Ques 3:

Two classes having variance = 1, mu=1 and mu = 2 respectively.

If we plot the gaussian curve for both classes, it will come to be this:

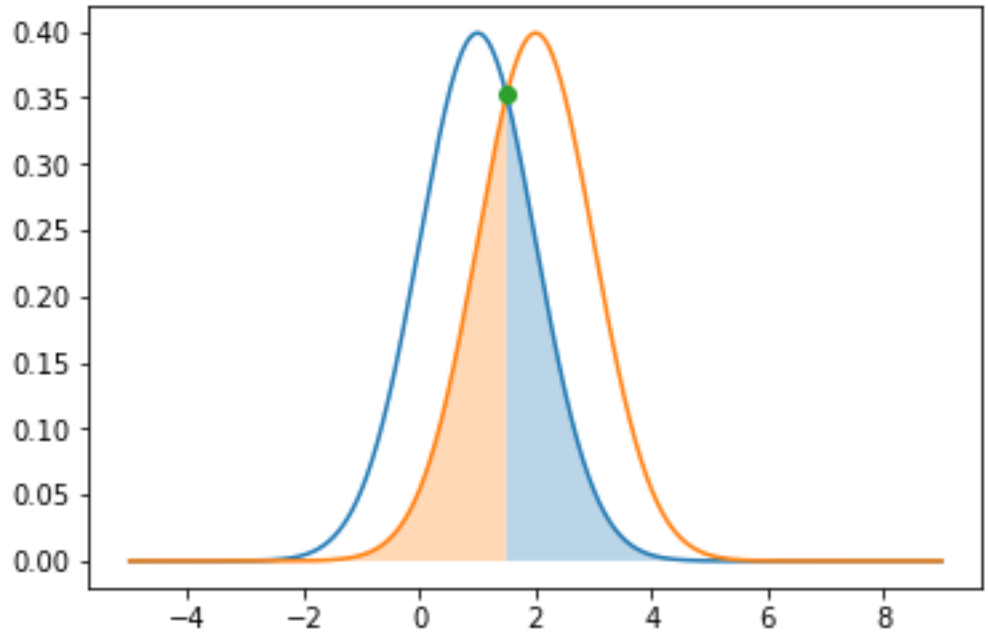
Used this reference website to plot the two gaussian curves, get the intersection point and get the area under the intersection of two bell curves.

SCIPY, O.

Overlapping probability of two normal distribution with scipy

In-text: (scipy, 2018)

Your Bibliography: scipy, O. (2018). *Overlapping probability of two normal distribution with scipy*. [online] Stackoverflow.com. Available at: https://stackoverflow.com/questions/32551610/overlapping-probability-of-two-normal-distribution-with-scipy [Accessed 5 Feb. 2018].



```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm
norm.cdf(1.96)
def solve(m1,m2,std1,std2):
  a = 1/(2*std1**2) - 1/(2*std2**2)
  b = m2/(std2**2) - m1/(std1**2)
  c = m1**2 /(2*std1**2) - m2**2 / (2*std2**2) - np.log(std2/std1)
  return np.roots([a,b,c])
m1 = 1
std1 = 1
m2 = 2
std2 = 1
#Get point of intersect
result = solve(m1,m2,std1,std2)
#Get point on surface
x = np.linspace(-5,9,10000)
plot1=plt.plot(x,norm.pdf(x,m1,std1))
plot2=plt.plot(x,norm.pdf(x,m2,std2))
plot3=plt.plot(result,norm.pdf(result,m1,std1),'o')
#Plots integrated area
r = result[0]
olap = plt.fill_between(x[x>r], 0, norm.pdf(x[x>r],m1,std1),alpha=0.3)
olap = plt.fill_between(x[x<r], 0, norm.pdf(x[x<r],m2,std2),alpha=0.3)
# integrate
area = norm.cdf(r,m2,std2) + (1.-norm.cdf(r,m1,std1))
print("Area under curves ", area)
plt.show()
```

```
Area under curves  0.6170750774519738
```

Explanation:

- The two gaussian curves having unique variance and mean 1 and 2 respectively are intersecting at a common point of 1.5
- The classification accuracy of the two bell curves is up to their intersection points.
- Area under the two curves = 0.617
- Area for one curve from intersection point to the curve end point = 0.617/2 = 0.308
- Remaining area for one bell curve = 1 – 0.308 = 0.692
- This remaining area 0.692 is the area for one curve upto its intersection point which can be classified accurately.
- **So, The percentage classification accuracy of one bell curve is 69.2% approx. The theoretical limit of gaussian classification in this case is 69.2%.**

Calculation Explanation:

- First find the intersection point of two bell curves which is in this case is 1.5
- Next step is to find the area under the intersection points of bell curves.

- Area from the intersection point upto right extreme point for a curve could be calculated by calculating the normal cumulative distributive function at intersection point r ie at 1.5 given a mean and standard deviation. (This is explained in the program above)
- Next we can find the remaining area by subtracting intersection area from 1 to get classification accuracy.