Midterm 1 2007

<Solutions>

<Oct. 1, 2007>

1 Instructions

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

2 More Joy of LaTeX

2.1 Math

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

2.2 Tables

Name	Score	Percent
Heather	42	84
Roger	39	78
Pham	50	100

Table 1: Physical fitness test results

3 R functions

Recall that the *geometric mean* of a set of strictly positive values is obtained by exponentiating the average of the natural logs of all the values. That is, the geometric mean of a set of positive values y_1, y_2, ...,, y_n is

$$exp\left(\frac{\sum log(y_i)}{n}\right)$$

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

```
> geommean <- function(x) {
+     if (!is.numeric(x) || any(x < 0))
+         stop("need vector of all positive numbers")
+     else exp(mean(log(x)))
+ }
> geommean(rivers)

[1] 481.0056
```

4 The bootstrap

```
> geommean2 <- function(x, ind) {
+     exp(mean(log(x[ind])))
+ }
> library(boot)
> boot.river <- boot(rivers, geommean2, 1000)
> boot.river

ORDINARY NONPARAMETRIC BOOTSTRAP


Call:
boot(data = rivers, statistic = geommean2, R = 1000)


Bootstrap Statistics :
    original     bias    std. error
t1* 481.0056 -0.5351836    23.57526

> boot.ci(boot.river)

BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 1000 bootstrap replicates

CALL :
```

```
boot.ci(boot.out = boot.river)

Intervals :
Level      Normal              Basic
95%   (435.3, 527.7 )   (434.2, 527.7 )

Level     Percentile            BCa
95%   (434.3, 527.8 )   (435.5, 529.2 )
Calculations and Intervals on Original Scale
```

The bootstrap analysis suggests that the point estimate is almost unbiased, in which case bias-correction would not be appropriate.