# HW: Bayesian Parameter Estimation

CSC 591: Algorithms for Data-Guided Business Intelligence

**Rachit Shah (**[**rshah25@ncsu.edu**](mailto:rshah25@ncsu.edu)**)**

February 17, 2019

## Problem

Data: i.i.d with known variance and unknown mean.

Prior Distribution is Gaussian

Likelihood from sample data which is Gaussian.

## Q1: Derive the formula for the posterior distribution of μ

**Q2: Show that the posterior distribution is the Gaussian, p(μ|X)~N(μn,σ2n)**

From equation 1, we can see that the posterior distribution is a Gaussian distribution as it is proportional to a Normal Distribution of the form

**Q3: Show the derivation and the final estimate for μn and 1/σ2n**

**Q4: If the mean of the posterior density (which is the MAP estimate), μn is written as the weighted average of the prior mean, μ0, and the sample (likelihood) mean, X¯, then what are the formulas for the weights?**

Therefore, the weights are:

**Q5: Are the weights in Question #4 directly or inversely proportional to their variances (justify)?**

As you can see from equation 4, weight corresponds to the weighted average of the prior mean and its equation has the variance in the denominator. While weight is the weighted average of the likelihood and its equation also has the corresponding variance in the denominator. Hence, both weights are **inversely proportional** to their variances.

**Q6: Do the weights in Questions #4 sum up to 1 (justify)?**

**Yes.**

**Q7: Is each weight between zero and one (justify)?**

From equation 5 in the previous equation, we can see that .

Weight has in both numerator and denominator but has an additional in the denominator. Since variance can’t be negative and n too can’t be negative, the denominator is always greater than or equal to the numerator. When n becomes too large, the denominator will approach infinity and will get closer to 0. While if n grows closer to 0, will approach 1. Hence, always has its value between 0 and 1.

Weight has in both numerator and denominator but has an additional in the denominator. Since variances can’t be negative and n too can’t be negative, the denominator is always greater than or equal to the numerator. When n is very large such that , the value of will approach 1. However, for low values of n, if value of becomes greater than , it will approach 0 but never become negative. Hence, always has its value between 0 and 1.

Hence, **Yes**, both weights always have value between 0 and 1.

**Q8: Given your answers for Questions #4-7, what can you say about the value of μn w.r.t. the values of μ0 and X¯**

From question 7, we see that as n (samples) increases, reaches 0 while reaches 1. And vice versa. Also, . Since, both weights have their value between 0 and 1, theoretically the maximum value can reach is which will happen when both weights are 1 or 0 if both weights are 0. But, since both weights are inversely dependent on n, they will never both be 1 or 0. Hence, the value of lies anywhere between the values of and .

**Q9: If σ2 is known, then for the new instance xnew, show that p(xnew|X)~N(μn,σ2n+σ2)**

is given to be a normal distribution and we found the posterior distribution to also be a normal distribution

Now, , where, is the normal distribution and is the posterior normal distribution Hence, we have reframed the equation to be a sum of the 2 normal (gaussian) distributions.

**Theorem:** If X1, X2, ... , Xnare mutually independent normal random variables with means μ1, μ2, ... , μnand variances σ21,σ22,⋯,σ2nσ12,σ22,⋯,σn2, then the linear combination: follows the normal distribution:

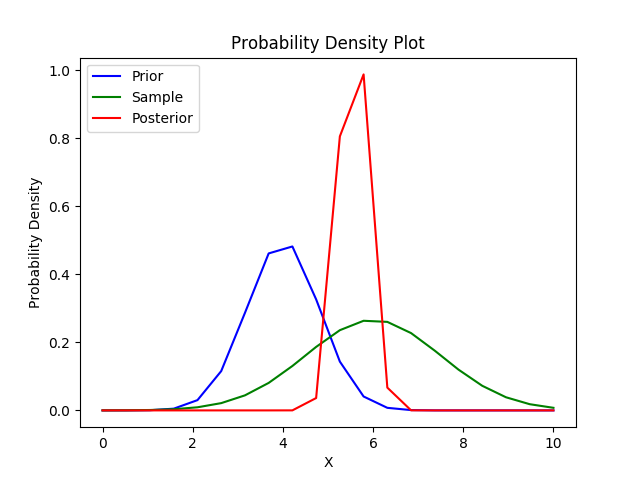
Since, the problem statement states that the data is i.i.d and follows gaussian distribution, the condition of the theorem is satisfied, and we can say that the sum of the 2 normal distributions is also a normal distribution.

**Q10: Generate a plot that displays p(x)~N(6,1.52), prior p(μ)~N(4,0.82), and posterior p(μ|X)~N(μn,σ2n) for n=20 sample points. What are the values for μn and σ2n?**

**Answer:**

After running the code below, the output is

('Mean of the posterior distibution is', 5.56336966620394, 'and variance is', 0.09568106312292358)



**Code: (Python3)**

**import** numpy **as** np

**from** matplotlib **import** pyplot **as** plt

**import** scipy**.**stats **as** st

samples **=** 20

x **=** np**.**linspace**(**0**,** 10**,** samples**)**

##Prior

mean\_0 **=** 4

sd\_0 **=** 0.8

prior\_dist **=** st**.**norm**(**mean\_0**,** sd\_0**).**pdf**(**x**)**

##Sample

mean\_x **=** 6

sd\_x **=** 1.5

sample\_dist **=** dist **=** st**.**norm**(**mean\_x**,** sd\_x**).**pdf**(**x**)**

##Posterior

x\_t **=** st**.**norm**(**mean\_x**,**sd\_x**).**rvs**(**samples**)**

var\_n **=** 1**/((**1**/**sd\_0**\*\***2**)+(**samples**/**sd\_x**\*\***2**))**

mean\_n **=** var\_n **\*** **((**mean\_0**/**sd\_0**\*\***2**)+(**np**.**mean**(**x\_t**)\***samples**/**sd\_x**\*\***2**))**

**print(**"Mean of the posterior distibution is"**,**mean\_n**,**"and variance is"**,**var\_n**)**

posterior\_dist **=** st**.**norm**(**mean\_n**,** np**.**sqrt**(**var\_n**)).**pdf**(**x**)**

##Plot

plt**.**plot**(**x**,**prior\_dist**,**"b-"**,**label**=**'Prior'**)**

plt**.**plot**(**x**,**sample\_dist**,**"g-"**,**label**=**'Sample'**)**

plt**.**plot**(**x**,**posterior\_dist**,**"r-"**,**label**=**'Posterior'**)**

plt**.**legend**(**loc**=**'upper left'**)**

plt**.**title**(**'Probability Density Plot'**)**

plt**.**ylabel**(**'Probability Density'**)**

plt**.**xlabel**(**'X'**)**

plt**.**show**()**