STAT 210
Applied Statistics and Data Analysis
Problem List 1 (due on week 2)

Fall 2022

Exercise 1

Using the functions `rep` and `seq`, generate the following sequences

Text

1. 10 10 10 10 10 9 9 9 9 8 8 8 7 7 6 5 4 4 3 3 3 2 2 2 2 1 1 1 1 1
2. 1 1 2 3 3 4 5 5 6 7 7 8 9 9 10
3. 100.0000 100.2222 100.4444 100.6667 100.8889 101.1111 101.3333 101.5556 101.7778 102.0000
4. 1.0 1.0 1.0 1.2 1.4 1.4 1.4 1.6 1.8 1.8 1.8 2.0
5. 1 2 3 4 5 2 3 4 5 6 3 4 5 6 7 4 5 6 7 8 5 6 7 8 9

Exercise 2

Use the Montecarlo method for estimating π.

Exercise 3

We will use the data set `mtcars`, that has information regarding fuel consumption and 10 related variables for 32 different car models.

1. Use the function `str` to explore the data set.
2. Using the function `subset`, create a new file named `file1` containing the variables `mpg, hp` and `wt`, but only for cars with 6 cylinders or more.
3. Using the functions `apply` and `mean`, calculate the mean value for each of the three variables in `file1`.
4. Using the function `sweep`, create a new object called `file2` with the data in `file1` after subtracting the means for each variable.
5. Using the function `within` create a new column in `file2` containing a new variable called `par1` calculated as `par1` = 1.4*`hp` - 0.5*`wt`.

Exercise 4

(a) Create a vector named `smpl1` with a sample of size 100 from the set of categories 'bad', 'reg', 'norm', 'good', 'exc'. The categories 'bad' and 'exc' should have probability 0.1, 'reg' and 'good' should have probability 0.2, and 'norm', probability 0.4.

(b) Create a factor named `fact1` using the vector `smpl1` as input.

(c) Create an ordered factor named `fact2` using the vector `smpl1` as input. The levels should be in increasing order.

(d) Now, you want to reduce the number of categories to three: 'bad' and 'reg' will now be 'poor', 'norm' will be 'normal' and 'good' and 'exc' will now be 'great'. One easy way to do this is to use the `labels` argument in the function `factor` to rename the levels. Look up the help page for `factor`; there is an example that will show you how to do this. Name the resulting ordered factor `fact3`.

(e) Use the function `table` to create tables for the three factors you have made. Comment on the differences.