**ISTANBUL TECHNICAL UNIVERSITY**

**COMPUTER ENGINERING DEPARTMENT**

**BLG 527E MACHINE LEARNING**

**CRN: 13817**

**Instructor: Zehra Çataltepe**

**Homework #1**

**October 4, 2017**

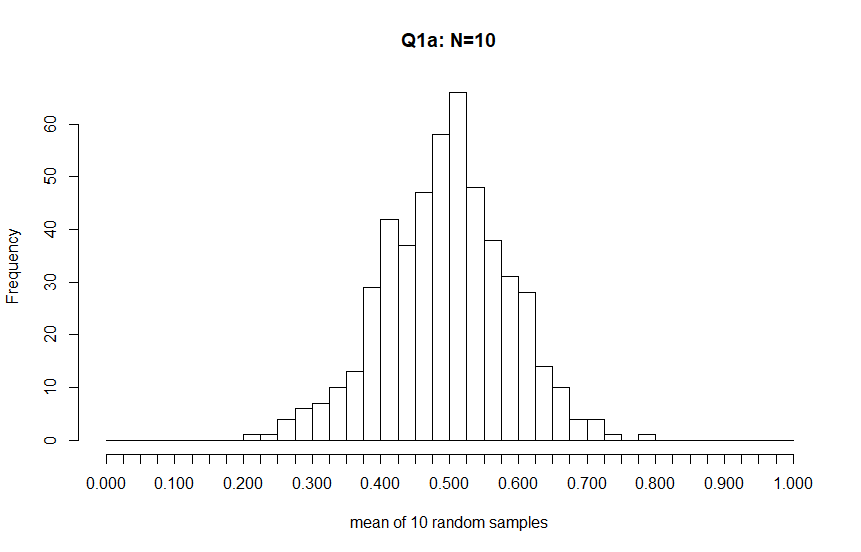
**Tuğrul Yatağan**

**504161551**

**Answers**

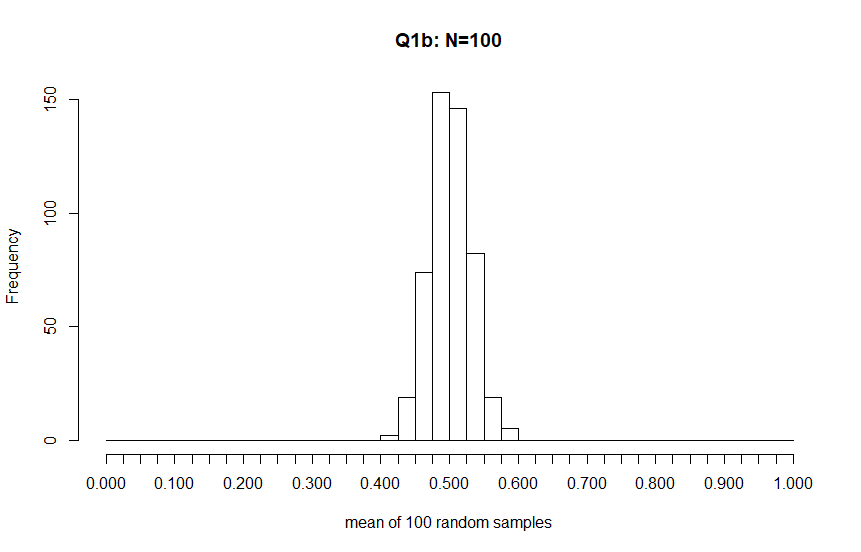
**Q1a)**

N=10

****

**Q1b)**

N=100

****

**Q1c)**

Both sampling distribution of the mean approaches a normal distribution. But, as N grows, the shape of the histogram resembles a Normal distribution more closely. N is the sample size for each mean.

