Assignment 1

STAT221

Due: 7 Aug 2020

Where you are asked to do some mathematical work you are expected to show all of your working in the assignment submission. Where you are asked to do something in R you should include ALL of the code used to produce the result in your assignment submission so it can be reproduced and checked by us. But do not include any code which is not required (i.e. be concise, just as you would be expected to do in writing essays or any other maths/stats work). Your code should include comments at the main steps to explain what it is doing.

Please submit your assignment in a single file (pdf) through Learn.

1) Congruential Random Number Generator [1 mark]

Write R code to generate $N = 1000$ pseudo-random numbers using a congruential random number generator. Choose reasonable values for the multiplier, increment, and modulo. Use your student ID as a seed.

2) Inversion sampler for the Pareto distribution [2 marks]

The Pareto distribution has two parameters, the location $a > 0$ and shape $b > 0$ parameters, and probability density function given by

$$f(x) = \begin{cases} \frac{a^b}{x^{b+1}} & \text{for } x \geq a \\ 0 & \text{for } x < a \end{cases}$$

and the cumulative distribution function is given by

$$F(x) = \begin{cases} 1 - \left(\frac{a}{x}\right)^b & \text{for } x \geq a \\ 0 & \text{for } x < a \end{cases}$$

a) Find the inverse CDF of the Pareto distribution
b) Write R code for an inversion sampler to generate a sample of $N = 1000$ random numbers from the Pareto distribution with $a = 2$ and $b = 3$. Use the generated CNRG pseudo-random numbers of Question 1. (do not use any functions, which directly implement the sampler like `rpareto()`)
c) Draw a histogram for the generated sample and add a line with the true density of the Pareto(2, 3) distribution.

3) Inversion sampler for a discrete random variable [2 marks]

A discrete random variable X has the probability mass function

x	1	2	3	4	5
f(x)	0.1	0.2	0.4	0.1	0.2

a) Write R code for an inversion sampler to generate a sample of 1000 random numbers from the distribution of X. Use the generated CNRG pseudo-random numbers of Question 1 (do not use any functions, which

directly implement the sampler like `rbinom()`)

b) Show the probability mass function of the sample of simulated random numbers. Include the true probability mass function in the plot for a comparison.

4) Inversion sampler for the binomial distribution [3 marks]

The binomial distribution has two parameters, the number of independent Bernoulli trials n and the probability of success on any one trial $0 < p < 1$, with probability mass function

$$f(x) = P(X = x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{for } x = 0, 1, \ldots, n \\ 0 & \text{otherwise.} \end{cases}$$

a) Write R code for an inversion sampler to generate a sample of 1000 random numbers from a binomial distribution with $n = 12$ Bernoulli trials and the probability of success on each trial of $p = 0.8$. Use the generated CNRG pseudo-random numbers of Question 1 (do not use any functions, which directly implement the sampler like `rbinom()`)

b) Show the probability mass function of the sample of simulated binomial random numbers. Include the true binomial probability mass function in the plot for a comparison.

5) Rolling dice [2 marks]

Simulate the sum of outcomes when rolling three dice, one with 6 sides, one with 12 sides, and one with 20 sides.

a) Generate a sample of $N = 1000$ totals of rolling the three dice, using an inversion sampler (use `runif()` and choose your student ID as seed).

b) Show the probability mass function of the sample of simulated totals.