General Notes

- You will submit a minimum of two files, the core files must conform to the following naming conventions (including capitalization and underscores). 123456789 is a placeholder, please replace these nine digits with your nine-digit Bruin ID. The files you must submit are:

  1. *123456789_stats102c_hw1.Rmd*: Your markdown file which generates the output file of your submission.
  2. *123456789_stats102c_hw1.html/pdf*: Your output file, either a PDF or an HTML file depending on the output you choose to generate.

  If you fail to submit any of the required core files you will receive ZERO points for the assignment. If you submit any files which do not conform to the specified naming convention, you will receive (at most) half credit for the assignment.

- **Your .Rmd file must knit**. If your .Rmd file does not knit you will receive (at most) half credit for the assignment.
  The two most common reason files fail to knit are because of workspace/directory structure issues and missing include files. To remedy the first, ensure all of the file paths in your document are relative paths pointing at the current working directory. To remedy the second, simply make sure you upload any and all files you source or include in your .Rmd file.

- Your coding should adhere to the tidyverse style guide: https://style.tidyverse.org/.

- Any functions/classes you write should have the corresponding comments as the following format.

```
my_function <- function(x, y, ...){
#A short description of the function
#Args:
#x: Variable type and dimension
#y: Variable type and dimension
#Return:
#Variable type and dimension
Your codes begin here
}
```

NOTE: *Everything* you need to do this assignment is here, in your class notes, or was covered in discussion or lecture.

- Please **DO NOT** look for solutions online.
- Please **DO NOT** collaborate with anyone inside (or outside) of this class.
- Please work **INDEPENDENTLY** on this assignment.
- **EVERYTHING** you submit **MUST** be 100% your, original, work. Any student suspected of plagiarizing, in whole or in part, any portion of this assignment, will be **immediately** referred to the Dean of Student's office without warning.

Problem 1: Suppose that X is a discrete random variable with probability mass function:

x	0	1	2	3	4
p(x)	0.1	0.15	0.2	0.25	0.3

(a) Write a function using the inverse transform method to generate random numbers from the distribution of X.

(b) Generate 10,000 random numbers and draw a bar chart.

(c) Compare the sample relative frequencies with the theoretical probability distribution.

Problem 2: Please write a function using the inverse cdf method to generate Poisson random numbers.

(a) Design an algorithm using the inverse cdf method.

(b) Implement your algorithm in R.

(c) Generate 10,000 random numbers with $\lambda = 3$ and compare your results with **rpois**'s.

Problem 3: A cumulative distribution function of X is given as following

$$F(x) = 1 - e^{-(x/\alpha)^\beta},\ x \geq 0,\ \alpha > 0,\ \beta > 0$$

(a) Please show that $Y = (\frac{X}{\alpha})^\beta$ follows an exponential distribution.

(b) Write a function to generate 100,000 random numbers with $\alpha = 2$ and $\beta = 4$, and plot the histogram.

Problem 4: For the acceptance-rejection method, please prove that the returned random sample from the target density f(x).

Problem 5: Write a function to generate random variables from the Beta(α, β) distribution using the acceptance-rejection method. You may use Unif(0,1) as the envelop distribution. You may set $\alpha = 3$ and $\beta = 2$.

(a) Calculate M before coding.

(b) Generate a random sample of size 100,000, and plot the histogram.

(c) Compute the empirical acceptance rate, and compare it with $1/M$.

Problem 6: The standard Laplace density is

$$f(x) = \frac{1}{2}e^{-|x|},\ x \in \mathbb{R}$$

(a) Design an algorithm to generate 100,000 random variables using the inverse CDF method and implement it in R.

(b) Design an algorithm to generate 100,000 random variables using the rejection method and implement it in R. You may use Normal(0, 3) as the envelop distribution.

(c) Compare your results of (a) and (b). Discuss their advantages and disadvantages.