To: Professor Yuce

       Applied Mathematics

       New York City College of Technology

From: Brian Holliday

       Applied Mathematics

       New York City College of Technology

Subject: Lab 20: Studentized and Wild Bootstrapping

Figure:

Figure 1: Bias for CHI8

Figure 2: Bias for EXP1

Figure 3: MSE for CHI8

Figure 4: MSE for EXP 1

Date: 4/23/20

Summary:

As we covered in Lab 19, Bootstrapping is a statistical resampling method in which we usually try to increase our confidence in descriptive statistic. Last time we just had the plain bootstrap with the mean square error and the bias on the same graph. Now we will be looking at the different types of possible bootstrap methods. We have the plain, normal, studentized and wild bootstrap. The studentized bootstrap takes the sample mean and subtracts it from the actual mean and then divides that value by the sample standard deviation. The wildbootstrap uses the random Poisson variables as part of its sampling, subtracting each variable by the sample mean and then multiplying that term by the random Poisson variable.

We want to see which one is the best at estimating the mean square error and the bias from both the chi-square distribution and the exponential distribution.

Bias for CHI8
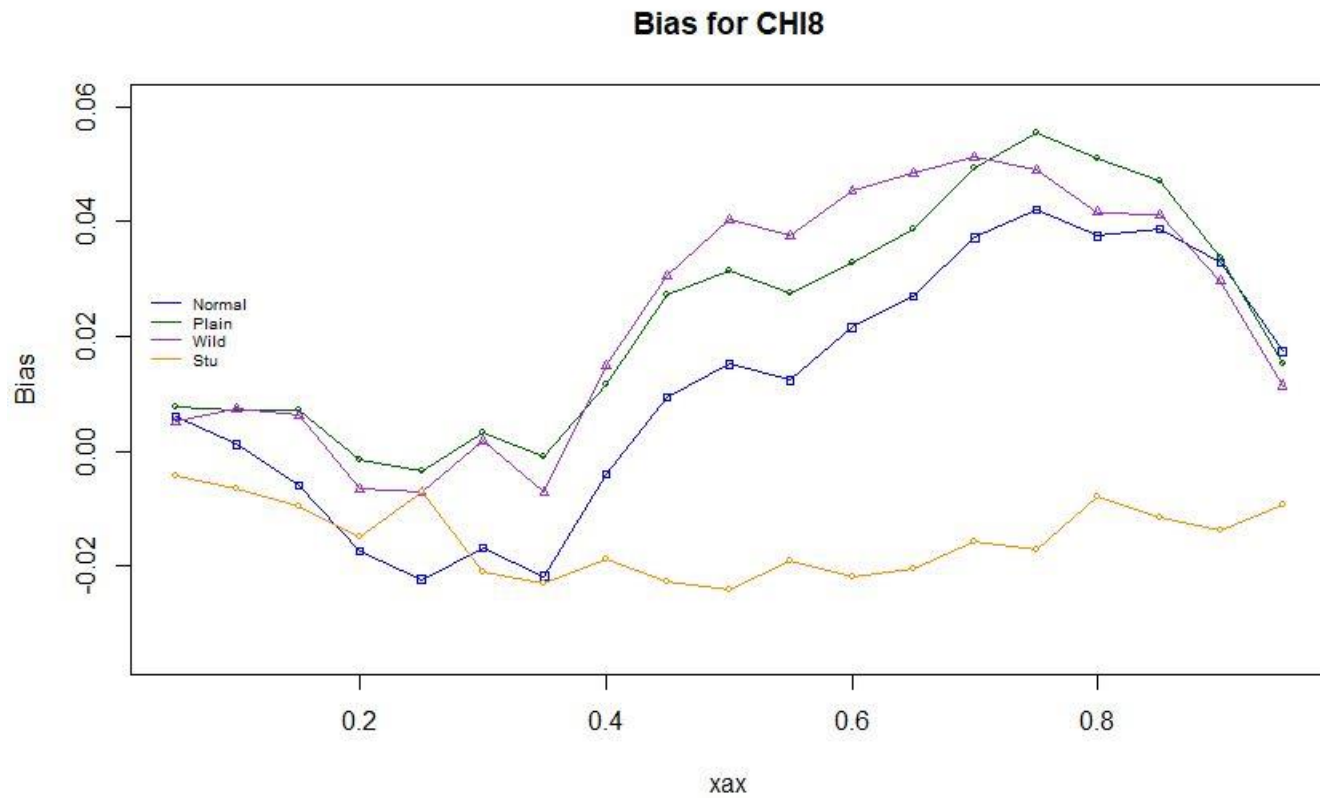
For the Bias with the chi-square distribution we could see that the wild and plain distribution are the most consistent across the quantiles, with similar results.

