

STAT 672, Spring 2011
Homework 9
due Thurs. April 28

Show all work. 2 Problems.

The numbers refers to *Hollander and Wolfe*, if not specified otherwise. Some problems may have additional parts.

1. In Homework 5, p. 131, 4.15 you constructed a 95% CI for Δ . Now construct a 95% CI for the differences in the population medians using the nonparametric bootstrap. Use the R `bootstrap` function to a construct 95% CI based on percentiles and $B = 1000$ bootstrap replicates. It is fine to call the `bootstrap` function two times. Make a histogram of the 1000 bootstrapped differences in the medians. Compare your results to Homework 5.

The data is available off the class web page at:

```
http://www.rohan.sdsu.edu/~babailey/stat672/t4-3.txt
```

2. Consider the R dataset `faithful`. We will look at the eruption times in minutes.

(a) Make a probability histogram of the eruption data to estimate the density. By default, how does the R function `hist` determine the number of bins? How many bins did it choose? Describe the characteristics of the density.

(b) The R function `density` is useful for nonparametric density estimation. If we use the commands:

```
> attach(faithful)
> temp <- density(eruptions)
> plot(temp)
```

You will have to be a detective for the following parts. (The help file is useful!) Lab4 is also available, if you need more help.

(i) What is in the R object `temp`?

(ii) What is the default kernel used by `density` ?

(iii) How is the bandwidth selected?

(iv) There is a file at:

```
http://www.rohan.sdsu.edu/~babailey/stat672/geyser.r
```

Source the file or copy and paste the lines of code in R. What plot is produced?