**Q2)**

gi(x) = ln(p(x|Ci)) + ln(P(Ci))

p(x|C1) = N(0,1)

p(x|C2) = N(1,2)

**Q2a)**

P(C1) = P(C2) = 0.5

g1(x) = ln(p(x|C1)) + ln(P(C1)) = ln(N(0,1)) + ln(0.5)

g2(x) = ln(p(x|C2)) + ln(P(C2)) = ln(N(1,2)) + ln(0.5)

Gaussian distribution is:

 P(x)=1/(sigmasqrt(2pi))e^(-(x-mu)^2/(2sigma^2)) 

So;

N(0,1) = . ln(N(0,1)) = ln() -

N(1,2) = . ln(N(1,2)) = ln() -

Then;

g1(x) = ln() - + ln(0.5) = ln() -

g2(x) = ln() - + ln(0.5) = ln() -

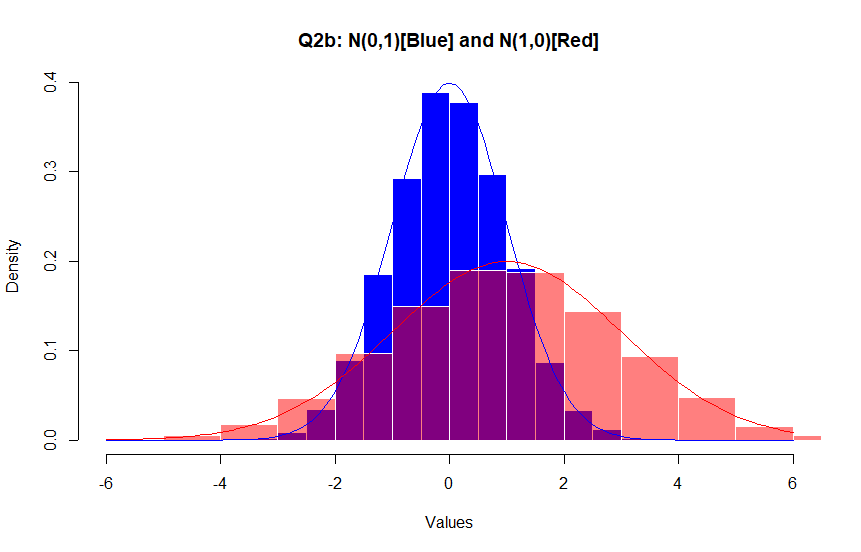
**Q2b)**

p(x|C1) = N(0,1)

p(x|C2) = N(1,2)

Histograms represent two Gaussian distribution with mean=0,1 and standard deviation=1,2

Lines represent density distribution of two Gaussian distribution with mean=0,1, standard deviation=1,2

****

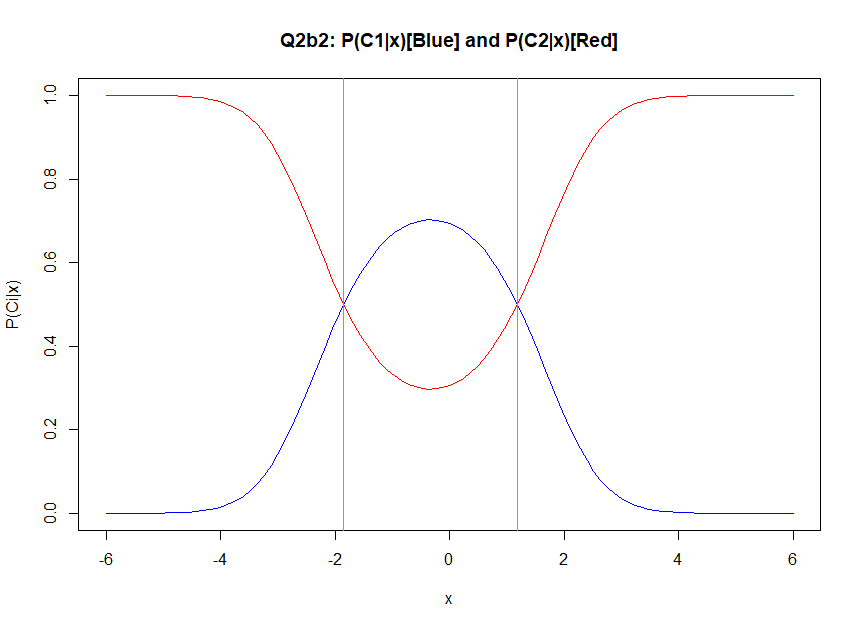
P(C1|x) = = =

P(C2|x) = = =

If we assume P(C1)=P(C2)=0.5, then;

