UNM Statistics Qualifying Exam Take-Home
January 2012

Due 12:00pm Jan 17 Tuesday, 2012. Return to Ana Parra Lombardin the Math/Stat Dept Office.

Directions: The answer to each problem should be presented as a summary. It should be word processed and double spaced. **A suggested length of the report to each problem is 4 pages**. Create brief, well-organized appendixes for each problem.

You may **not** consult any other person when working on this exam or discuss your exam with anyone else regardless of whether or not the person is taking the exam. You may use your course notes as well as any available books or web resources for the exam. Questions pertaining to clarification about these questions can be directed to Yan Lu, luyan@math.unm.edu.

1. (50 pts) The data for this problem is available at
   "http://www.math.unm.edu/∼luyan/bostonhousing.txt". This data set contains housing data for 506 observations on 13 variables from Boston.
   crim: per capita crime rate by town
   zn: proportion of residential land zoned for lots over 25,000 sq.ft
   indus: proportion of non-retail business acres per town
   chas: Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)
   nox: nitric oxides concentration (parts per 10 million)
   rm: average number of rooms per dwelling
   age: proportion of owner-occupied units built prior to 1940
   dis: weighted distances to five Boston employment centres
   rad: index of accessibility to radial highways
   tax: full-value property-tax rate per USD 10,000
   ptratio: pupil-teacher ratio by town
   lstat: percentage of lower status of the population
   medv: median value of owner-occupied homes in USD 1000's, target (or response) variable

   Using the data, build an appropriate regression model to predict medv using a subset of the rest variables. Summarize your findings. Make sure you tell the reader what you found and how you found it. Make sure you include important results in your report and put some necessary outputs in appendix.

Table 1:		
Premix speed	Finish mix speed	Center heights
1	1	24,17
1	2	23,22
1	3	19,15
2	1	23,30
2	2	22,22
2	3	20,20
3	1	19,10
3	2	21,22
3	3	27,29

2. (50 pts) An experiment was conducted on the effect of premixing speed and finish mixer speed on the center heights of cakes. Three different levels of speed were chosen for each of the two variables; increasing speed level indicates increasing mixing speed. Two cakes were baked at each level of premixing and finish mixing speed. The data are as in Table 1.

(a) (5 pts) What is the experimental design of the data in Table 1.

(b) (35 pts) Keep the experimenter's goal in mind, write a thorough report. Make sure you tell the reader what you found and how you found it. Make sure you include important results in your report and put some necessary outputs in appendix.

(c) (10 pts) Consider two experiments:

Experiment 1: a randomized complete block design with 3 treatments and 3 blocks, and

Experiment 2: a 2-way factorial design where factor A has 3 levels and factor B has 3 levels

What would you do differently in assigning experimental units for these two designs?