HW2: Inverse CDF Method, writing R functions

Due: Wednesday, 2/7/2018, in class.

Note: Students are required to submit their homework as both PDF file and hardcopy by the due date of the assignment. Please turn in code separately and electronically. All electronic submissions should be made via D2L Dropbox and should follow the following naming convention: last name, first name, assignment number, proper extension. So, for example, if John Smith is turning in Homework 2, he would name the file Smith_John2.pdf. The associated code would be Smith_John2.Rmd. If you wish to break up your code into separate files, you may submit them as Smith_John2a.Rmd, Smith_John2b.Rmd, and so on. There will be a 20% penalty per day that your homework is late. Homework in the wrong format will not be given credit.

Problem 1: Inverse CDF method for random number generation

Weibull (a, b) distribution, parameters $a > 0$ and $b > 0$ has the following probability density function:

$$f(x) = ab^{-a}x^{a-1}\exp^{-(x/b)^a},\ x > 0.$$

1.a Obtain the cumulative density function (CDF).

1.b Solve $U = F(X)$ for X.

1.c Suppose that $U = 0.7$ is generated. Given $a = 1$ and $b = 1$, What is X?

Note: Since $1 - U \sim U(0, 1)$ as well, can replace $1 - U$ by U to get the final algorithm.

1.d Draw the histogram from 1000 sampled Weibull $(a = 1, b = 1)$ random variables based on 1.b.

Problem 2: Inverse CDF method for random number generation

Suppose

x	$p(x)$
-1	0.6
2.5	0.3
4	0.1

If $U = 0.63$, what do we take as X?

Problem 3

Determine the number of days between the following datas: January 1 in the year 2018, and January 1 in the year 2028.

3.a Write your own `R` function called, `countdays`.

3.b Use the built-in `R` function `as.Date` to compute.

Problem 4

4.a Write an `R` function, `f4`, that prints, with their row and column labels, only those elements of a correlation matrix for which $|\text{correlation}| \geq 0.9$.

4.b Using `cars` data, print the resuluts of `f4`.

Problem 5

5.a Derive the mean and variance of the Binomial distribution (n, p).

5.b Write an `R` function, `f5`, that simulates a student guessing at a True-False test consisting of 40 questions.

5.c Compute the number of student's correct answers. Assuming $n = 40$ and $p = 0.5$, compare with the theoretical values.

Problem 6 (Optional) The geometric distribution with pmf

$$P(X = k) = (1 - p)^{k-1} p, \; k = 1, 2, \ldots.$$

5.a Use the inverse CDF method to generate random variable X. When $p = 0.2$ and $U = 0.62$, what is X?

5.b Draw the histogram from 1000 sampled geometric ($p = 0.2$) random variables based on 5.a.