

IT 542: Pattern Recognition and Machine Learning

Assignment 4

ID:201811024

Name- Rajat Kumar

(1) Draw samples from $p(x) = P1.p1(x) + P2.p2(x)$

Where, $P1 = 0.6$, $P2 = 0.4$, $p1(x) \sim N(5,10)$ and $p2(x) \sim N(10,15)$.

Draw a histogram for the data.

Code:

```
from scipy.stats import norm
```

```
import matplotlib.pyplot as plt
```

```
import numpy as np
```

```
P1=0.6
```

```
P2=0.4
```

```
mu1,mu2,sigma1,sigma2=5,10,np.sqrt(10),np.sqrt(15)
```

```
s1 = np.random.normal(mu1, sigma1, 20)
s2 = np.random.normal(mu2, sigma2, 20)
s=P1*s1+P2*s2
#s=s1+s2

n, bins, patches = plt.hist(x=s, bins='auto',
color='#0504aa',alpha=0.7, rwidth=0.85)

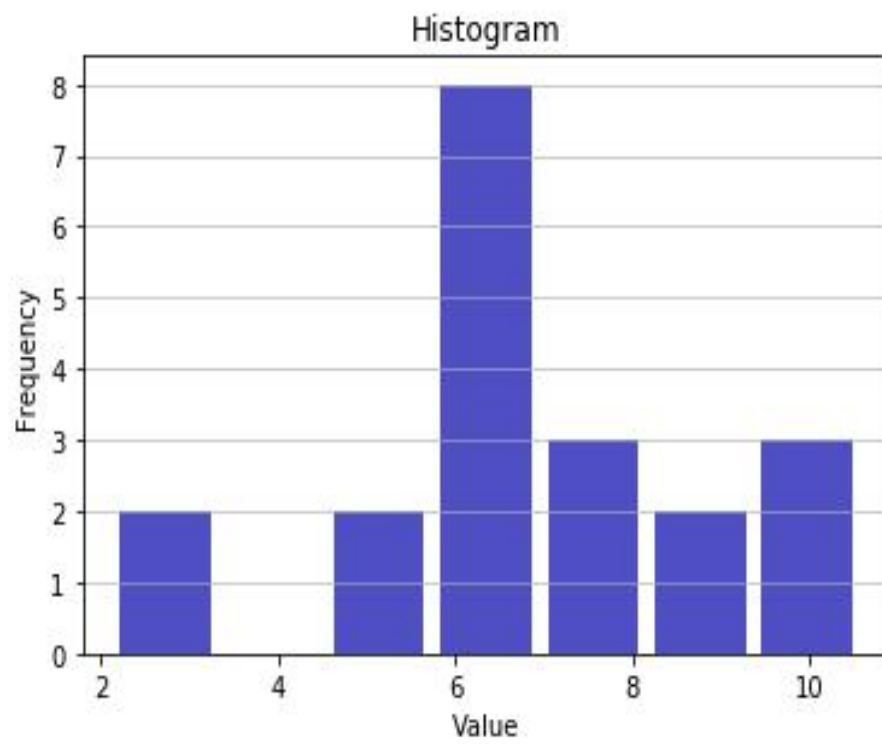
plt.grid(axis='y', alpha=0.75)

plt.xlabel('Value')
plt.ylabel('Frequency')

plt.title('Histogram')

print(n)
print(bins)
print(patches)
```

OUTPUT:



2. Use the KDE to estimate pdf of given samples:

Code:

```
import numpy as np

import matplotlib.pyplot as plt

x1= np.random.normal(5,1,100)
x2=np.random.normal(-5,1,100)

data=x1

data= np.append(data,x2)

#data=0.6*x1+0.4*x2


print(np.shape(data))

N=200

print(data)

def phi(x):

    return (np.exp(-0.5*(x**2))/(np.sqrt(2*np.pi)))    #
Normal Density-> Mean =0, var=1
```

```
def tpdf(x):
```

```
    return (phi(x-5)*0.5) + (phi(x+5)*0.5) # True
```

```
Density-> Mean 10,-10, var=1
```

```
def kernel(x,h):
```

```
    return np.mean(phi((x-data)/h)/h) # Kernel
```

```
Density
```

```
def kpdf(x,h):
```

```
    temp=[]
```

```
    for i in range(len(x)):
```

```
        temp.append(kernel(x[i],h))
```

```
    return temp
```

```
x = np.linspace(-25,+25,100)
```