Bias for EXP1
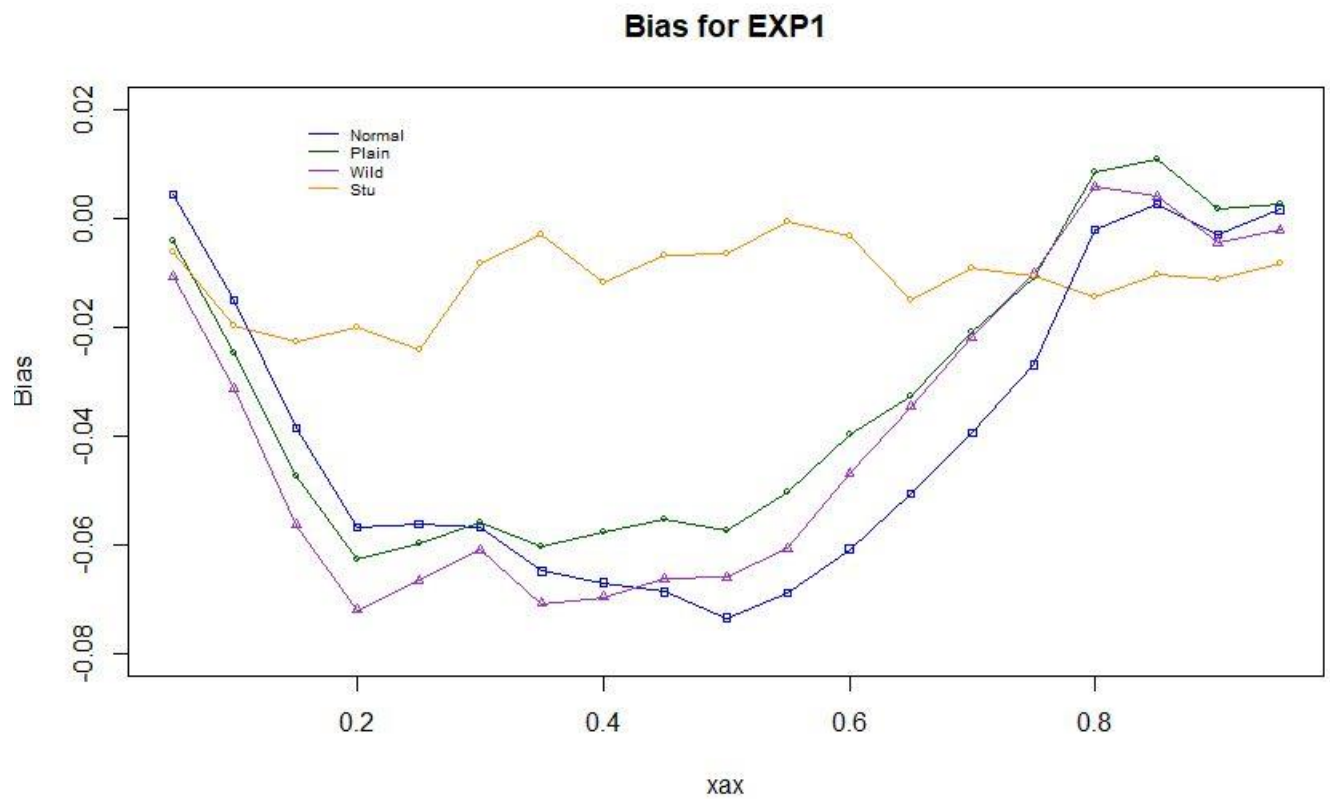
The normal and wild again appear to be good for the bias in exp. Normal can be used as well. The only one I would not use for both when it comes to bias is the studentized.

