'data.frame': 78 obs. of 2 variables:

$ cells: num 75.7 161.6 151.3 147.1 177.9 ...

$ group: Factor w/ 2 levels "after","before": 1 1 1 1 1 1 1 1 1 1 ...

**> sp.lm <- lm(cells~group,data=sp)**

**> anova(sp.lm)**

Analysis of Variance Table

Response: cells

Df Sum Sq Mean Sq F value Pr(>F)

group 1 39321 39321 12.006 0.000875 \*\*\*

Residuals 76 248910 3275

**> summary(sp.lm)**

Call:lm(formula = cells ~ group)

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 140.900 9.164 15.375 < 2e-16 \*\*\*

groupbefore -44.905 12.960 -3.465 0.000875 \*\*\*

Residual standard error: 57.23 on 76 degrees of freedom

Multiple R-Squared: 0.1364, Adjusted R-squared: 0.1251

F-statistic: 12.01 on 1 and 76 DF, p-value: 0.000875

**> confint(sp.lm)**

2.5 % 97.5 %

(Intercept) 122.64844 159.15156

groupbefore -70.71673 -19.09353

**[24 pts]** The information above should be used to answer questions 1-7 below on a separate sheet of paper. For each question, as appropriate, **show your work, refer to a specific part of the output (by the R function name), and describe your rationale**.

1. Not including the background information, what **two** pieces of evidence indicate that there are two groups?
2. What **two** pieces of evidence indicate that there are 78 individuals?
3. What conclusion (about spider webs and alcohol) should be made from the p-value in the ANOVA table?
4. What is the sample mean size of cells for the spiders before consuming alcohol?
5. What is the sample mean size of cells for the spiders after consuming alcohol?
6. What is the 95% CI for the after-before parameter?
7. **[1 pt - XC]** Interpret the multiple R2 value?

A simple experiment was conducted to examine the effect of the presence of fish on the use of shelters by crayfish. A total of 30 tanks stocked with 60 crayfish each were used for the experiment. Each tank was randomly selected to be stocked with two rock bass, two black crappies, or one smallmouth bass (so that each species was stocked into 10 tanks each). After an acclimatization period the number of crayfish using shelters at a randomly selected time was recorded. The data from this experiment were entered into R and analyzed with the commands below. Answer the questions further below on a separate sheet of paper **with the fullest amount of detail that you can – be specific, refer to results where appropriate.**

**> HCF <- read.table("HiddenCrayfishData.txt",head=TRUE)**

**> str(HCF)**

`data.frame': 30 obs. of 2 variables:

$ hidden: int 32 35 36 17 26 30 29 33 26 42 ...

$ group : Factor w/ 3 levels "BC","RB","SMB": 2 2 2 2 2 2 2 2 2 2 ...

**> lm1 <- lm(hidden~group,data=HCF)**

**> levene.test(lm1)**

Df F value Pr(>F)

group 2 0.6337 0.5383

27

**> ad.test(lm1$residuals)**

data: lm1$residuals

A = 0.319, p-value = 0.5181

**> outlier.test(lm1)**

max|rstudent| = 1.925558, degrees of freedom = 26,

unadjusted p = 0.0651627, Bonferroni p > 1

Observation: 25

**> anova(lm1)**

Response: hidden

Df Sum Sq Mean Sq F value Pr(>F)

group 2 1160.07 580.03 8.9752 0.001027 \*\*

Residuals 27 1744.90 64.63

**> mc <- glht(lm1,mcp(group="Tukey"))**

**> summary(mc)**

Estimate Std. Error t value p value

RB - BC == 0 -2.900 3.595 -0.807 0.70222

SMB - BC == 0 11.500 3.595 3.199 0.00957 \*\*

SMB - RB == 0 14.400 3.595 4.005 0.00116 \*\*

**> confint(mc)**

Estimate lwr upr

RB - BC == 0 -2.9000 -11.8144 6.0144

SMB - BC == 0 11.5000 2.5856 20.4144

SMB - RB == 0 14.4000 5.4856 23.3144

1. **[8 pts]** Use the results and description from above to fully assess ALL assumptions of a one-way ANOVA.
2. **[10 pts]** Complete the following questions with the partial generic ANOVA table below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **Df** | **SS** | **MS** | **F** | **p-value** |
| **Among Groups** | 7 |  | 12.23 |  |  |
| **Within Groups** |  | 192.16 |  |  |  |
| **Total** | 23 |  |  |  |  |

* 1. Fill in the remainder of the ANOVA table above (excluding the p-value).
  2. How many groups were in this analysis? \_\_\_\_\_\_\_\_\_
  3. What is the pooled estimate of the variance among individuals in each group? \_\_\_\_\_\_\_\_\_

1. **[15 pts]** Answer 3 of the 6 questions below (a-f) with COMPLETE SENTENCES.
   1. 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.*
   2. Completely compare and contrast the meanings of MSWithin, MSTotal, and MSAmong. *Your statements should be general but you may refer to a specific instance as an example.*
   3. Mathematically prove that SSTotal partitions exactly into SSWithin and SSAmong. *You do not need complete sentences for this question but you should show each step in the proof.*
   4. Define experiment-wise and comparison-wise error rates. Identify the relative size of each error rate.
   5. Describe when a Tukey HSD and when a Dunnet’s procedure would be appropriate to use. For the Dunnet’s situation describe why the Dunnet’s method is “better” than Tukey’s method for the same situation.
   6. 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 show each step in the proof.*
2. **[3 pts]** Explain why you are taking this course (“because it is required” is not an acceptable answer).