x

h=0.3

```
print(len(kpdf(x,h)))
```

```
print(len(tpdf(x)))
```

Kernel - Bandwidth estimated by Silverman's Rule of Thumb

```
h=np.power((4/3*N),1/5)
```

```
ax1 = plt.axes([0, 0, 3, 0.5])
```

```
plt.plot(kpdf(x,h),'r') # Kernel Density Estimated
```

```
plt.plot(tpdf(x),'g') # True density
```

```
plt.title(['h=',h])
```

```
plt.show()
```

```
h=np.linspace(0.01,5,num=50)
```

```
for i in range(len(h)):
```

```
    ax1 = plt.axes([0, 0, 3, 0.5])
```

```
    plt.plot(kpdf(x,h[i]),'r') # Kernel Density Estimated
```

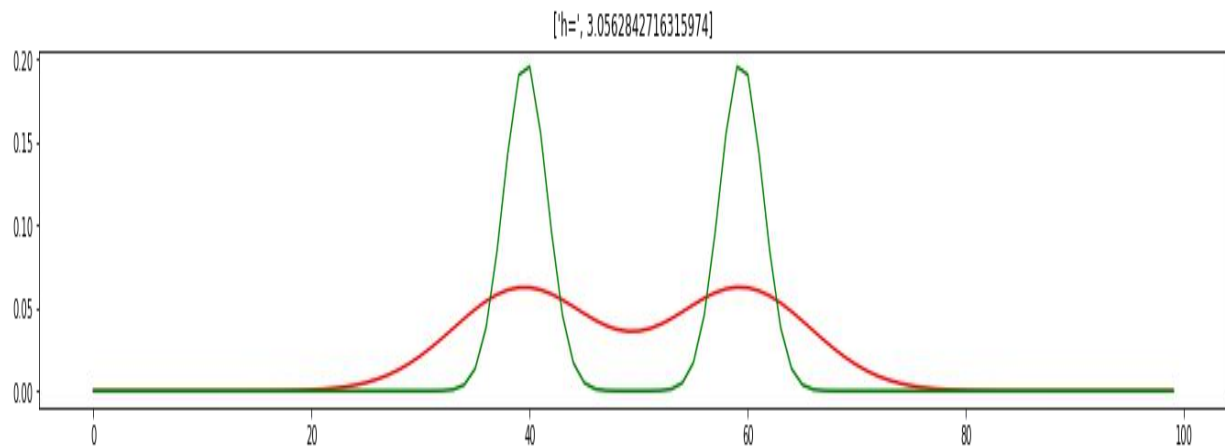
```
plt.plot(tpdf(x),'g') # True density
```

```
plt.title(['h=',h[i]])
```

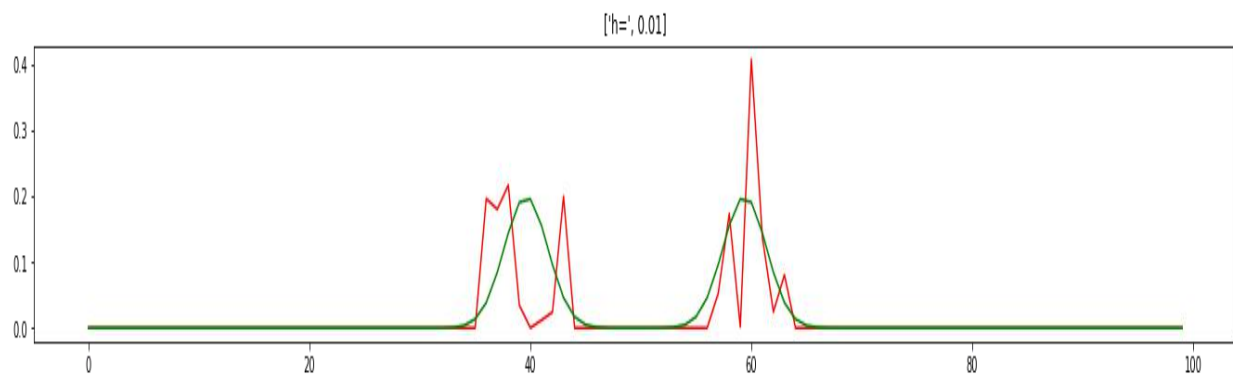
```
plt.show()
```

Output:

