STAT 621 HOMEWORK 3
Due: Wednesday Sept. 25

1. Complete Problem 7 from last week's homework (Crossvalidation).

2. (This is a real example – from Hollander, Wolfe and Chicken) A 1989 study by Molitor aimed to see if children who watched violent TV were significantly more tolerant of real-life violent behavior than children who watched nonviolent TV. Half of the 42 children in the study were shown violent TV (an edited version of the Karate Kid), and the other half watched exciting but nonviolent footage of the 1984 Olympic Games. Afterwards, each child was asked to supervise two younger children via television monitor. The child supervisers were told to go find the research assistant if the younger children got into trouble. What each child witnessed while alone was actually a videotaped sequence showing two young children at play, that got progressively more violent. That is, the two children called each other names, pushed each other, chased each other, fought, and then supposedly broke the video camera while fighting.

   The time (sec) that each child remained in the room after the first act of violence was recorded. Data are saved in the file `violent.txt` on Blackboard.

   (a) Use the Wilcoxon-Mann-Whitney test to determine whether children who view violent TV take longer to seek help (i.e., are more tolerant) than children who view nonviolent TV. State your hypotheses, report both the values of W and U, give the p-value and conclusion. Use $\alpha = .05$.

   (b) Test the same hypothesis using the large sample approximation. Compare your results.

3. Hodges and Lehman proposed the following estimator of Δ, the parameter describing the shift in the distributions of the the two populations compared with the Wilcoxon-Mann-Whitney statistic.

$$\widehat{\Delta} = \text{median}\{(Y_j - X_i), i = 1, \ldots m; j = 1, \ldots, n\}$$

   That is, $\widehat{\Delta}$ is the median of all possible pairwise differences between the Ys and Xs.

   The data below come from a study of wing length (mm) in migratory and nonmigratory dark-eyed juncos.

Migratory (X)	84.5	81.0	82.6	80.5	82.1	83.4	79.7
Nonmigratory (Y)	82.4	83.9	86.3	86.6	87.8	84.1	

   (a) Compute the estimate $\widehat{\Delta}$.

   (b) Use the bootstrap to approximate the sampling distribution of $\widehat{\Delta}$. Make a histogram, report the mean and standard deviation.

(c) Report a 95% bootstrap confidence interval for Δ and give an interpretation. Does your interval suggest an overall difference in wing length?

4. Consider the random sample of 4 Xs and 3 Ys shown below.

```
x=c(2.1, 1.9, 2.6, 3.3)
y=c(1.9, 2.6, 3.7)
```

(a) Compute the observed value of W, the Wilcoxon statistic. Note there are ties here: substitute average ranks for the ranks of tied observations.

(b) Given these data, find the distribution of W under the hypothesis that there is no difference in the X and Y distributions, $\Delta = 0$. Note: this can get tedious; the `combn` function R may help.

(c) How extreme is the observed value of W? Report the p-value for testing the alternaive that $\Delta > 0$.

5. A study investigated the use of three species of pine trees for habitat modification in the Alum Creek Lake in Westerville, Ohio. The distances between branches on the tree (interstitial lengths) are thought to affect fish behavior, feeding and spawning. Average interstitial length (mm) was recorded for randomly selected branches on 12 each of scotch pine, blue spruce and white pine trees. Data are saved on Blackboard as `pines.txt`.

(a) Consider the Kruskal-Wallace procedure. Evaluate the data with regard to the assumptions of this procedure. Compute any relevant summary statistics or plots. What about using a standard ANOVA model?

(b) Use the KW the procedure to test whether there are differences in median interstitial lengths among the three tree species.

6. Pretherapy training of clients has been shown to beneficial for new psychotherapy patients. A study (Sauber 1971) investigated four approaches to pretherapy training: Control, Reading, Video, and Group Interaction. Nine patients were randomly assigned to each pretherapy treatment, and scored on their "psychotherapeutic attraction" (higher scores suggest better outcomes). Data are on Blackboard as `therapy.txt`. Use the KW test to determine if there is a difference in median scores among the treatment groups. Use a large sample approximation with correction for ties discussed in class.