MSE for CHI8
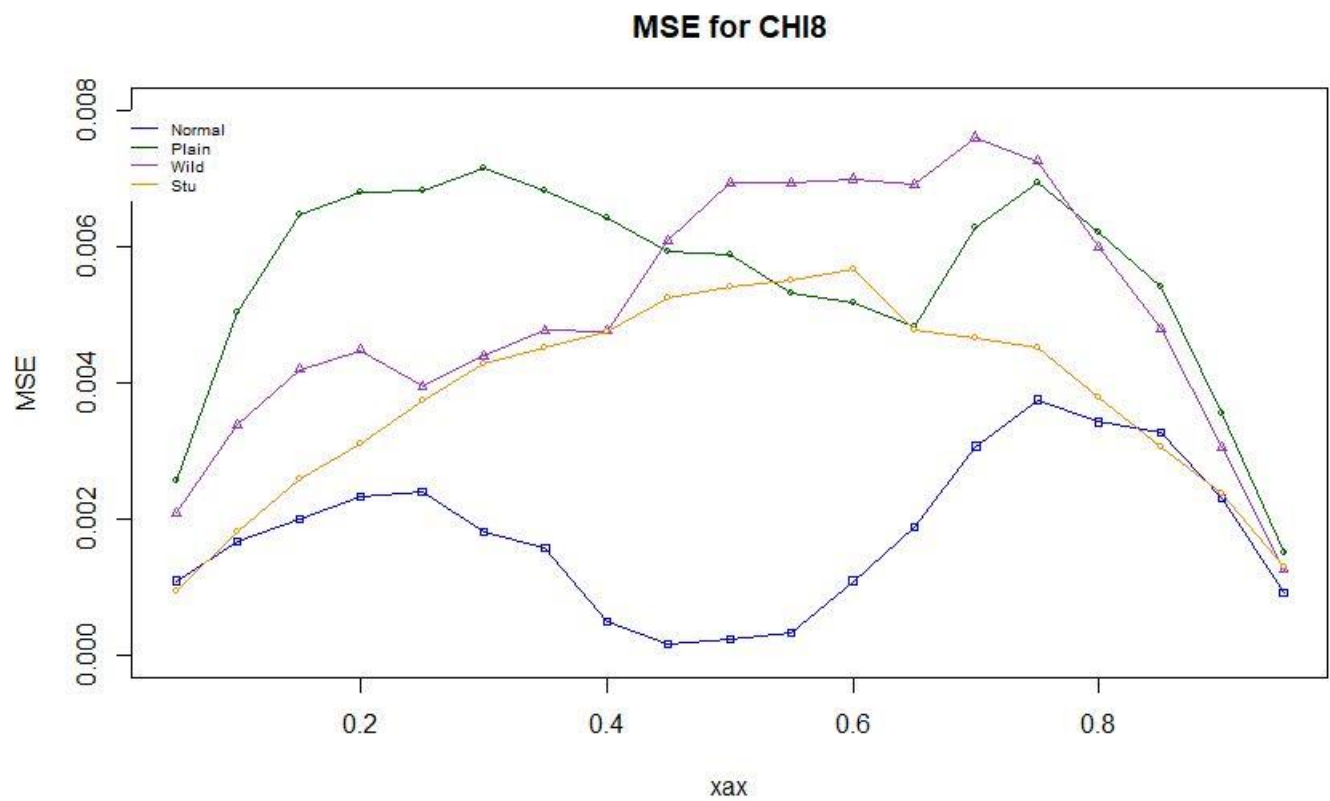
The data varies between all four of the types of bootstrap methods, but the studentized seems to be the most consistent across the quantiles.