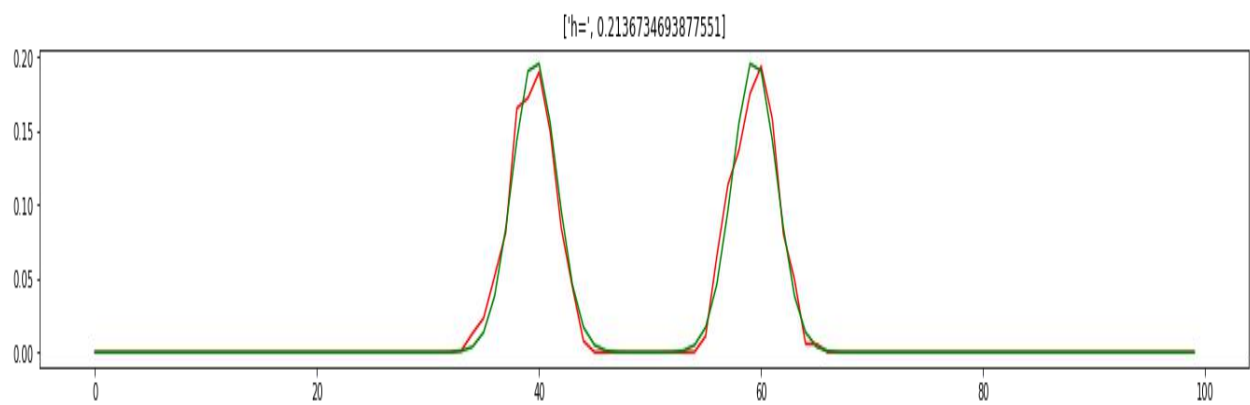
h= 3.05



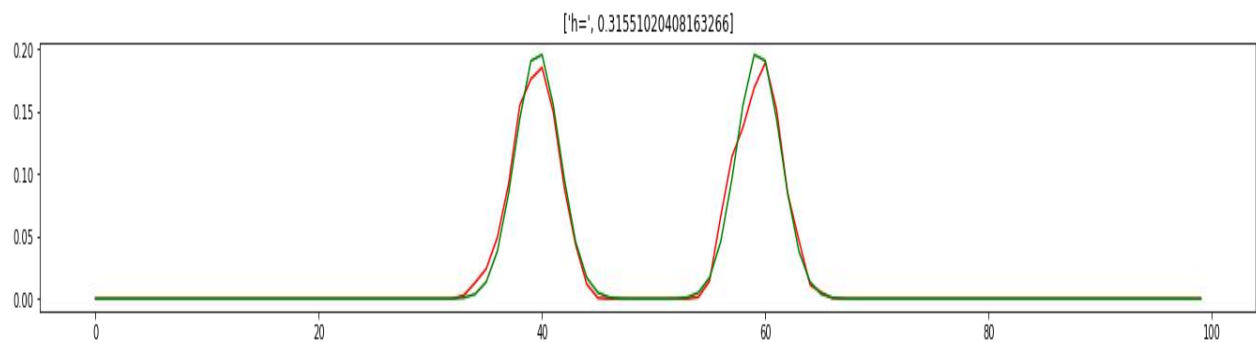
h= 0.01



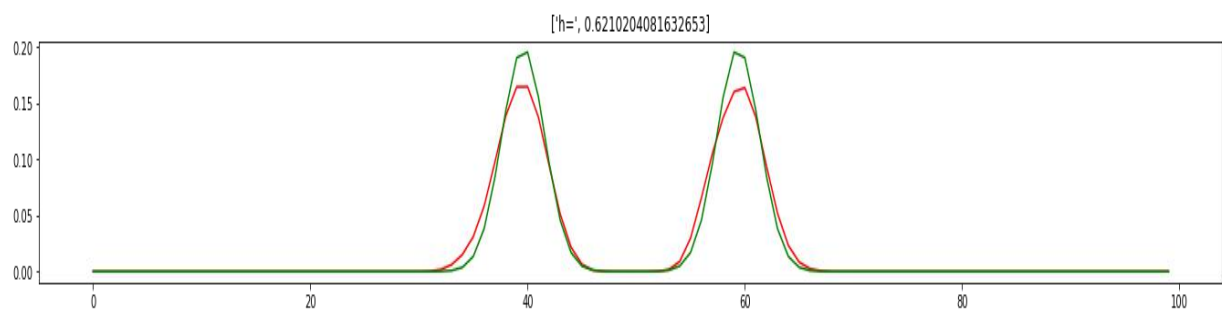
h= 0.21



