Stat414/614; Fall 2024; Worksheet 04; 20 Points; NAME:

This worksheet, based on the Module04 material, is available in Chapter 4 of Millard and Neerchal. The slides elaborate on these ideas and provide greater details. Also posted are three .Rmd files with r un-ready codes with extensive comments. Some of the problems below would require you to thoroughly study the codes, run them and understand the output.

1. In this first problem, we are drawing samples from an infinite population which is assumed to adequately modelled by a Normal distribution with $\mu = 10$ and $\sigma = 2$. Objective is to study the sampling distribution of the widely used statistics mean($\bar{X}$), median($\hat{X}$), minimum (Min) and maximum (Max) of the observed sample.

   (a) Generate 10,000 random samples, each of size 16 from a normal population with mu=10, and sigma=2, and compute min, median, maximum, and mean from each sample. Store these results in a dataframe (10,000 rows and four columns appropriately labeled).

   (b) Summarize the sampling distributions and plot histograms.

   (c) Compare the summary statistics and comment on the shapes of the histogram

   (d) Repeat (a)-(c) for a larger sample size.

2. Repeat steps of problem # for a skewed parent distribution such as a lognormal or a gamma distribution. Explain how would choose the parameters of these distributions so that they can be compared to the results from normal distribution.