Empirical Methods (CS 5453)
Homework 2

March 10, 2011

This homework assignment is due on Thursday, March 31st at 5:00pm. Your work may be handed in electronically (use the **Homework 2** digital dropbox on D2L) or in hardcopy form.

This assignment must be done individually: do not share/discuss your answers with others or look at the answers of others.

All data sets are contained within the hw2_?.mat and assessment.xls files available on the main homework page.

Question 1

1. (10pts) Suppose that we are performing a robot navigation experiment in a busy building. The foot traffic level is a potential factor that could influence our performance metric (people could interfere with the robot's ability to navigate from one location to another). Should we consider this factor as an *extraneous* or a *noise* variable? Explain in detail.

2. (10pts) What is the relationship between a hypothesis and *extraneous variables*?

3. (10pts) Suppose that we were to consider foot traffic as a noise variable. Explain the process of controlling it.

4. (10pts) Suppose that we were to consider foot traffic as an extraneous variable. Explain the process of controlling it.

5. (10pts) Suppose that in comparing the performance of navigation algorithms A and B, we find that we have a ceiling effect. What could we change about the experimental design to solve this problem?

6. (10pts) Define *censoring*. Why would we need to use censoring in this experiment? Can we design the experiment so as to always avoid the need for censoring?

7. (10pts) Define *sampling bias*. Could a sampling bias exist in this experiment?

Question 2

The matlab variables "dat1," "dat2," and "dat3" contain a set of 3-tuple discrete observations (each of the data sets is represented as a single matrix). Columns 1 and 2 are independent variables; column 3 is a dependent variable.

1. (20pts) For dat1, describe in detail (including appropriate statistical tests) the influences that the independent variables have on the dependent variable.

2. (20pts) For dat2, describe in detail (including appropriate statistical tests) the influences that the independent variables have on the dependent variable.

3. (20pts) For dat3, describe in detail (including appropriate statistical tests) the influences that the independent variables have on the dependent variable.

4. (20pts) Describe a reasonable approach for formally testing whether there is an interaction effect between the independent variables in their influence of of the dependent variable. (you will need to invent this procedure.)

5. (10pts) For dat1, is there an interaction effect between the two independent variables?

6. (10pts) For dat2, is there an interaction effect between the two independent variables?

7. (10pts) For dat3, is there an interaction effect between the two independent variables?

Question 3

The file *assessment.xls* contains data collected from a recent classroom study. For each of two projects (labeled A and B), groups were assessed using three different types of metrics: *quality, originality* and *elegance.* For each metric, groups were compared to the population of "all" computer scientists (*absolute* measure) and their classroom peers (*relative*). For each of these six measures, groups were assigned a score from one (lowest) to five (highest).

Project A was performed first, followed by project B. Between the two projects, an intervention was performed. The intervention included a different way of defining and explaining the project.

We would like to argue that 1) the students performed better on project B than project A in these metrics, and 2) the intervention was responsible for this improvement.

1. (10pts) Draw an abstract graphical model that captures this hypothesis and expresses the relationships between the key variables. You may compress the performance metrics into a smaller set of metrics and you many introduce other reasonable variables.

2. (20pts) Do the students perform better in project B than in project A? Make this argument in detail, including appropriate hypothesis tests.

3. (10pts) What do you conclude about the hypothesis?