Midterm 1 2007

<Your Name>

<date when you took exam>

1 Instructions

Produce a LaTeX document formatted like this one. Include this paragraph as it appears here.

You do not have to copy the rest of these instructions or any of the questions into your document. Just make the sections and subsections with the same names that I have here. Use Sweave to insert your code and output.

Submit your exam by uploading your .Rnw file and your final PDF file into ICON.

2 More Joy of LaTeX

2.1 Math

Typeset the following matrix in LaTeX:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

2.2 Tables

Typeset the following table with caption:

Name	Units	Price
Heather	3	3.75
Roger	12	15.00
Pham	5	6.25

Table 1: Purchases

3 R functions

Recall that the *harmonic mean* of a set of strictly positive values is the number of values divided by the sum of the reciprocals of the variables. That is, the harmonic mean of a set of positive values y_1, y_2, ...,, y_n is

$$\frac{n}{\frac{1}{y_1} + \frac{1}{y_2} + \ldots + \frac{1}{y_n}}$$

Write an R function that will accept a numeric vector, test that all elements in the vector are positive, and if so, return the harmonic mean of the values in the vector.

The dataset called `rivers` is built into R. To get a description of the dataset, enter

```
> help(rivers)
```

Use your function to calculate the harmonic mean of the data in this dataset.

Include in your exam document the code for your function, your command line to apply your function to the `rivers` data, and the resulting output.

4 The bootstrap

Now suppose that the `rivers` data could be considered a sample of major rivers from around the world. Thus the value you calculated in the previous question could be considered a point estimate of the harmonic mean of the population of lengths of all major rivers.

Use the `boot` function in the `boot` package to estimate:

- the bias in your point estimate
- the standard error of your point estimate
- a 95% percentile-method confidence interval for the population geometric mean

Use at least 500 replicate datasets.

Include in your document any R code and output that you obtain. Write a sentence or two stating whether bias correction is needed in this problem.