Comp 528 Computer Systems Performance Analysis
Spring 2005

Homework Assignment 3
Due: March 17

Design and Analysis of Experiments

1. **Constructing a procedure for building experimental design sign tables.** (25 points) Write a MATLAB function

   ```
   function [a] = sign_table_2pk(p,k)
   ```

   to construct a sign table for an arbitrary two-level experimental design. The procedure should be capable of generating sign tables for full factorial and fractional factorial designs, with or without replications. The function arguments are `k`, the number of factors, and `p`, $\texttt{p} \geq 0$, where $2^{\texttt{k}-\texttt{p}}$ is the number of experiments to be performed. The function returns the sign table `a`.

   For credit, submit the procedure, the name of a file on the RTC where I can find the code, and a brief write-up that explains the process of constructing sign tables.

2. **Constructing a procedure for analyzing experimental designs.** (25 points) Write a MATLAB procedure to analyze an arbitrary two-level experimental design. It must analyze full factorial and fractional factorial designs, with or without with replications. (Hint: all designs can be analyzed in the same way; there is no need to distinguish among the cases.) The program should take two arguments: `a`, a sign table, and `y` a matrix of responses. Each column in `y` represents the results from a set of experiments for different factor combinations; each row in `y` represents the replications of a particular experiment. The program should return three vectors: `q`, a vector of coefficients representing the effect of each factor or interaction, `pv` a vector that contains the percent of variation associated with each factor or interaction, and `c` a vector that contains a confidence interval for each q_i coefficient. (If there are no replications, the confidence interval you return should be zero length, with the start and endpoints equal to q_i.) I recommend applying the routine to several of the case studies in Chapters 17, 18 and 19 in Jain to check your work.

   I will evaluate your solutions to problems 1 and 2 by using it to evaluate case study 19.2 in Jain. Explain how I should use your procedures to analyze case study 19.2.

   For credit, submit the procedure, the name of a file on the RTC where I can find the code, and a brief write-up that explains what your analysis routine does and why.

3. **Analyzing experimental designs.** (15 points)

   (a) Apply your procedures to analyze the following $2^2 3$ experimental design.

Workload	Processor A	Processor B
I	(41.2, 39.0, 42.2)	(64, 59.1, 63.4)
J	(51.4, 52.5, 50.5)	(48.1, 49.0, 47.2)

   Write the model equation that the results represent. How much variation is explained by each of the factors and their interaction?

   (b) Use visual tests to compare the residuals with the predicted response and determine if the errors are normally distributed. Submit your plots.

   (c) Interpret your results.

4. **Designing an experiment for analyzing the effects of factors.** (15 points) Design an experiment to compare implementations of *heapsort* and *quicksort*. The design should consider the following factors:

   - short vs. long input files,
   - almost sorted input records vs. essentially random input records,
   - string keys of arbitrary length vs. integer keys, and
   - keys only vs. records.

   The experiment should produce confidence intervals for the effects. Don't use a full factorial design.

   (a) Describe your design. Show the sign table for the design labeled with the factors.

   (b) What confoundings are present in the design. Argue why you believe your choice of confoundings is a good one.

   (c) What is the resolution of the design? Is there one with higher resolution? If so, why didn't you choose it?

   (d) Describe your considerations for picking the inputs representing the levels for each factor to be evaluated with the design. (Hint: Think about the execution times your experiments will yield. Inputs will need to be selected appropriately so that the analysis of the results is sensible.)

   (e) Is the model additive or multiplicative and why? What is your model equation?

5. **Designing a study.** (20 points) Design a study to compare the "performance" of the *Windows Media Player* and *Quicktime Player* network browser plugins. Describe what you think the "performance" of a media player means. How can you quantify what you perceive as important? Is performance easily summarized by a single number or multiple numbers? What factors should you consider? Is your experimental plan represented by a single experimental design, or multiple designs? Why? Describe your design or designs briefly. Does your design include replications? (You do not need to perform the comparison, just design it!). Describe your study in no more than 1.5 pages.