STATISTICS 402 - Assignment 5

Due February 18, 2009

1. A psychologist does a study involving memory. One theory regarding memory states that verbal material is remembered as a function of the degree to which it is processed when it is initially presented. To test this theory a study is run by placing older literate adults into five learning groups. The Counting group reads through a list of words and counts the number of letters in each word. The Rhyming group reads a list of words and thinks of words that rhyme with each word on the list. The Adjective group reads a list of words and thinks of adjectives that modify the words, one adjective for each word. The Imagery group reads a list of words and forms a vivid image for each word. None of these four groups is told that they will later have to recall the words. Subjects in a fifth group, the Intentional group, are asked to memorize a list of words for later recall. After subjects read through a list of words three times, they are asked to write down all the words they can remember. The number of words correctly recalled is noted for each subject.
   a) Why is this study an experiment and not an observational study?
   b) Identify the response, conditions and experimental units.
   c) Give an example of two outside variables that should be controlled in this experiment. Explain briefly how each can be controlled.
   d) If the experimenter wishes to detect a difference in mean number of words correctly remembered as small as 1 standard deviation with Alpha=0.05 and Beta=0.05, how many study subjects are needed?
   e) The experimenter is able to recruit 50 study subjects. Give two combinations of Alpha, Beta and $\frac{\Delta}{\sigma}$ that correspond to this number of volunteers. Briefly explain the trade-off between your two choices.
   f) Describe in detail how you would randomly assign the study subjects to the five groups so that there are an equal number of volunteers in each group. Once you have described what you will do, actually do the randomization. Include your randomized assignment of the 50 study subjects to the 5 groups.

2. A completely randomized experiment, as described in problem 1, is run with the following results.

	Counting	Rhyming	Adjective	Imagery	Intentional
	9	7	11	12	10
	8	9	13	11	19
	6	6	8	16	14
	8	6	6	11	5
	10	6	14	9	10
	4	11	11	23	11
	6	6	13	12	14
	5	3	13	10	15
	7	8	10	19	11
	7	7	11	11	11
Mean, $\bar{y}$	7.0	6.9	11.0	13.4	12.0
Std. Dev., s	1.826	2.132	2.494	4.502	3.742

a) Plot the data. If you use a computer program to plot the data, cut and paste the output on your answer sheet. Based on the plot what can you say about the effectiveness of the various learning methods? What can you say about the consistency of the learning methods?
b) Estimate the effect of each of the learning methods.
c) Construct an analysis of variance table giving sources of variation, degrees of freedom, sums of squares, mean squares, appropriate F statistic and associated P-value. If you use a computer package, you can copy the results from the output onto your answer sheet.
d) Give the value of R^2 and an interpretation of this value.
e) Are there statistically significant differences amongst the five groups in terms of mean counts of remembered words? Support your answer by referring to the appropriate test of the null hypothesis: $\mathbf{H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5}$.
f) If there are statistically significant differences, which groups are different? Support your answer with a multiple comparisons method. Be sure to discuss the chance of committing at least one error in the set of pair-wise comparisons of groups.
g) Look at the distribution of residuals. What does this tell you about the conditions necessary for the analysis of variance? Be sure to tell me what you see in the plots and how that relates to the specific conditions.
h) Write a brief summary (one or two sentences is enough) of the findings of the experiment. In this summary make a recommendation based on your analysis as to what learning method should be used.

Note: If you use a computer program to analyze the data, include only the appropriate output to answer each question. It is helpful if the output used to answer a question is specifically referenced.