   Let \( x \) have an exponential density
   \[
   p(x | \theta) = \begin{cases} 
   \theta e^{-\theta x} & x \geq 0 \\
   0 & \text{otherwise}
   \end{cases}
   \]
   a. Plot \( p(x | \theta) \) versus \( x \) for \( \theta = 1 \)
   b. Plot \( p(x | \theta) \) versus \( \theta \) over \( (0 \leq \theta \leq 5) \) for \( x = 2 \).
   c. Suppose that \( n \) samples \( x_1, \ldots, x_n \) are drawn independently according to \( p(x | \theta) \)
      Show that the Maximum-likelihood estimate for \( \theta \) is given by:
      \[
      \hat{\theta} = \frac{1}{n} \sum_{k=1}^{n} x_k
      \]

2. Duda, Hart, Stork 3.35.
   Let the sample mean \( \mu_n \) and the sample covariance matrix \( C_n \) for a set of \( n \) samples
   \( x_1 \ldots x_n \) (each of which is \( d \)-dimensional) be defined by
   \[
   \mu_n = \frac{1}{n} \sum_{i=1}^{n} x_i
   
   C_n = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \mu_n)(x_i - \mu_n)^t
   \]
   We call these the “nonrecursive” formulae.
   (a) What is the computational complexity of calculating \( \mu_n \) and \( C_n \) by these formulae?
   (b) An alternative recursive technique is based on successive addition of new
      samples \( x_{n+1} \) according to the following formulae.
      \[
      \mu_{n+1} = \mu_n + \frac{1}{n+1}(x_{n+1} - \mu_n)
      
      C_{n+1} = \frac{n}{n+1} C_n + \frac{1}{n+1} (x_{n+1} - \mu_n)(x_{n+1} - \mu_n)^t
      \]
      Show that the mean computed using the recursive method is the same as the nonrecursive method. (The Covariance can be similarly shown, but it’s more tedious).
(c) What is the computational complexity of finding $\mu_n$ and $C_n$ by these recursive methods?
(d) Describe situations where you might prefer to use the recursive method for computing $\mu_n$ and $C_n$, and ones where you might prefer the nonrecursive method?

3. Consider a normal $p(x) = N(\mu, \sigma^2)$ and Parzen window function $\varphi(x) = N(\mu, 1)$. Show that the Parzen window estimate

$$p_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} \varphi \left( \frac{x - x_i}{h_n} \right)$$

has the following property

$$E[p_n(x)] = N(\mu, \sigma^2 + h_n^2)$$

4. Consider the following set of two-dimensional vectors from three categories:

<table>
<thead>
<tr>
<th></th>
<th>$\Omega_1$</th>
<th>$\Omega_2$</th>
<th>$\Omega_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>$X_2$</td>
<td>$X_1$</td>
<td>$X_2$</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>0</td>
<td>-10</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>-2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>10</td>
<td>-4</td>
</tr>
</tbody>
</table>

(a) Plot the decision boundary resulting from the nearest neighbor rule just for categorizing $\omega_1$ and $\omega_2$. Find the sample mean $m_1$ and $m_2$ and on the same figure sketch the decision boundary corresponding to classifying $x$ by assigning it to the category of the nearest sample mean.
(b) Repeat part (a) for categorizing only $\omega_1$ and $\omega_3$.
(c) Repeat part (a) for categorizing only $\omega_2$ and $\omega_3$.
(d) Repeat part (a) for three-category classifier, classifying $\omega_1$, $\omega_2$ and $\omega_3$. 