MSE for EXP1
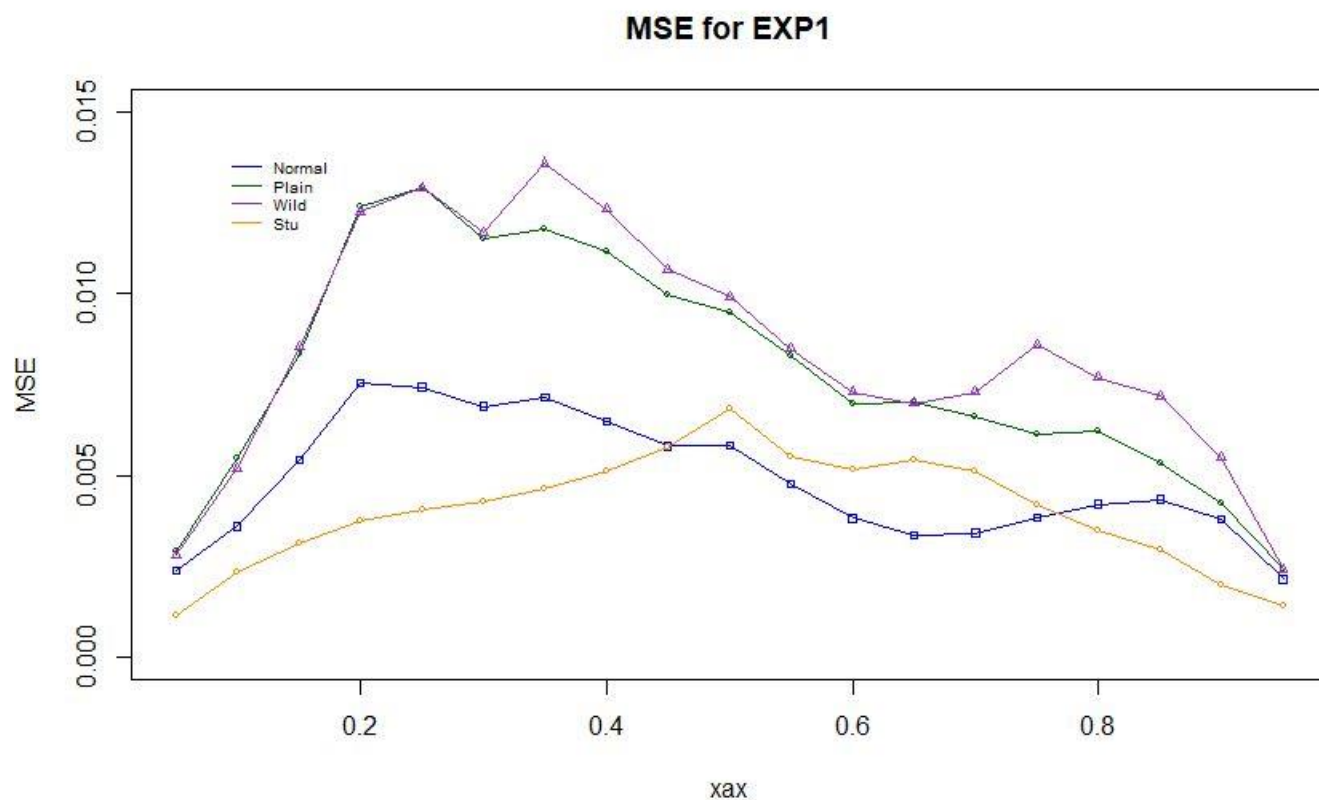
The variances in the normal and studentized appears to have the least variation.

The studentized appears to work better for MSE, while normal, plain, and wild all work well with bias. What bootstrap method you use will depend on the statistic that you are measuring and the distribution that data comes from.