**Winter, 2010 Biometry Quiz #3**

1. **[30 pts]** Welbergen (2010; Journal of Mammalogy 91:38-47) examined the relationship between body mass (grams) and forearm length (mm) for three groups of gray-headed flying foxes (*Pteropus poliocephalus*) – adult males, adult females, and juveniles (labeled as “young”). Welbergen’s primary interest with these data was to see if body mass could be predicted from forearm length and whether those predictions were dependent on the group (variable called “sex”) of flying fox. Use this information to answer the following questions (these do not have to be complete sentences).
   1. **[2 pts]** Construct all necessary indicator variables assuming that the “young” flying fox group is to be the reference group [*Give an explicit name and definition to each variable.*]
   2. **[3 pts]** Construct the ultimate full model for these data. [*Variables should be entered into the model according to the order discussed in class.*]
   3. **[3 pts]** Construct all submodels for these data. [*Show your work.*]
   4. **[3 pts]** Fully and carefully interpret the meanings of , 1, 2, and 1.
   5. **[2 pts]** Show the simple and full models for the equal intercepts test (assuming parallel lines). [*Carefully label each model.*]
   6. **[2 pts]** Show the simple and full models for the parallel lines test. [*Carefully label each model.*]
2. **[10 pts]** Answer ***two*** of the three following questions.
3. Suppose that you have collected a quantitative response variable. In addition you may have collected (i) a quantitative explanatory variable, (ii) a categorical explanatory variable, (iii) both a quantitative and categorical explanatory variable, or (iv) two categorical explanatory variables. Identify which statistical procedure/test you would use in each of these four situations and describe the type(s) of question(s) that you would be able to answer with each.
4. The relationship between the number of decaying objects in a discrete set (e.g., number of tulip bulbs in a box that rot) and time can be modeled with



where N0 (the constant initial number of objects) and  (the mean lifetime) are parameters to be estimated, Nt (number of objects at time t) and t (time) are variables, and e is the “natural number” constant. Algebraically show how to linearize this function and then explain how the two parameters can be estimated from the slope and intercept of the fitted transformed function.

1. Suppose that you have conducted a one-way ANOVA and received a particular result (either the groups have different means or they do not). Now suppose that you introduce a quantitative covariate to the analysis that is significantly related to the response variable (and the relationship between the covariate and the response does not differ by group). The introduction of this covariate to the analysis could lead you to reverse your original conclusion – i.e., changing from concluding that the groups differ to concluding that they do not differ (or vice versa). Explain how each of these reversals could happen (you likely will want to refer to variabilities “explained” and “not explained” and may want to draw some explanatory graphics).
2. **[30 pts]** Welbergen’s data (from the question #1) was entered into R and analyzed with the commands shown on the following pages. Use these results to answer the following questions.
   1. Two models are presented in the results. Which model do you choose to analyze and why? Be as specific as possible with evidence to support your choice.

Use the model you identified above to answer the following questions. Note that I have put the variable names in quotes because you may be referring to a transformed version of the variable (and you should do so in your answer).

* 1. What can you say about the relationship between “body mass” and “forearm length” for these flying foxes? Be as explicit as you possibly can but make sure to comment on any differences between groups (i.e., sexes) and state specifically what the relationship is for each group. Provide evidence for your conclusions.
  2. What can you say, if anything, about differences in “body mass” among groups (i.e., sexes) after having factored out the effect of “forearm length” on “body mass”? Again be as explicit as you possibly can. Provide evidence for your conclusions.
  3. At the end of the results for your chosen model are results from predict(). Interpret each of these values and comment on whether these values are statistically significantly different or not (with supporting evidence).

**> library(NCStats)**

**> ff <- read.table("http://www.ncfaculty.net/dogle/R/Data\_Master/R/FlyingFox1.txt",header=TRUE)**

**> str(ff)**

'data.frame': 223 obs. of 3 variables:

$ forearm.len: num 152 154 157 157 158 ...

