Homework # 3
due on 03/08/2018

Instructions: The solutions to this HW will be posted in one week. Please do not write your solutions in red ink as this HW should be self graded or peer graded in red ink. Grades will be awarded for the completion of the problems and having graded the solutions. Correctness or incorrectness of the solutions will not be considered for the overall HW grade.

1 Problem 1

Let $X$ be a continuous random variable with the following p.d.f.:

$$f_X(x) = \begin{cases} 2x, & \text{if } 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

Also suppose that: $Y|X = x \sim \text{Geometric}(x)$. Find the MAP estimate of $X$ given $Y = 3$.

2 Problem 2

In the previous homework we computed the MLE estimate $\hat{\theta}_{MLE}$ derived from i.i.d. observations $X_{(1)}, \ldots, X_{(n)}$ under the following p.d.f.:

$$f(x) = \theta^2xe^{-\theta x}, \quad 0 < x, 0 < \theta < \infty \quad (2)$$

a) Find the bias, variance, and mean-square error (MSE) of $\hat{\theta}_{MLE}$.

b) Is $\hat{\theta}_{MLE}$ unbiased? If no, provide an unbiased estimator for $\theta$. What is the mean-square error (MSE) of this new estimator? Is it higher or lower than the MSE of $\hat{\theta}_{MLE}$?

c) Now, given $\hat{\theta}_{MLE}$, we would like to shrink it (that is scale it by a constant $k$) in order to obtain an estimator of minimum mean-square error (MSE). What is the value of $k$ that achieve this? What is the bias and variance of this estimator?
3 Problem 3

River floods are often measured by their discharges (in units of feet cubed per second). The following table gives the flood discharges of the Blackstone River in Woonsocket, Rhode Island, in each of the years from 1956 to 1965.

<table>
<thead>
<tr>
<th>Year</th>
<th>Flood Discharge (ft$^3$/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1956</td>
<td>8710</td>
</tr>
<tr>
<td>1957</td>
<td>3850</td>
</tr>
<tr>
<td>1958</td>
<td>4970</td>
</tr>
<tr>
<td>1959</td>
<td>5398</td>
</tr>
<tr>
<td>1960</td>
<td>4780</td>
</tr>
<tr>
<td>1961</td>
<td>4020</td>
</tr>
<tr>
<td>1962</td>
<td>5790</td>
</tr>
<tr>
<td>1963</td>
<td>4510</td>
</tr>
<tr>
<td>1964</td>
<td>5520</td>
</tr>
<tr>
<td>1965</td>
<td>5300</td>
</tr>
</tbody>
</table>

Take the yearly flood discharge as a random variable $X$. Use the above data to estimate the cdf $F_X(5000)$ using empirical distribution function.

4 Problem 4

In the context of Question 3, estimate the density at 5000, i.e. $f_X(5000)$, using histogram method with bin edges $\{0, 2000, 3000, 4000, 4500, 5000, 5500, 6000, 10000\}$.

5 Problem 5

In the context of Question 3, estimate the density at 5000 using kernel density approach. Suppose we choose the uniform kernel and $h = 300$. 