**Homework 1**

**Statistical Genomics (545)**

**Spring 2023**

Instructor: Zhiwu Zhang

Due on Monday, February 6, 2023, 3:10PM PST

**Objectives**: 1) random variables; 2) derivation of random variables; 3) distribution of random variables; 4) derive statistics from samples; 5) sample statistics; and 6) define R functions.

**Hand in:** Email your R RMD and Knit (PDF) with subject of “StaGen545 HW1” to Zhiwu.Zhang@WSU.edu. Name your files as

“Homework1\_ firstname\_lastname”. PDF file is limited to one page per question (5 points off for violation). Your PDF file should focus on design, results, and interpretation by text, tables and figures. Leave the R code in the RMD file.

1. Define a random variable that is a function of random variables with known distributions such as uniform, binomial, Poisson, normal, Chi square, F, or t distributions. Name the distribution of your new random variable as your last name and develop an R function to generate the random variable. The input of your R function should include n, which is the number of variables to be generated, and parameters for the distribution you defined. Your distribution should be different from the listed known distributions (20 points).
2. Sample ten thousand observations from the distribution you defined. Make scatter, histogram, density, and cumulative density plots (20 points).
3. Create tables for your variable at different percentiles (1%, 5%, 10%, 50%, 90%, 95%, and 99%), and describe the impact of the parameters on your distribution (20 points).
4. Give an example to apply your distribution (**Extra credit**: 20 points).
5. Generate one or multiple samples with sizes of your choice from the distribution you defined, and define a novel statistic (not mean, sd, range) from your samples (20 points).
6. Create ten thousand replicates of your statistic and make the same plots listed in (2) (20 points).
7. Give an example to explain your statistic if it has a distribution that is not in the known distribution list (**Extra credit**: 20 points).