

## Math 6070-1, Spring 2014; Assignment #2

Assigned on: Friday January 31, 2014  
Due: Wednesday February 5, 2014

In this assignment, you are asked to experiment with some of the initial practical aspects of density estimation. For this you need to generate 1,000 samples from a  $N(0, 1)$  distribution.

1. Complete reading Sections 1 and 2 of the Density estimation module at <http://www.math.utah.edu/~davar/math6070/2014/DensityEstimation.pdf>.
2. Create a histogram of your data, with  $x_0 = 0$ , using several different bin sizes. Since you know the exact sampling distribution, experiment with the bin sizes until you have a feeling for which bin size[s] produce more accurate representations of your data. Explain your findings carefully.  
Remember that you have to write your own code for this [though you are encouraged to run a canned program to check your work for yourself].
3. Produce a kernel density estimate of the  $N(0, 1)$  pdf using your sample, using the following different kernels:

- (a)  $K(x) := (2\pi)^{-1/2} \exp\{-x^2/2\}$  [the standard normal kernel];
- (b)  $K(x) := \frac{1}{2} \mathbf{I}\{|x| < 1\}$  [the Unif(-1, 1) kernel];
- (c)  $K(x) := \mathbf{I}\{0 < x < 1\}$  [the Unif(0, 1) kernel];
- (d)  $K(x) := \exp(-x) \mathbf{I}\{x > 0\}$  [the one-sided exponential kernel];
- (e)  $K(x) := \frac{1}{2} \exp(-|x|)$  [the double-exponential kernel];
- (f)  $K(x) := \pi^{-1} \{1 + x^2\}^{-1}$  [the Cauchy kernel].

Do this for an array of different bandwidths in each case. Experiment until you find a satisfactory choice of  $K$  and  $h$ . Explain your findings carefully.

4. Generate 1,000 samples from a Unif(0, 1) distribution next, and redo all the steps of problem #3. Do you notice a difference in your findings?