STAT 6336
23 October 2019
Mary Lena and Mackay

Results

We sampled n=10,20,30,and 50 observations from a uniform[0,1] distributions, 100 000 times each. We then sampled again with the same n values until we had 100 000 that passed the Shapiro-Wilk test for normality, at p = 0.05. Running t-tests on both sets of 100 000 show that the Shapiro-Wilk filtered data have a false positive rate further away from the reported alpha than the unfiltered data, deviating further and further with increasing sample size. Reducing alpha to 0.001 on the Shapiro-Wilk test yields results that are closer by about three orders of magnitude. Results with this Shapiro-Wilk alpha are also better than the unfiltered data.