UNM Statistics Qualifying Exam Take-Home
Spring 2008

Due Friday January 18, 2008. Return to Donna George in 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. An appendix is allowed for each problem but will be examined only at the discretion of the graders. The better organized your appendix is, the more likely it is to get examined.

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 Curt Storlie, `storlie@stat.unm.edu`.

1. The data for this problem is from a random sample of records of resales of homes in Albuquerque from Feb 15 to Apr 30, 1993 from the files maintained by the Albuquerque Board of Realtors. This type of data is collected by multiple listing agencies in many cities and is used by realtors as an information base. There are 117 cases on eight variables. The data is available for download at

   `http://www.stat.unm.edu/~storlie/abq_homes_data.txt`

   The data set contains the variables:

   Price - Selling price in hundreds of dollars.

   SQFT - Square feet of living space

   Age - Age of home (years)

   Features - Number out of 11 features (dishwasher, refrigerator, microwave, disposer, washer, intercom, skylight(s), compactor, dryer, handicap fit, cable TV access)

   NE - Located in northeast sector of city (1) or not (0)

   CUST - Custom built home (1) or not (0)

   Corner - Corner location (1) or not (0)

   Taxes - Annual taxes in dollars

   There are several missing values denoted by a "∗" in the data file. Place yourself back in May of 1993. At that time TJ Reiter was an up and coming realtor in Albuquerque interested in using this data to help him price a house that he was selling. The house in question had the following values of the covariates ($SQFT = 2400$, $Age = 1$, $Features = 6$, $NE = 0$, $CUST = 1$, $Corner = 0$, $Tax = 1520$).

   Please help TJ by building a model that best describes the relationship between *Price* and the possible covariates. Use this model to give him any information he may find useful. Make sure that you carefully assess all assumptions and write a succinct, coherent, and complete summary of your analysis.

2. A biologist performed an experiment to assess the weight gain for rats fed on four different diets: combinations of two protein sources and two protein amounts. The design is a completely randomized design with ten rats randomly allocated to each of the four treatments. There are forty observations on the following four variables:

PreWt - The weight before the experiment (grams).

PostWt - The weight after the experiment (grams).

Protein - The protein source: either Beef or Cereal.

Amount - The amount of protein: either High or Low.

The data is available for download at

`http://www.stat.unm.edu/~storlie/rat_data.txt`

Build a model that best describes the relationship between the weight gain (as a percentage of intial weight) and the factors *Protein* and *Amount*. Make sure that you carefully assess all assumptions and write a succinct, coherent, and complete summary of your analysis.