Project 1: Generating pseudo random numbers from various distributions

Instructions: Submit a hard copy of your project, including the code. Send a softcopy of your code by e-mail with subject line "code for Project 1 – Your name". Please choose filenames that clearly indicate which program belongs to which (part of a) problem [e.g. "prob3bc.R" for problem 3 section c]. I will accept code in Matlab, R, or Mathematica. However, I will provide help only to those that use R, just to encourage you to use R. In fact the project is written with references to the R language, but you can use equivalent commends in Matlab or Mathematica.

1. `[10 points]` The distribution function for the exponential distribution with parameter λ is given by

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & \text{Otherwise.} \end{cases}$$

a. Write an `R` program that uses F^{-1} and the uniform random number generator in `R` (`runif`) to generate n numbers from the exponential distribution with λ. Use `set.seed(`*your birthday MMDDYY*`)`.

b. Generate $n = 10000$ values from the exponential random variable with parameter $\lambda = 2$, using your program (don't give me the numbers generated!). Use the `hist` command in R to plot the *density* histogram of the *relative frequencies* for the generated data. Overlay this histogram by the graph of the density of the exponential random variable with parameter $\lambda = 2$. The commands `dexp` and `lines`, or `curves` can be used. Select an appropriate number of bins for your histogram. Briefly comment on the relationship between the histogram and the density curve.

2. `[20 points]` Let $\phi_0(x)$ and $\phi_1(x)$ denote the density functions for two normal random variables with $\sigma = 1$, and respective means 0 and 3. The density for a mixture distribution of these is defined as

$$f(x) = \gamma\phi_0(x) + (1 - \gamma)\phi_1(x) \quad -\infty < x < \infty.$$

for a given admixture parameter $0 < \gamma < 1$.

a. Write an `R` program, using the normal random generator `rnorm` and the uniform random number generator `runif`, to generate n random numbers from this mixture. The inputs to your program should be n, and γ. Generate two sets of 10,000 data points; one with $\gamma = .75$, and another with $\gamma = .25$.

b. For each of the data sets generated, graph the density histogram and superimpose it by the density $f(x)$ defined in the equation above. Make sure to choose the number of bins for the histogram appropriately. In each case explain why the shape of the density that you obtain is expected.

3. [30 points] Problem 42 on page 111 of your text gives the pdf for the double exponential density with parameter λ. It also suggests a method to generate random numbers from the double exponential family using two random variables W and T, described in the problem.

a. Write a program that generates random numbers from the double exponential family. The input to the program should be the parameter λ, and the number of random numbers to be generated, n. The output should be n pseudo random numbers from the double exponential with parameter λ. You are only allowed to use `runif` in your program for random number generation.

b. Write a program that uses the Accept/Reject algorithm efficiently to generate n observations from the standard normal density $N(0, 1)$, using random numbers that are generated from the uniform(0,1) and the double exponential random variate with parameter $\lambda = 1$ [your program in part (a)]. Your program should also count and report the proportion of values that are rejected. Give the density histogram of $n = 10,000$ numbers generated from your program and superimpose it by the standard normal pdf.