$ body.mass : num 779 788 839 761 763 ...

$ sex : Factor w/ 3 levels "female","male",..: 2 2 2 2 2 2 2 2 2 2 ...

**> ff$sex <- factor(ff$sex,levels=c("young","male","female"))**

**> table(ff$sex)**

young male female

79 75 69

**> lm1 <- lm(body.mass~forearm.len\*sex,data=ff)**

**> trans.chooser(lm1,show.stats=TRUE)**



**> anova(lm1)**

Response: body.mass

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

forearm.len 1 9051340 9051340 1792.2167 < 2e-16 \*\*\*

sex 2 987212 493606 97.7368 < 2e-16 \*\*\*

forearm.len:sex 2 30210 15105 2.9908 0.05232 .

Residuals 217 1095928 5050

**> comp.slopes(lm1)**

Multiple Slope Comparisons

comparison diff lwr upr raw.p adj.p

1 male-young 5.67494 0.80082 10.54905 0.02270 0.06178

2 female-young -0.08093 -4.76632 4.60445 0.97287 0.97287

3 female-male -5.75587 -11.27945 -0.23228 0.04119 0.06178

Slope Information

level slopes lwr upr raw.p adj.p

3 female 7.2419 3.4534 11.030 0.00021 0.00021

1 young 7.3229 4.5661 10.080 0.00000 0.00000

2 male 12.9978 8.9782 17.017 0.00000 0.00000

**> comp.intercepts(lm1)**

Tukey HSD on adjusted means assuming parallel lines.

comparison diff lwr upr p.adj

1 male-young 274.22 247.00 301.44 0

2 female-young 117.43 89.61 145.25 0

3 female-male -156.79 -184.95 -128.63 0

Mean adjusted values at a covariate value of 154.9

young male female

490.56 764.78 607.99

**> fit.plot(lm1,legend="topleft",col="gray")**



**> predict(lm1,data.frame(forearm.len=c(155,155,155),sex=c("young","male","female")))**

1 2 3

471.0677 723.9145 619.1563

**> ff$log.bm <- log(ff$body.mass)**

**> ff$log.fl <- log(ff$forearm.len)**

**> lm2 <- lm(log.bm~log.fl\*sex,data=ff)**

**> anova(lm2)**

Response: log.bm

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

log.fl 1 31.4080 31.4080 2588.8126 < 2e-16 \*\*\*

sex 2 1.8891 0.9445 77.8534 < 2e-16 \*\*\*

log.fl:sex 2 0.0606 0.0303 2.4973 0.08467 .

Residuals 217 2.6327 0.0121

**> comp.slopes(lm2)**

Multiple Slope Comparisons

comparison diff lwr upr raw.p adj.p

1 male-young -0.43715 -1.6134 0.73908 0.46465 0.46465

2 female-young -1.26771 -2.3880 -0.14740 0.02675 0.08025

3 female-male -0.83056 -2.2220 0.56091 0.24070 0.36105

Slope Information

level slopes lwr upr raw.p adj.p

3 female 1.7103 0.75957 2.6610 0.00048 0.00048

2 male 2.5409 1.52483 3.5569 0.00000 0.00000

1 young 2.9780 2.38538 3.5706 0.00000 0.00000

**> comp.intercepts(lm2)**

Tukey HSD on adjusted means assuming parallel lines.

comparison diff lwr upr p.adj

1 male-young 0.43712 0.39502 0.47921 0

2 female-young 0.24246 0.19944 0.28548 0

3 female-male -0.19465 -0.23821 -0.15110 0

Mean adjusted values at a covariate value of 5.039

young male female

6.1314 6.5685 6.3739

**> fit.plot(lm2,legend="topleft",col="gray")**



**> predict(lm2,data.frame(log.fl=c(5.04,5.04,5.04),sex=c("young","male","female")))**

1 2 3

6.172781 6.575204 6.419336