STAT 740 - Spring 2004 - Take Home 1

Due: 5:00 pm Friday March 5th

- Do not consult with anyone except me on these problems.
- Adequate comments and a modicum of efficiency are required for a perfect score.
- E-mail the code that is needed so that I can **exactly** reproduce your results.

1) Consider for example the harmonic series

$$\sum_{i=1}^{\infty}\frac{1}{i}=1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{i}+\cdots$$

Write a Fortran program to estimate the sum of the harmonic series as a single precision floating point real number by beginning with the largest number (1) and then adding the progressively smaller numbers (1/2, 1/3, etc...) in order. Have the program determine what the maximum value of i is that causes the estimated sum to change. Also have the program add the numbers in reverse order (from smallest to largest) beginning at $1/i_{max}$ and again beginning at $1/(2i_{max})$. Briefly discuss why each of the three results is larger or smaller than the other two, and compare their results to the true sum of the harmonic series. Can you think of some other method of adding the numbers between 1 and $1/i_{max}$? (You do not need to program this last method.)

2) Consider the distribution with probability density function:

$f(x) = x$ for $0 \le x \le 1$ and $f(x)=2-x$ for $1 \le x \le 2$ and 0 elsewhere

Briefly describe an algorithm that could be used to generate random numbers from this pdf. Program this algorithm in either R or Fortran using no other random number generators except one for a single uniform(0,1) (e.g. `runif(1,0,1)` in R or on `ran2` in Fortran).

3) In class we have seen the R functions for making a bootstrap confidence interval for the median using the normal method, basic method, and BCa method.

Modify this code to create a single R function that will construct bootstrap confidence intervals for the standard deviation using the normal method, basic method, percentile method, and BCa method using the same set of bootstrap samples.

Run a simulation study to see how the effectiveness of these four methods change with B, using B= 100, 250, 500, and 1000. For each case use 100 simulation runs and samples of size 10 from a Gamma distribution with mean 6 and variance 12. Instead of using inferential statistics construct a graphical display to demonstrate your findings and briefly summarize it in a paragraph.

4) This problem continues the final problem from homework 3. Use good experimental design to set up a simulation study to determine how the Wilcoxon signed rank test and the t-test perform in terms of type I error rate (nominal α=0.05) for data like yours (using 1,000 simulation runs). Briefly summarize your findings, including: the use of the appropriate statistical test of the null hypothesis that the two procedures have the same obtained significance levels for data like yours; and, confidence intervals for the individual estimates of the type I error rates.