Assignment 1

In this assignment you are required to write 3 functions:

Q1: Application of Knuth's algorithm for generating pseudo random numbers.

Q2: Envelope rejection sampling to generate random numbers from the Gamma distribution

Q3: Ratio-of-uniforms method to generate random numbers from the Normal distribution.

The template for the functions in given in the file "Assignment_1_123456789.R". You should change (only) the number "123456789" in the file name to your ID number. Do not make any other changes in the name of the file.

In each one of the functions that are in the file you are required to replace the place-holder "return(NA)" by the body of the code of the function that you write. You must not change the name of the function nor the arguments of the function. Do not add any comments to the code and do not add any code outside the body of the functions. Unless otherwise stated, your code cannot assume availability of or call any package besides the standard packages in the base distribution of R that are uploaded automatically when an R session opened.

Question 1 (Q1):

Knuth's algorithm for generating pseudo random numbers applies the recursion formula:

$$X_n = \left(X_{n-j} + X_{n-k}\right) \bmod m$$

for j, k, m three integers, with $j < k$. The initiation of the recursion requires a sequence $X_{-k+1}, X_{-k+2}, \ldots, X_0$ that contains k integers.

Write a function by the name "a1q1" that produces a sequence of pseudo random numbers according to Knuth's algorithm in two steps.

In the first step the initializing sequence of length k is produced by the linear congruential generator with $X_0 = \text{seed}$ and:

$$X_{-(i+1)} = (1103515245 \times X_{-i} + 12345) \bmod 2^{32}, \qquad 1 \leq \mathrm{i} \leq \mathrm{k} - 1.$$

In the second step a sequence of length n of integers is produced according to Knuth's algorithm.

The arguments of the function are:

n = The length of the sequence.

j = The first increment.

k = The second increment.

m = The modulus.

seed =The seed (a single integer).

The output of the function should be a vector of length n that contains only the numbers that were produced in the second step.

Question 2 (Q2):

Consider the Gamma distribution of unit rate, which has a density given by:

$$f(x) = \frac{x^{r-1}e^{-x}}{\Gamma(r)}, \qquad x > 0.$$

In this question you should apply the envelope rejection sampling algorithm in order to generate random numbers from the Gamma distribution for a non-integer $r > 1$.

Let $m = \lfloor r \rfloor$ be the integer part of r and define $\lambda = m/r$. Generate proposals from the density:

$$g(x) = \frac{\lambda^m x^{m-1} e^{-\lambda x}}{\Gamma(m)}, \qquad x > 0.$$

Observe that this is the distribution of the sum of m independent exponential random variables with rate λ. You may use the fact that:

$$\frac{f(x)}{g(x)} = \frac{\Gamma(m)}{\Gamma(r)} \lambda^{-m} x^{r-m} e^{-(1-\lambda)x} \leq \frac{\Gamma(m)}{\Gamma(r)} \lambda^{-m} \left(\frac{r}{e}\right)^{r-m},$$

for all $x > 0$.

Write a function by the name “a1q2” that outputs the entire sequence of proposals that were generated from density g and the sequence of random numbers from the distribution with density f.

The arguments of the function are:

n = The length of the sequence from the distribution f.

shape = The value of the shape parameter r.

The function should return a data frame with two columns and n rows. The names of the columns should be:

X = The accepted proposals.

U = The uniform random variables that are associated with the accepted proposals.

Remark: You may use the R function “rexp”. You must not use the function “rgamma”.

Question 3 (Q3):

Consider the Normal kernel:

$$h(x) = e^{-\frac{x^2}{2}}, \qquad -\infty < x < \infty.$$

In this question you should apply the ratio-of uniforms algorithm in order to generate random numbers from the standard Normal distribution.

Observe that:

$$m = \sup h(x) = 1$$

$$a = \inf\left(x\sqrt{h(x)}\right) = -\sqrt{2/e}$$

$$b = \sup\left(x\sqrt{h(x)}\right) = \sqrt{2/e}.$$

It follows that if $U_1 \sim U(0, m)$ and $U_2 \sim U(a, b)$ are independent then the conditional distribution of $X = U_2/U_1$, given the event $A = \{U_1^2 \leq \exp(-0.5U_2^2/U_1^2\}$, is the standard Normal distribution.

Write a function by the name "a1q3" that outputs a sequence of numbers from the standard Normal distribution.

The argument of the function is:

n = The length of the sequence random numbers.

The function should return a data frame of length n, with the generated U1 and U2, and X - the pseudo-random numbers from the standard Normal distribution. The dimension of the data frame is nx3

Remark: You must not use the function "rnorm".