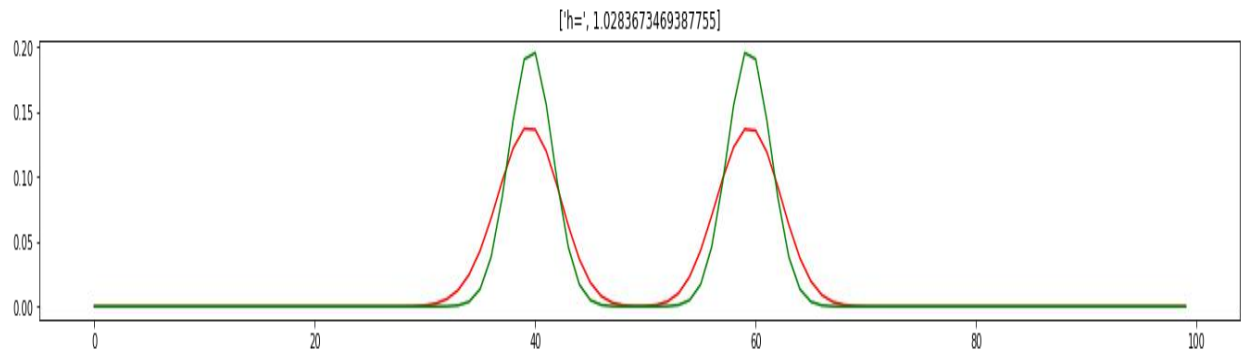
h= 0.31



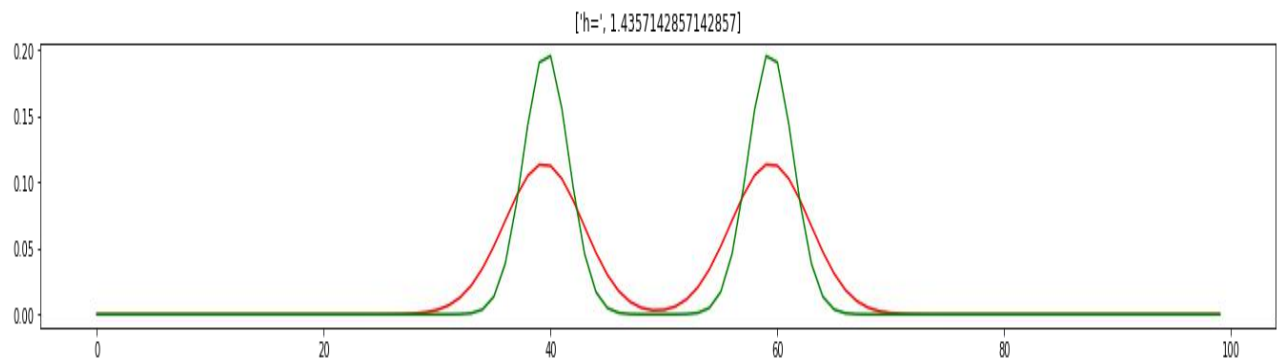
h= 0.62



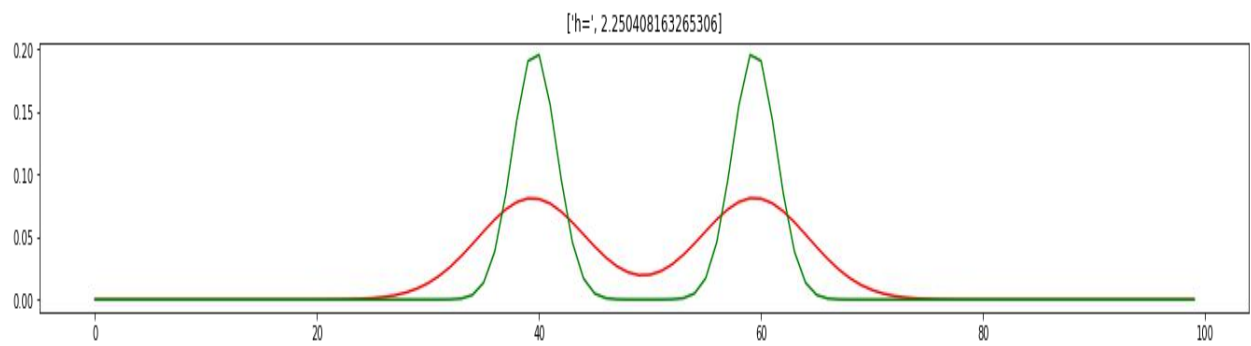
h= 1.02



h= 1.43



h= 2.25



h= 3.9