P(C1|x) = = = =

P(C2|x) = = = =

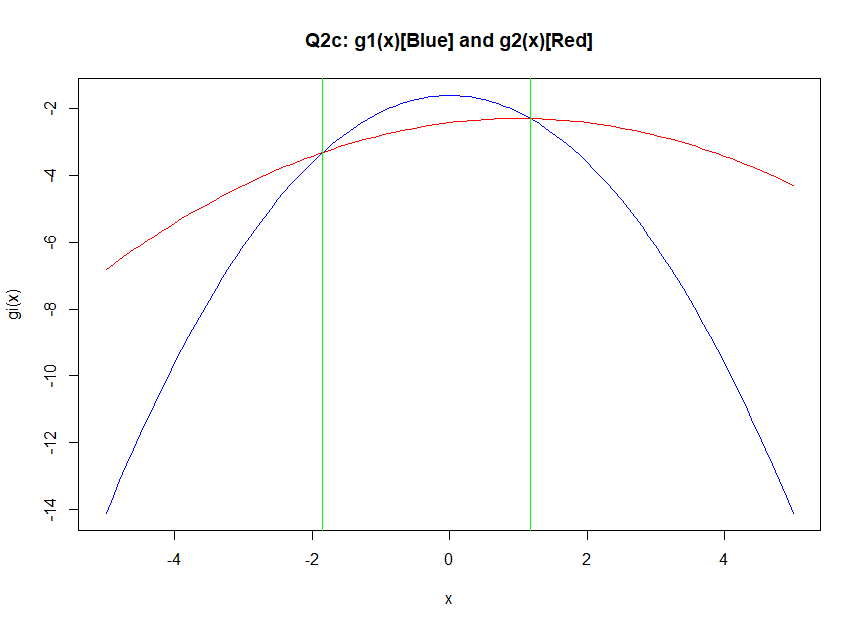


Regions where P(C1|x) is greater than P(C2|x) is class C1, Regions where P(C2|x) is greater than P(C1|x) is class C2.

**Q2c)**

g1(x) = ln() - + ln(0.5) = ln() -

g2(x) = ln() - + ln(0.5) = ln() -



Regions where g1(x) is greater than g2(x) is class C1, Regions where g2(x) is greater than g2(x) is class C2.

**Q2d)**

P(C1) = 0.2

P(C2) = 0.8

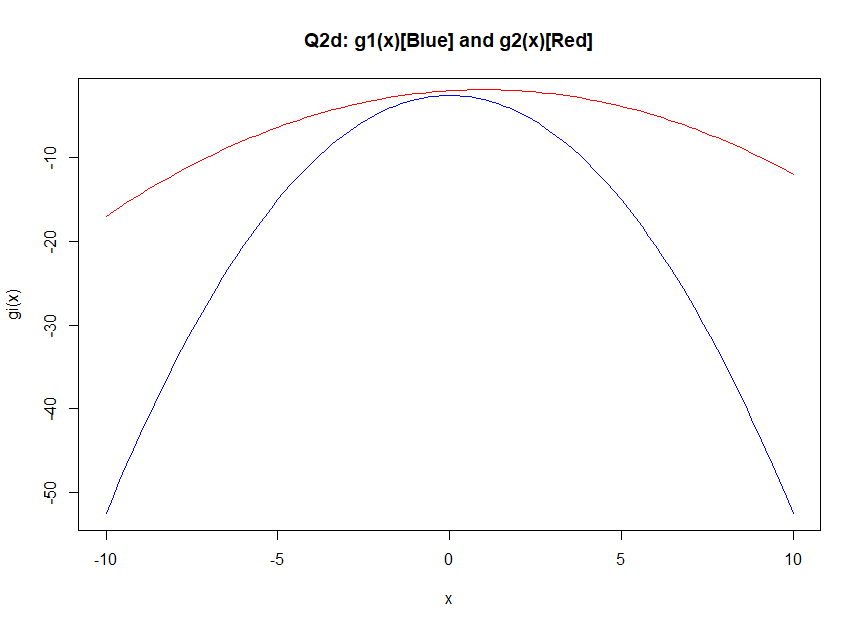
We know that from Q2a;

ln(N(0,1)) = ln() -

ln(N(1,2)) = ln() -

g1(x) = ln(p(x|C1)) + ln(P(C1)) = ln(N(0,1)) + ln(0.2) = ln() - + ln(0.2)

g2(x) = ln(p(x|C2)) + ln(P(C2)) = ln(N(1,2)) + ln(0.8) = ln() - + ln(0.8)



g2(x) is always greater than g1(x) so C2 is always identified.