Exercises

Exercise 1

Step 1: Find out what does the R function **runif()** do. (Hint: try "?runif")

Step 2: Run the codes below, figure out what it does and get a sense of how to do subsetting in R. (When you don't understand one specific line or parameter, think about the previous hint.)

```r
A <- matrix(runif(20), ncol=4);
A[4,3]
A[c(1,3),c(2,3)]

A[2,] <- c(1,2,3,4);
A <- replicate(4,A[,2])
```

Step 3: Which of the following will give you all elements that are larger than 0.5 in A?

1. A>0.5
2. A[A[,2]>0.5,]
3. A[A>0.5]

Exercise 2

Write a program which generates a 10-element vector of random binary values(0 and 1) such that each element has the same probability of being 1 or 0. (Bernoulli distribution, p = 0.5) (Hint: try **runif()**)

Exercise 3

Write a function, *my_bernoulli()*, to do the same thing as in the previous exercise. The function should take two inputs: p(probability of 1, default value is 0.5) and n(length of output vector).

Exercise 4

First download the data from the Internet. Then change the working directory to where you place the file. Use the function **read.csv()** to read data locally. If you feel unfamiliar with the function, use "?read.csv" to get help.

Exercise 5

First, dig into the function **paste()**. Then run the following code.

```r
a <- 123;
b <- "Hello, world!";
c <- as.character(a);
char <- paste(c, b, sep = "_");
```

Substitute all numbers in the variable *char* with asterisks.

Exercise 6

Follow the steps below to plot the total precipitation for March, April and May.

1. We already have all the data we need in **total_pre**. Add a new column to the variable to indicate the month information. (Hint: you can obtain the month information by using the **format()** function, like format(total_pre$date[1], "%m"). The **sapply()** function will be very useful here.)

2. Get the precipitation data for March, April and May. You can either assign these to a new variable or still use the original variable with subsetting.

3. First plot March data. Then use function **lines()** to add lines for both April and May. (Hint: Since there are different numbers of days in these months, when you first call the **plot** function, you need to specify the parameter *xlim*). (This step is a bit hard. Be sure to read the help files for these functions when you are stuck.)