STA 290 — Homework 6

Please include all of your C code in an appendix.

1. Suppose we are interested in a random variable with inverse CDF

$$F^{-1}(u) = \begin{cases} 0 & u < 0 \\ u^4 & 0 \leq u \leq 1 \\ 1 & u > 1 \end{cases}$$

   Estimate the mean and variance of this random variable by simulating a million draws and computing the sample mean and sample variance.

2. For the data on mercury levels in fish, suppose you just want to see if there is a difference in mercury level for the different stations at which the fish were caught. If this were the only variable of interest, we would do an analysis of variance.

   (a) Fit an ANOVA in S-Plus to produce the analysis of variance table.

   (b) Recreate this table in C (except for the p-value, which is more difficult to get in C).

3. In professional boxing, each judge scores each round with the winner getting ten points and the loser typically getting nine points, although there are some circumstances where this could be less or the round could be scored a tie. Championship bouts last for twelve rounds if there is no knockout. The data in this problem comes from the first heavyweight title fight between Evander Holyfield and Lennox Lewis in 1999. Most people watching the fight thought that Lewis had clearly won, but the fight was ruled a draw because of the three official judges, one scored it for Holyfield, one for Lewis, and one as a draw. There was subsequently much speculation of corruption and accusations of misjudging on the part of judge Eugenia Williams, who had scored the fight for Holyfield. The data here consists of the round-by-round scores of the three official judges as well as the scores assigned by seven "experts" in the media (e.g. the official commentator of the TV broadcast, several newspaper reviewers, etc.). All rounds were scored as either 10-9 or 10-10. The actual data are:

	Individual Round Scores												
	1	2	3	4	5	6	7	8	9	10	11	12	**Total**
Official Judges													
Williams	L	L	H	H	H	L	L	H	H	H	H	L	115-113
O'Connell	L	L	H	L	L	H	E	H	H	E	H	L	115-115
Christodoulu	L	L	H	L	L	L	L	H	H	H	E	L	113-116
Media & Unofficial Scorecards													
HBO-Lederman	L	L	H	L	L	L	L	L	H	H	L	L	111-117
Sportsticker	L	L	H	L	L	L	L	L	H	H	L	L	111-117
Boxing Time	L	L	H	L	L	L	L	L	H	H	L	L	111-117
Sportsline	L	L	H	L	L	H	L	L	L	H	L	L	111-117
AP	L	L	H	H	L	L	L	L	L	H	L	L	111-117
ESPN	L	L	L	L	L	L	L	L	L	H	L	L	109-119
Glyn Leach	L	L	H	E	L	E	L	H	H	H	H	L	115-115

   where L denotes a round scored for Lewis, H a round for Holyfield, and E a tied (even) round.

   The statistical problem is to see if the scorecard of judge Williams looks extreme relative to the other scorecards. Thus we have a problem in inter-rater agreement, although we don't have the standard setup of a small number of judges and a large number of samples. To simplify things, we treat all even rounds as rounds scored for Holyfield (since Lewis was widely viewed as having "won"), resulting in binomial data. There are a number of possible models that we could fit. Perhaps the simplest is to assume that there is no difference between the judges, i.e., the scores given by each judge for a particular round are independent bernoulli observations with a common probability (which may differ from round to round). We can take this as a null hypothesis, and compute a p-value like quantity by estimating the probability of observing a scorecard with at least as many rounds for Holyfield (seven) as were on Williams' scorecard.

To be more precise, the model is as follows. Let $x_{ij} = 1$ if judge j scores round i for Holyfield, and $x_{ij} = 0$ if it is scored for Lewis. Let p_i be the probability that a judge scores round i for Holyfield. Under the null hypothesis, $x_{ij} \sim Bern(p_i)$. We can estimate p_i with the MLE $\hat{p}_i = \frac{1}{J}\sum_{j=1}^{n} x_{ij}$ where J is the total number of judges (here ten). We then estimate the probability of observing a scorecard at least as extreme as Williams' by computing the probability of observing a scorecard with at least seven rounds for Holyfield, using $\hat{p}_1, \ldots, \hat{p}_12$. We can do this with an exhaustive enumeration – for each possible scorecard of 12 bernoulli variables, if the scorecard has at least 7 rounds for Holyfield, compute its probability and add all of these probabilities. This is where C is more efficient than S – you can quickly loop over all 2^{12} possible scorecards.

A data file is available on the course web page with the data already coded as 0 and 1. Since you don't actually need the names of the judges for this analysis, you might find it easier to remove that column from the data file before you try to read it in to C. Estimate the p-value as described above and report your findings.

If you are interested in more details on the problem, take a look at a paper I have on it `http://www.stat.duke.edu/courses/Fall00/sta290/Homeworks/boxingpaper.pdf`.