Nonparametric Statistics Takehome Exam II.

Due date: November 26 2:30pm. 1%/hour deduction afterward. Worth 20% of your final grade.

Submission guideline: For each problem, you need to write a report about one page long detailing your hypotheses, assumptions, statistics and conclusions, etc. Justify every assumption and step you are taking. Each problem is worth about 5 % = 4 % (correctness)+ 1% (presentation). Partial grade given at 1% increment. For exceptional work, additional bonus points will be given.

1. Cuckoos are known to lay their eggs in the nests of other birds. The eggs are then adopted and hatched by the host birds. Is there any relationship between the length of eggs and the host birds? What do you conclude? Measurements are in millimeters. The data file is at http://www.stat.wisc.edu/∼mchung/teaching/data/cuckoo.data.

2. In 1970, Congress instituted a random selection process for the military draft. All 366 possible birth dates were placed in plastic capsules in a rotating drum and were selected one by one. The first date drawn from the drum received draft number one and eligible men born on that date were drafted first. Is the draft lottery fair? If there any relationship between the birth date and the draft number, what would that be? The data file is at http://www.stat.wisc.edu/∼mchung/teaching/data/draft.data.

3. Henry Cavendish in 1798 measured the density of the earth and reported it as a multiple of the density of water. Estimate the probability density of the measurements. What is the expected density of the earth? Test if the the density of earth is 5.46. What is the P-value? Use *kernel methods* for all your computations. The data file is at http://www.stat.wisc.edu/∼mchung/teaching/data/earth.data.

4. Let Y_j be the j-th observation at time t_j. We have kernel smoothing curve given by $\mu(t) = \sum_{j=1}^n W_j(t)Y_j$, where the weights $W_j(t) = K(t - t_j)/\sum_{i=1}^n K(t - t_i)$ and $K(t)$ is a kernel. Show that $\mu(t_i)$ will lie between the minimum and the maximum observations for all i, i.e. $Y_{(1)} \le \mu(t_i) \le Y_{(n)}$ for all i (Hint: $\sum_{j=1}^n W_j(t) = 1$ for all t. Kernel is a symmetric nonnegative probability density function).