1. (8 pts total) Thirty trainees are randomly divided into three groups of 10 and each group is given instruction in the use of a different word processing system. At the end of the training period, each trainee is given the same “benchmark” word-processing project to complete and the time to completion is recorded. The cell means one-way ANOVA model will be used to test whether or not the mean completion time is the same for the three systems.
   a. (1 pt) Identify the response variable.
   b. (2 pts) Identify the factor and factor levels.
   c. (1 pt) Is the factor an experimental or observational factor?
   d. (1 pt) Suppose each trainee is allowed to select the word-processing system of his or her choice rather than being randomly assigned to a system. Discuss any potential problems with this approach.
   e. (3 pts) Construct a partial ANOVA table which includes the sources of variation and degrees of freedom for each source.

2. (5 pts) Suppose that you are designing an experiment with 6 treatments. You have reason to believe from a pilot study that a good estimate for \( \sigma \) will be 2. You are interested in detecting any difference of size 3 or greater between the treatment means. Use table B.12 in KNNL to determine the number of observations needed per treatment group to detect such differences with powers of 0.70, 0.80, and 0.90, at significance levels of 0.05 and 0.01. For a power of 80% and significance level of 0.05, how may observations would you need in total?
3. (22 pts total) Use the dataset from Problem 16.11 described on page 725 in KNKL (CH16PR11.TXT) and available on my website. The columns are weight, machine number, and observation number. Read the data into SAS accordingly. Assume a significance level of \( \alpha = 0.05 \). This analysis may be continued in a later homework, so make sure to save your SAS code!

   a. (3 pts) Create aligned (side-by-side) boxplots for the six filling machines (you can use PROC BOXPLOT for this). Just from looking at the plots, do any of the factor level means appear to be different?

   b. (2 pts) Use PROC MEANS to obtain the sample sizes, means, and standard deviations for the six different filling machines. Note, if you are using my sample code, you may want to take out the ‘noprint’ option, so that you can view all of the output.

   c. (3 pts) Use PROC REG to obtain the estimated mean weights for the six machines, by creating indicator variables. Show your work here. Check to see if these correspond to the means from part b.

   d. Examine the question of whether or not the six filling machines give the same average amount of detergent.

      i. (5 pts) Write the cell means model for this analysis (including assumptions). State the null and alternative hypotheses in terms of the cell means model parameters, give the test statistic with degrees of freedom, the p-value, and your conclusion.

      ii. (5 pts) Write the factor effects model for this analysis (including assumptions). State the null and alternative hypotheses in terms of the factor effects model parameters, give the test statistic with degrees of freedom, the p-value, and your conclusion.

   e. (2 pts) For the factor effects model, obtain the factor effects and show that they sum to zero. Show your work.

   f. (2 pts) Obtain confidence intervals of the mean weight for each of the six filling machines, using the Bonferroni procedure to control the overall Type I error rate at 0.05. Give the standard error of the factor level means (should be the same for all six machines).