Module 2: Exercises

Exercise I (Basics)

a) Make the following three vectors without using the `c()` command: $x = (1, 1, 1, 1, 1)$, $y = (1, 3, 5, 7, 9)$, $z = (2, 2^2, 2^3, 2^4, 2^5)$ (Hint: `seq()` and `rep()` may be useful – check the help.)
b) Make a matrix `X` with columns `x`, `y`, and `z`.
c) Try to add two vectors and/or matrices that do not match in dimensions, and see if you can figure out what R does.

Exercise II (Distributions)

a) Plot the density function for the beta distribution for a few different values of the two shape parameters (note the support of the density in the help file).
b) Generate 100 realizations from a beta distribution, and make a histogram. Add the theoretical density to the histogram. (Hint: the `curve` command has an argument `add=TRUE` that allows you to add a plot on top of the histogram; note that this is only useful if the histogram is normalized to integrate to one, which can be achieved by including the argument `probability=TRUE` to the `hist` command)
c) Calculate the average of the 100 realizations. Can you guess what the theoretical mean is for your parameter values (feel free to repeat the experiment and/or increase the number of realizations)? Using other parameter values can you guess the general formula for the mean in terms of the parameters (without looking it up somewhere)?

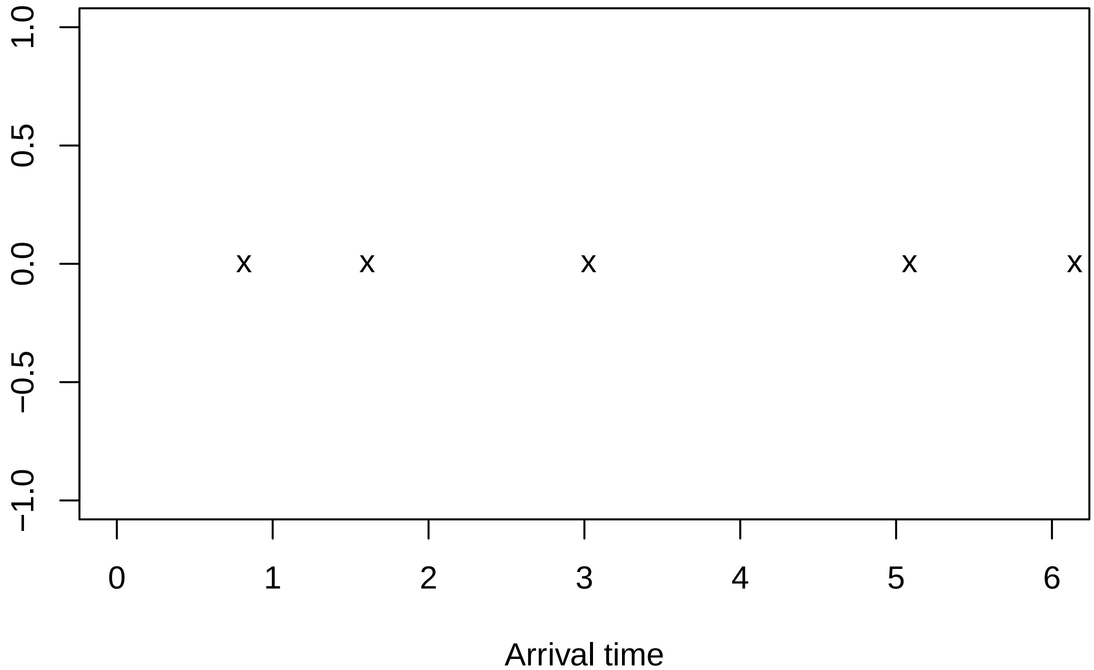
Exercise III (Functions and loops)

a) Make a function with a for loop that can calculate the product of all the entries of an input vector. Compare with the built-in function `prod` (don't call your function `prod`, or you won't able to use the built-in function easily).
b) Make a function that will calculate the Fibonacci numbers up to n (an input parameter). Does it handle n=1 and 2 correctly? (Hint: An `if` statement may be useful here.)

Exercise IV (Arrival times)

Consider a process of arrivals on the real line where the inter-arrival time is exponentially distributed with rate 1. E.g.

```r
set.seed(54321) # For reproducibility
t1 <- rexp(1)
t2 <- t1 + rexp(1)
t3 <- t2 + rexp(1)
t4 <- t3 + rexp(1)
t5 <- t4 + rexp(1)
plot(c(t1, t2, t3, t4, t5), rep(0, 5), pch = "x",
     xlim = c(0,6), ylab = "", xlab = "Arrival time")
```



- Make a function f that sequentially generates inter-arrival times and counts the number of arrivals in the interval $[0, T]$ where T is the only parameter of the function. E.g. for the example above $f(1) = 1$, $f(2) = f(3) = 2$, $f(4) = f(5) = 3$, $f(6) = 4$. (Hint: Either use `while`-loop instead of `for`-loop or simply use a conservatively long `for`-loog – maybe `1:100` – and break out of the loop once T is exceeded.)
- Generate 1000 realizations of the random number of arrivals in $[0, 5]$. Calculate the empirical mean and variance. Compare with a Poisson distributed random variable with rate parameter 5.

Exercise V (Uniform distribution)

- Make a short report in Rmarkdown about the uniform distribution on $[A, B]$ where $A < B$ should be variables definied in the very beginning of the document. The report should at least include:
  - A histogram of a large sample from the distribution.
  - A plot of the density (preferably overlayed on the histogram).
  - The difference between the sample mean and the theoretical mean as well as the difference between the sample variance and the theoretical variance. (Hint: To calculate the theoretical mean and variance think about the relation between Unif(A,B) and Unif(0,1) and use the results from module 1.)
- Rerun your report with a different choice of A, B and check that everything is still correct.