**Quiz 3 ANOVA**

**#1.**  (29 points) The following data consists of a 3 (age) x 2 (sex) natural design in which the proportion of pretend play between parents and infants changes as a function of age. The DV (Y) is the observed proportion of parent-infant play that consists of pretend play. The first independent variable (X1) is age, where 1 = 7-month old infants; 2 = 10-month old infants; and 3 = 13-month-old infants. The second independent variable (X2) is sex, where 1=girls and 2=boys. There are 8 infants in each cell of the design.

The data are listed below in triplets to be read as (X1, X2, Y) (e.g. input x1 x2 y @@;)

1 1 0.02 1 1 0.01 1 1 0.07 1 1 0.04 1 1 0.01 1 1 0.09 1 1 0.05 1 1 0.06

1 2 0.05 1 2 0.01 1 2 0.04 1 2 0.03 1 2 0.02 1 2 0.02 1 2 0.13 1 2 0.06

2 1 0.15 2 1 0.11 2 1 0.22 2 1 0.05 2 1 0.09 2 1 0.05 2 1 0.15 2 1 0.11

2 2 0.14 2 2 0.21 2 2 0.06 2 2 0.12 2 2 0.11 2 2 0.19 2 2 0.12 2 2 0.04

3 1 0.09 3 1 0.03 3 1 0.18 3 1 0.12 3 1 0.18 3 1 0.43 3 1 0.24 3 1 0.4

3 2 0.02 3 2 0.19 3 2 0.15 3 2 0.07 3 2 0.45 3 2 0.2 3 2 0.49 3 2 0.19

The questions to be answered for this dataset are as follows:

(+5) a. Reproduce a basic ANOVA Table (main effects + interaction) for this analysis.

(+2) b. What is the value of RMSE for this dataset?

(+4) c. What is the F and p-value for the age main effect? What are the models being compared?

(+2) d. What is the F and p-value for the sex main effect?

(+4) e. What models are being compared for the test of the interaction effect? What is the F and p-value for this test?

(+2) f. Based on your analysis, what do you conclude is the final model for this dataset?

( +10) g. Your instructor is on an island, and can only be rescued if you solve the following problem. Test the significance of the linear and quadratic trends in age using a) regression modeling, and b) ANOVA contrasts (the contrasts are linear = -1 0 1 and quadratic = 1 -2 1). What, if any, differences do you see and why?

(3 point bonus) h. Now, test the age linear trend by sex interaction in the data above, using ANY METHOD you are comfortable with. Report the results and the hypotheses being tested in this model.

**#3. (27 points – 3 for each) The two-way ANOVA, non-orthogonal case, has been a vexing problem for ANOVA researchers for many years. Please answer the following questions concerning the two-way non-orthogonal ANOVA.**

**a. Explain the meaning of the word non-orthogonal.**

**b. What condition(s) must exist for non-orthogonality to occur? Be specific.**

**c. Why is the non-orthogonal ANOVA more difficult than the orthogonal ANOVA?**

**d. What is a Type I effect for the main effects of A and B in terms of model comparisons?**

**e. What is a Type II effect for the main effects of A and B in terms of model comparisons?**

**f. What is a Type III effect for the mains effects of A and B in terms of model comparisons?**

**g. What assumption is made for a Type II effect or Type II means to be legitimate?**

**h. What assumption is made for a Type I effect or a Type I mean to be legitimate?**

**i. What is the equivalent problem in regression?**

**#3. Consider the following 2-way ANOVA Table with the group number listed in the cells of the table.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Factor** | **B=1** | **B=2** | **B=3** | **B=4** |
| **A = 1** | **Group = 1** | **Group=2** | **Group = 3** | **Group=4** |
| **A = 2** | **Group=5** | **Group=6** | **Group = 7** | **Group=8** |

**(+16) a. Construct an ANOVA Design matrix using ) right-reference cell coding , and b) no replications within each cell. Below each column of the design matrix, put the ANOVA model parameter that the column values refers to in the matrix multiplication solution of the least-squares estimation problem. (Hint: There should be no more than 8 columns and 8 parameters).**

**(+6) b. To fit a cell-mean model in SAS, write the appropriate GLM code below.**

**( +6) c. Using the cell-mean model, write a contrast statement that tests for 1) the A main effect and separately, 2) the B main effect.**

**(+2 BONUS). Using the cell-mean model, write a contrast statement that tests for a linear trend in the B effect.**