

HW5

2024-09-29

Homework 5

Due October 15

Q1

Q:

Use the `state.x77` data matrix and the `tapply()`, and separately, the `aggregate()` function to obtain

a.

The mean per capita income of the states in each of the four regions defined by the factor `state.region`,

b.

The maximum illiteracy rates for states in each of the nine divisions defined by the factor `state.division`,

c.

The number of states in each region,

d.

The names of the states in each division,

e.

The median high school graduation rates for groups of states defined by combinations of the factors `state.region` and `state.size`.

A:

a.

b.

c.

d.

e.

Q2

Q:

For a sample x_1, x_2, \dots, x_n , the MAD estimator of scale is defined as $1.4826 \cdot \text{median}\{|x_i - \bar{x}|\}$ where $\bar{x} = \text{median}\{x_i\}$

Use the matrix object `mtcars` to compute the MAD estimator of scale for the columns in `mtcars`

a.

Using the `apply()` function with the `mad()` function

b.

By calculating it directly from the definition (i.e., not using the `mad()` function). You may use the `apply()` and the `sweep()` functions but avoid using any loops

A:

a.

b.

Q3

Q:

Let $h(x, n) = 1 + x + x^2 + \dots + x^n = \sum_{i=0}^n x^i$ Answer the following questions

a.

Write code to do the above in a for loop.

b.

Rewrite the code that you wrote to use a while loop.

c.

For each of the $x = 0.3, 1.01$ evaluate the performance for $n = 500, 5000$ in terms of time taken by the software. To do so, wrap the code around the R function `system.time()` with the same code as above inside the parentheses. Report the values returned in the output as per the user time field.

d.

Compare the above with results obtained avoiding loops.

A:

a.

b.

c.

d.

Q4

Q:

The Lotka-Volterra model for a predator-prey system assumes that $x(t)$ is the number of prey animals at the start of year t and that $y(t)$ is the number of predators at the start of year t . Then the number of prey animals and predators at the end of the following year is given by:

$$x(t+1) = x(t) + b_x x(t) - d_x x(t)y(t)$$

$$y(t+1) = y(t) + b_y d_x x(t)y(t) - d_y y(t)$$

Where b_x , b_y , d_x and d_y are as follows: - b_x is the natural birth rate of the prey animals in the absence of predation - d_x is the death rate of prey animal in an encounter with the predator. - d_y is the natural death rate of the predators in the absence of food (prey animals) - b_y is the efficiency of turning predated animals into predators

Let $b_x = 0.04$, $d_x = 0.0005$, $b_y = 0.1$, and $d_y = 0.2$.

Suppose that there were 4000 animal of the prey variety at the beginning of the time period. Also, suppose that there were only 100 predators. Write a while loop to show the predator-prey system as long as there are over 3900 prey animals. Save the prey/predator output in a list and plot them using lines on the same plot.

A:

Q5

Q:

The game of craps is played as follows: first, Player 1 rolls two six-sided die; let x be the sum of the die on the first roll. If $x = 7$ or $x = 11$, then Player 1 wins, otherwise the player continues rolling until (s)he gets x again, in which case also Player 1 wins, or until (s)he gets 7 or 11, in which case (s)he loses. Write R code to simulate the game of craps. You can simulate the roll of a fair die using the `sample()` function in R.

A:

Q6

Q:

Suppose that $(x(t), y(t))$ has polar coordinates given by $(\sqrt{t}, 2\pi t)$. Write code to plot the curve $(x(t), y(t))$ for $t \in [0, 1]$.

A:

Q7

Q:

Consider the following code:

```
x <- matrix(rnorm(n = 500), ncol = 5)
varx <- var(x)
```

Starting with varx, use two applications of the sweep() function, one dividing each row of the matrix and the other dividing each column, of a covariance matrix to obtain R, the correlation matrix.

A: